INTERNATIONAL CONFERENCE ON COMPUTATIONAL INTELLIGENCE AND INNOVATIVE APPLICATIONS (CIIA2022)

The 2nd International Conference on Advanced Technology and Sustainable Development – 2022

PROCEEDING – ICATSD2022

November 24-26, 2022

Organized by
Industrial University of Ho Chi Minh City & Eastern International University
Vietnam

International Conference on Computational Intelligence and Innovative Applications (CIIA2022)

INDUSTRIAL UNIVERSITY OF HO CHI MINH CITY
PUBLISHING HOUSE

THE 2nd INTERNATIONAL CONFERENCE ON ADVANCED TECHNOLOGY AND SUSTAINABLE DEVELOPMENT 2022 (ICATSD 2022)

INTERNATIONAL CONFERENCE ON COMPUTATIONAL INTELLIGENCE AND INNOVATIVE APPLICATIONS (CIIA2022)

INDUSTRIAL UNIVERSITY OF HO CHI MINH CITY PUBLISHING HOUSE

THE 2nd INTERNATIONAL CONFERENCE ON ADVANCED TECHNOLOGY AND SUSTAINABLE DEVELOPMENT – 2022 (ICATSD 2022)

ORGANIZING COMMITTEE

- *Dr. Phan Hong Hai, Industrial University of Ho Chi Minh City, Vietnam*
- *Prof. Dr. Le Van Tan, Industrial University of Ho Chi Minh City, Vietnam*
- *Assoc. Prof. Dr. Dam Sao Mai, Industrial University of Ho Chi Minh City, Vietnam*
- *Dr. Nguyen Xuan Hong, Industrial University of Ho Chi Minh City, Vietnam*
- *Assoc. Prof. Dr. Trinh Ngoc Nam, Industrial University of Ho Chi Minh City, Vietnam*
- *Dr. Le Ngoc Son, Industrial University of Ho Chi Minh City, Vietnam*
- *Mr. Nguyen Dinh Hien, Industrial University of Ho Chi Minh City, Vietnam*
- *Mr. Nguyen Tan Loi, Eastern International University, Vietnam*
- *Dr. Nguyen Van Phuc, Eastern International University, Vietnam*
- *Dr. Ngo Minh Duc, Eastern International University, Vietnam.*
- *Mr. Nguyen Manh Cuong, Agency for Southern Affairs of Ministry of Science and Technology, Vietnam*
- *Dr. Le Thanh Minh, Department of Science and Technology Ho Chi Minh City, Vietnam*

SCIENTIFIC COMMITTEE

- *Dr. Phan Hong Hai, Industrial University of Ho Chi Minh City, Vietnam*
- *Prof. Dr. Le Van Tan, Industrial University of Ho Chi Minh City, Vietnam*
- *Assoc. Prof. Dr. Dam Sao Mai, Industrial University of Ho Chi Minh City, Vietnam*
- *Assoc. Prof. Dr. Duong Hoai Nghia, Eastern International University, Vietnam*
- *Prof. Dr. Le Tien Thuong, University of Technology, Vietnam*
- *Prof. Dr. Ho Pham Huy Anh, HCMC University of Technology, Vietnam*
- *Prof. Dr. Yongwan Won, Chonnam National University, South Korea*
- *Assoc. Prof. Dr. Nguyen Tan Luy, Industrial University of Ho Chi Minh City, Vietnam*
- *Assoc. Prof. Dr. Huynh Trung Hieu, Industrial University of Ho Chi Minh City, Vietnam*
- *Prof. Dr. Bernard Tychon, Liege University, Belgium*
- *Prof. Dr. Petra Schneider, Applied Sciences University Magdeburg-Stendal, Germany*
- *Prof. Dr. Byeong-Kyu Lee, University of Ulsan, South Korea*
- *Prof. Dr. Tomio Miwa, Nagoya University, Japan*
- *Prof. Dr. Ping Chung Kuo, National Cheng Kung University, Taiwan*
- *Prof. Dr. Lin Woo Kang, KonKuk University, South Korea*
- *Prof. Dr. Min-Ren Yan, Member of The Editorial Board of Measuring Business Excellence*
- *Prof. Dr. Tony Chieh-Tse Hou, NDHU National Dong Hwa University, Taiwan*
- *Assoc. Prof. Dr. Tran Thi Giang Tan, University of Economics Ho Chi Minh City, Vietnam*
- *Prof. Dr. Yuan Ju Long, Zhejiang Uni. of Technology, China*
- *Prof. Dr. Wang Dung An, National Chung Hsing Uni., Taiwan*
- *Prof. Dr. Choi Seung Bok, The State University of New York, Korea*
- *Prof. Dr. Lyu Bing Hai, Zhejiang University of Technology, China*
- *Dr. Miang Koon Ng Judy, HBU-UCLan, Associate Fellow Higher Education Academy, UK*
- *Dr. Tran Thi-Lan, Department of East Asian Language and Literature, National University of Kaohsiung, Taiwan*
- *Prof. Dr. Nguyen Van-Cong, Industrial University of Ho Chi Minh City, Vietnam*
- *Assoc. Prof. Dr. John Luke Gallup, Portland State University, USA*
- *Prof. Dr. Fabien De Geuser, Dean at CFVG, French Vietnamese School of Management, ESCP Europe*

2022 INTERNATIONAL CONFERENCE ON COMPUTATIONAL INTELLIGENCE AND INNOVATIVE APPLICATIONS (CIIA2022)

GENERAL CHAIRS

- *Prof. Dr. Le Tien Thuong, Ho Chi Minh City University of Technology, VNU-HCM, Vietnam*
- *Dr. Phan Hong Hai, Industrial University of Ho Chi Minh City, Vietnam*

PROGRAM COMMITTEE CHAIRS

- *Prof. Dr. Ho Pham Huy Anh, Ho Chi Minh City University of Technology, VNU-HCM, Vietnam*
- *Prof. Dr. Yongwan Won, Chonnam National University, South Korea*
- *Assoc. Prof. Dr. Nguyen Tan Luy, Industrial University of Ho Chi Minh City, Vietnam*
- *Assoc. Prof. Dr. Hieu Trung Huynh, Industrial University of Ho Chi Minh City, Vietnam*
- *Dr. Tu Nguyen, Sandia National Laboratories, United States*

STEERING COMMITTEE

- *Prof. Dr. Narayan C. Debnath, Eastern International University, Vietnam*
- *Assoc. Prof. Dr. Manuel Clavel, Vietnamese-German University, Vietnam*
- *Prof. Dr. Genci Capi, Hosei University, Japan*
- *Dr. Marco Abbiati, Embassy of Italy in Vietnam*
- *Assoc. Prof. Dr. Ton Do, Nazarbayev University, Kazakhstan*

PROGRAM COMMITTEE

- *Prof. Narayan C. Debnath, Eastern International University, Vietnam*
- *Assoc. Prof. Dr. Linh Nguyen, University of Idaho, United States of America*
- *Assoc. Prof. Dr Cheng-Hu Chen, National Ilan University, Taiwan*
- *Assoc. Prof. Dr. Pham The Bao, Sai Gon University, Vietnam*
- *Dr. Kha-Tu Huynh, International University, VNU-HCM, Vietnam*
- *Assoc. Prof. Dr. Manuel Clavel, Eastern International University, Vietnam*
- *Dr. Marco Abbiati, Embassy of Italy in Vietnam*
- *Assoc. Prof. Dr. Ton Do, Nazarbayev University, Kazakhstan*
- *Dr. Le Lam Son, Ho Chi Minh City University of Technology, VNU-HCM, Vietnam*
- *Assoc. Prof. Dr. Tran Khanh Dang, Ho Chi Minh City University of Food Industry, Vietnam*
- *Prof. Dr. Genci Capi, Hosei University, Japan*
- *Dr. Tan Loc Nguyen, Chonnam National University, South Korea*
- *Dr. Le Thanh Sach, Ho Chi Minh City University of Technology, VNU-HCM, Vietnam*
- *Dr. Dung H. Duong, University of Wollongong, Australia*
- *Dr. Trinh Thanh Son, Sungkyunkwan University, South Korea*
- *Dr. Quang-Vinh Dang, Industrial University of Ho Chi Minh City, Vietnam*
- *Assoc. Prof. Dr. Quan Thanh Tho, Ho Chi Minh City University of Technology, VNU-HCM, Vietnam*
- *Dr. Phuc Dang, Industrial University of Ho Chi Minh City, Vietnam*
- *Dr. Nguyen Ngoc Son, Industrial University of Ho Chi Minh City, Vietnam*
- *Dr. Minh-Khai Nguyen, Queensland University of Technology, Australia*
- *Dr. Le Van Dai, Industrial University of Ho Chi Minh City, Vietnam*
- *Assoc. Prof. Dr. Nguyen Van Nho, Ho Chi Minh City University of Technology, VNU-HCM, Vietnam*
- *Dr. Nguyen Vinh Hao, Ho Chi Minh City University of Technology, VNU-HCM, Vietnam*
- *Dr. Truong Quang Vinh, Ho Chi Minh City University of Technology, VNU-HCM, Vietnam*
- *Dr. Phan Van Vinh, Eastern International University, Vietnam*

- *Dr. Huynh Tan Phuoc, Eastern International University, Vietnam*
- *Dr. Shreya Banerjee, Eastern International University, Vietnam*
- *Dr. Tran Van Luan, Eastern International University, Vietnam*
- *Dr. Nguyen Hoang Sy, Eastern International University, Vietnam*
- *Dr. Nguyen Dinh Vinh, Eastern International University, Vietnam*
- *Prof. Dr. Kurokawa Hiroaki - Computational Intelligence, Soft Computing, Artificial Neural Networks*

LOCAL ORGANIZING COMMITTEE

- *Assoc. Prof. Dr. Trinh Ngoc Nam, Industrial University of Ho Chi Minh City, Vietnam*
- *Dr. Phuc Dang, Industrial University of Ho Chi Minh City, Vietnam*
- *Dr. Nguyen Ngoc Son, Industrial University of Ho Chi Minh City, Vietnam*
- *Assoc. Prof. Dr. Nguyen Hoa, Sai Gon University, Vietnam*

CONTENTS

ICATSD2F.101

SEMANTIC VEHICLE SEGMENTATION FROM AERIAL IMAGES USING DEEP LEARNING

VAN LUAN TRAN[1], HUEI-YUNG LIN[2], MANH-HUNG HA[3]

[1]*School of Engineering, Eastern International University, Vietnam*

[2]*Department of Computer Science and Information Engineering, National Taipei University of Technology, Taiwan*

[3]*Faculty of Applied Sciences, International School, Vietnam National University, Hanoi*

luan.tran@eiu.edu.vn, lin@ntut.edu.tw, hunghm@vnuis.edu.vn

Abstract. Nowadays, with the increasing vehicle density, aerial images are essential to monitor and regulate traffic. Computer vision and deep learning have contributed significantly to the development of intelligent transportation systems. This paper proposes a technique for semantic vehicle segmentation from aerial images based on a deep convolutional neural network. We trained and tested our dataset for vehicle detection and classification, VAID (Vehicle Aerial Imaging from Drone). In this dataset, the images are collected by a Drone. We evaluated and compared our results of semantic vehicle segmentation on aerial images with other network architectures. The experimental results demonstrated that the proposed method with our dataset could get high semantic vehicle segmentation accuracy with the different traffic environments.

Keywords. Deep Learning, Aerial Image, Semantic Vehicle Segmentation.

1 INTRODUCTION

Recently, the classification of vehicles for traffic is an important field of research in intelligent transportation systems (ITS). Computer vision and deep learning are effective in a variety of applications such as traffic monitoring systems, intelligent traffic surveillance systems, route monitoring, network traffic analysis, etc. Applications from aerial imagery require wider geographical coverage, such as traffic management division to improve transportation mobility [1], [2]. Classification of vehicles in aerial images is essential in building an intelligent traffic surveillance system [3]. In [1], Ammar *et al.* proposed a solution for vehicle detection from aerial images based on deep convolutional neural networks (YOLOv3, YOLOv4, and Faster R-CNN). They reported that YOLOv3 [5] and YOLOv4 [6] evaluated better performance than Faster R-CNN [4] in most configurations. The dataset's attributes and the representativeness of the training images significantly impact on the effectiveness of these methods.

Figure 1: Overview of our vehicle classification on aerial images, the left image is semantic vehicle segmentation, and the right image is vehicle detection on YOLOv4.

The differences between vehicle detection and semantic vehicle segmentation are shown in Figure 1. The target of vehicle detection is to detect the location of all vehicles in images, which included the center of objects in bounding boxes and confidence scores [7]. Semantic vehicle segmentation is to identify the image area of a certain object, a task to classify each pixel in an image from a predefined set of classes. Semantic vehicle segmentation can produce results for pixel-wise extraction that are more precise than vehicle detection [7]. Recently, deep convolutional network advancements made more recently have made it feasible to complete end-to-end semantic vehicle segmentation. For instance, Audebert *et al.* [8] suggested using deep networks like SegNet for semantic segmentation. They can be used to separate even small objects, like vehicles and trucks, and are useful for understanding scenes in remote sensing data. In addition, they demonstrated that by using this simple linked component extraction and high-resolution fractionation, vehicle detection can be achieved with results superior to other methods. They used two difficult remote sensing datasets on the cities of Potsdam and Christchurch under different conditions to represent this precise segmentation. They demonstrated that these high-resolution semantic maps are sufficient to extract object-level boundaries that outperform conventional vehicle recognition techniques using a morphometric approach.

The proximity of objects in the aerial images makes it difficult for most vehicle detection methods to separate individual objects. While in most cases, we need not only to know the exact number of objects but also to present which pixel belongs to the object to determine the oriented objects. These problems prompted us to perform semantic vehicle segmentation. Semantic vehicle segmentation aims to identify each pixel's semantic class and link each one to a real-world representation of a vehicle. The distinction between vehicle detection is demonstrated via semantic segmentation [7]. In addition, deep networks have recently achieved great success in several remote sensing applications using high-resolution images [9], [16], [17]. In this paper, we propose a technique for semantic vehicle segmentation from aerial images based on a deep convolutional neural network. We trained and tested our dataset for vehicle segmentation with four categories of Sedan, Bus, Truck, and Trailer. We evaluated and compared our results of semantic vehicle segmentation in aerial images with other network architectures. The experimental results demonstrated the proposed method with our dataset can get high semantic vehicle segmentation accuracy with the different traffic environments.

2 SEMANTIC VEHICLE SEGMENTATION

For the semantic vehicle segmentation from aerial images, this work is specifically concerned with the method to extract accurate object masks. It is used to identify the locations and density of vehicles on an aerial image. ResNet [14] is one of the well-known fully convolutional networks currently accessible in research in this area. Additionally, it was the first architecture to include batch normalization. The accuracy of prior network models is frequently enhanced by adding more convolutional neural network layers. The original information may be lost if there are too many layers of depth. Some researchers use shortcut connections to maintain the information from the earlier levels to the later layers to solve this difficulty [10]. The convolution block has two convolution branches, one of which does a 1x1 convolution before being immediately joined to the other branch. The identity block immediately adds the value of that branch to the value without using the 1x1 convolution.

In our semantic vehicle segmentation network as presented in Figure 2, a deep convolutional network is constructed for our semantic vehicle segmentation network. It has a 512x512x3 input layer and five stages to downscale the features. In the first step, 64 output filters in the shapes of (3, 3), stride (2, 2), BatchNorm, and MaxPooling are convoluted by (2, 2). A convolutional block with three filters of varying sizes is used in the second step (64, 64, 256). This level uses the filters of four identity blocks (64, 64, 256). A convolutional block with three filters of size comprises the third step (128, 128, 512). Three identification blocks with filters are present at this level (128, 128, 512). A convolutional block with three filters of the size of the fourth stage was used (256, 256, 1024). One identification block with filters is present at this level (256, 256, 1024). A convolutional block with three filters of sizes (512, 512, 2048) is used in the final step. There are two identification blocks with filters at this level (512, 512, 2048). An Upsample module is attached to the last stage's output. Two 3x3 convolutions with 512 output filters make up the system. The 2x2 kernel and stride are then utilized to do a conv2Dtranspose with 256 output filters (2, 2). In the fourth stage, the output of the Upsample module and the output of the identity block are combined. In the third

stage, it is upsampled and concatenated with the output of the identity block. The output of the identity block in the second stage is concatenated with the output of the subsequent output before being upsampled once again. The convolutional layer with softmax activation is coupled to the output of the Upsample module. The output in the final layer is 512x512x4, which has the same resolution as the input image.
For model training, we used a categorical cross-entropy loss function. The formula for the categorical cross-entropy (CCE) is given by

$$CCE = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} y_{cp}\,.\,log(y_{cg}) + (1 - y_{cp}).log(1 - y_{cg})$$

where y_{cp} and y_{cg} denote the predicted and ground-truth values of the object in the image, respectively. N is the number of observations, and C is the number of categories. A softmax activation function cooperates with the categorical cross-entropy cost function. For multi-class classification, it calculates the loss between network predictions and target values.

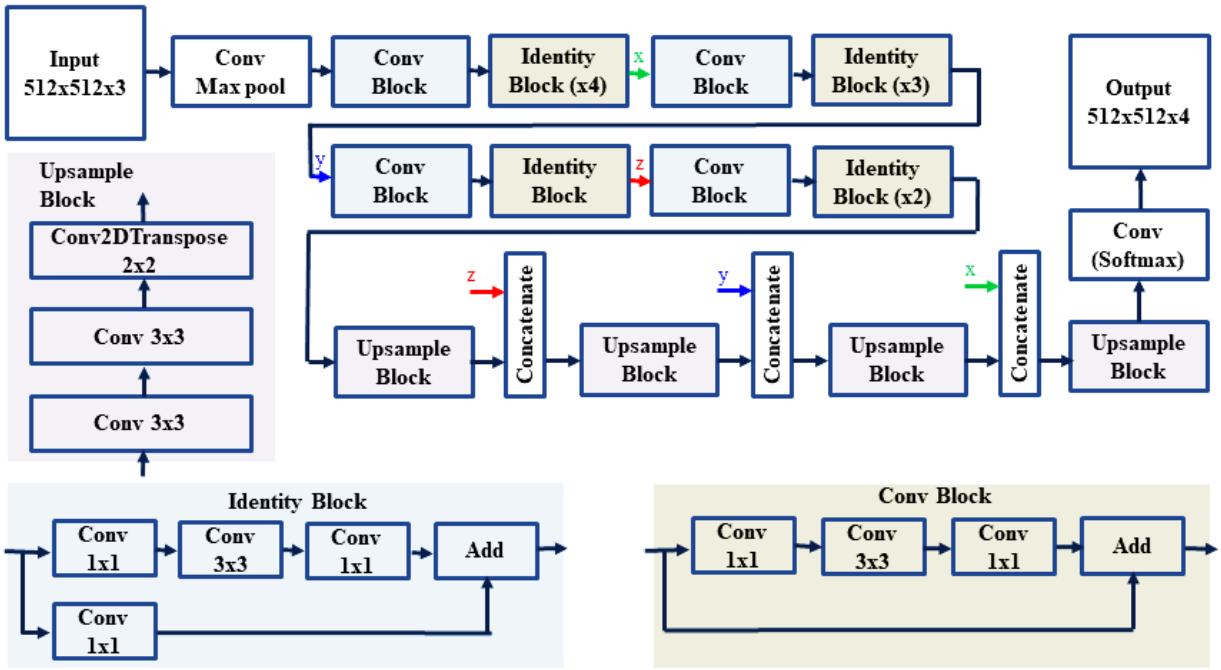


Figure 2: The end-to-end architecture of our semantic vehicle segmentation network. This architecture simulates our network model based on three main modules, which are Conv Block, Identity Block, and Upsample Block.

3 EXPERIMENTAL RESULTS

We collected a VAID dataset (Vehicle Aerial Imaging from Drone), a novel vehicle identification dataset based on aerial images taken by a drone, then trained and tested our technique on this dataset [15]. They were gathered from various locations in Taiwan using roughly 6000 aerial images taken at various viewing angles and lighting levels. The images were captured in JPG format with a resolution of 1137x640 pixels. Seven different types of vehicles are included in the VAID dataset: sedans, minibuses, trucks, pick-up trucks, buses, cement trucks, and trailers. It includes 4,600 images for training, 700 images for verification, and 700 images for testing. In this semantic vehicle segmentation, the training images were labeled with four categories of Sedan, Bus, Truck, and Trailer.
We trained our network on a computer with 32GB RAM and an Intel i7-7700 CPU running at 3.6 GHz, The GPU is an NVIDIA GeForce RTX 3060 with 12GB RAM. In the network training, the initial learning rate is set as 3e-4 for Adam, and the number of epochs and batch size are set to 100 and 2, respectively. In Figure 3, we presented some of our experimental results on aerial images using our semantic vehicle segmentation network. We obtained a high accuracy of the semantic vehicle segmentation. We also

performed the vehicle detection based on YOLOv4 to compare as reported in Figure 3. In this dataset, the density of Sedan cars is more than other objects. So, the accuracy of Sedan detection is higher than others.

Figure 3: Our experimental results from our dataset, the left images are our semantic vehicle segmentation, and the right image is our vehicle detection based on YOLOv4.

We evaluated and compared our semantic vehicle segmentation on this dataset with other methods as reported in Table 1. This table tabulates the evaluation results of Ours, Unet++ [12], Unet [11], and PSPNet [13] on 700 images for testing. We can achieve the high accuracy by mean intersection over union (mIoU) of Sedan, Bus, Truck, and Trailer with 94.3%, 83.1%, 85.4%, and 86.7, respectively. The average mIoUs of Ours, Unet++, UNet, and PSPNet are 87.7%, 86.9% and 83.7%, respectively. Our network model's mIoU of 87.7% is higher than the other approaches, such as Unet++, UNet, and PSPNet.

Table 1: Evaluation and comparison of our semantic vehicle segmentation and the evaluation results of Unet++, UNet, and PSPNet by mIoU.

Method	Sedan	Bus	Truck	Trailer	mIoU
PSPNet	87.9%	77.8%	79.0%	80.4%	81.3%
UNet	90.9%	81.2%	80.5%	82.0%	83.7%
Unet++	93.8%	83.2%	84.9%	85.8%	86.9%
Ours	94.3%	83.1%	85.4%	86.7%	87.7%

4 CONCLUSIONS

In this paper, we provided a method for building a semantic vehicle segmentation network from aerial images using a modified version of the ResNet framework. In order to test and compare the outcomes of semantic vehicle segmentation in aerial images, we employed training datasets and well-known network designs. The results show that under varied traffic and road situations, the suggested approach and dataset can provide high semantic vehicle segmentation rates. The proposed method provides better results compared to different techniques in the semantic vehicle segmentation from aerial images. Semantic vehicle segmentation will be incorporated into the traffic management system in the future to monitor parking lots and traffic flow.

ACKNOWLEDGMENT

The support of this work in part by the Ministry of Science and Technology of Taiwan under Grant MOST 106-2221-E-194-004, the Advanced Institute of Manufacturing with High-tech Innovations (AIM-HI) in Taiwan, and Eastern International University, Binh Duong Province, Vietnam is gratefully acknowledged.

REFERENCES

[1] A. Ammar, A. Koubaa, M. Ahmed, A. Saad, and B. Benjdira, "Vehicle detection from aerial images using deep learning: A comparative study," Electronics, vol. 10, no. 7, 2021. [Online]. Available: https://www.mdpi.com/2079-9292/10/7/820.

[2] Y. Tan, Y. Xu, S. Das, and A. Z. Chaudhry, "Vehicle detection and classification in aerial imagery," in 2018 IEEE International Conference on Image Processing, ICIP 2018, Athens, Greece, October 7-10, 2018. IEEE, 2018, pp. 86–90. [Online]. Available: https://doi.org/10.1109/ICIP.2018.8451709.

[3] S. Javadi, M. Dahl, and M. I. Pettersson, "Vehicle detection in aerial images based on 3d depth maps and deep neural networks," IEEE Access, vol. 9, pp. 8381–8391, 2021. [Online]. Available: https://doi.org/10.1109/ACCESS.2021.3049741.

[4] S. Ren, K. He, R. B. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," CoRR, vol. abs/1506.01497, 2015. [Online]. Available: http://arxiv.org/abs/1506.01497.

[5] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," CoRR, vol. abs/1804.02767, 2018. [Online]. Available: http://arxiv.org/abs/1804.02767.

[6] A. Bochkovskiy, C. Wang, and H. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," CoRR, vol. abs/2004.10934, 2020. [Online]. Available: https://arxiv.org/abs/2004.10934.

[7] L. Mou and X. X. Zhu, "Vehicle instance segmentation from aerial image and video using a multi-task learning residual fully convolutional network," CoRR, vol. abs/1805.10485, 2018. [Online]. Available: http://arxiv.org/abs/1805.10485.

[8] N. Audebert, B. L. Saux, and S. Lefevre, "Segment-before-detect: `Vehicle detection and classification through semantic segmentation of aerial images," Remote. Sens., vol. 9, no. 4, p. 368, 2017. [Online]. Available: https://doi.org/10.3390/rs9040368.

[9] N. Merkle, S. M. Azimi, S. Pless, and F. Kurz, "Semantic vehicle segmentation in very high resolution multispectral aerial images using deep neural networks," in 2019 IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2019, Yokohama, Japan, July 28- August 2, 2019. IEEE, 2019, pp. 5045–5048. [Online]. Available: https://doi.org/10.1109/IGARSS.2019.8898513.

[10] P.-S. Wang, Y. Liu, and X. Tong, "Deep octree-based cnns with outputguided skip connections for 3d shape and scene completion," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern

Recognition Workshops, 2020, pp. 266–267.

[11] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," CoRR, vol. abs/1505.04597, 2015. [Online]. Available: http://arxiv.org/abs/1505.04597.

[12] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," CoRR, vol. abs/1807.10165, 2018. [Online]. Available: http://arxiv.org/abs/1807.10165.

[13] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," CoRR, vol. abs/1612.01105, 2016. [Online]. Available: http://arxiv.org/abs/1612.01105.

[14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016. IEEE Computer Society, 2016, pp. 770–778. [Online]. Available: https://doi.org/10.1109/CVPR.2016.90.

[15] C. Li and H. Lin, "Vehicle detection and classification in aerial images using convolutional neural networks," in Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and
Computer Graphics Theory and Applications, VISIGRAPP 2020, Volume 5: VISAPP, Valletta, Malta, February 27-29, 2020, G. M. Farinella, P. Radeva, and J. Braz, Eds. SCITEPRESS, 2020, pp. 775–782.
[Online]. Available: https://doi.org/10.5220/0008941707750782.

[16] M. Hofmarcher, T. Unterthiner, J. Arjona-Medina, G. Klambauer, S. Hochreiter, and B. Nessler, "Visual Scene Understanding for Autonomous Driving Using Semantic Segmentation", pages 285–296. 09-2019.

[17] D. Marmanis, J. Wegner, S. Galliani, K. Schindler, M. Datcu, and U. Stilla, "Semantic segmentation of aerial images with an ensemble of cnns", ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, III-3:473–480, 06 2016.

ICATSD2F.102

ESTIMATING LASER IRRADIATION COORDINATES FROM TEMPERATURE DISTRIBUTION USING CONVOLUTIONAL NEURAL NETWORK

MIKI NAKAONE[1], TOMOMASA OHKUBO[1*], YUKI UENO[1], KEN GOTO[2], YUTAKA KAGAWA[3]

[1]*Graduate School of Engineering, Tokyo University of Technology*

[2]*Institute of Space and Astronautical Science, Japan Aerospace Exploration Agency*

[3]*Katayanagi Institute, Tokyo University of Technology*

**ookubotmms@stf.teu.ac.jp*

Abstract. An accelerated heating test system named selective laser thermoregulation (SLT) was developed to clarify the mechanical properties at high temperatures of SiC/SiC ceramic composites. The SLT system controls the temperature distribution by scanning the laser. To achieve an optional temperature distribution, we must determine the irradiation conditions of the laser, but they are highly flexible. Therefore, we need AI to estimate irradiation conditions that reproduce the required temperature distribution. In this study, we developed AI to estimate irradiation coordinates from the required temperature distribution using a fully connected neural network (FCNN) and a convolutional neural network (CNN). Both AI was able to learn the relationship between the temperature distribution and the irradiation coordinates in a short time, the learning time of the FCNN was about 11 minutes and the CNN was about 18 minutes. The mean absolute error (MAE) of the FCNN was 0.22 mm when the AI predicted the x coordinate and that of y coordinate was 0.19 mm. The MAE of the CNN was 0.22 mm when the AI predicted the x coordinate and that of y coordinate was 0.20 mm. According to these results, the FCNN and CNN were able to predict with almost equally high accuracy.

Keywords. ceramic matrix composites (CMCs), deep learning, fiber laser, Galvano scanner, selective laser thermoregulation (SLT) system

1 INTRODUCTION

Aircraft are essential for the movement of people and trade in the current era of globalization. However, aircraft consume a lot of energy. The efficiency of aircraft engines can be improved by using materials with high specific strength and high heat resistance. SiC/SiC ceramic composites (CMCs) which have these characteristics are expected to replace Ni-based alloys currently used in aircraft engines [1] [2]. To actually use SiC/SiC CMCs in aircraft engines, it is necessary to perform tests under high-temperature environments to clarify their mechanical properties.

Several studies have been carried out to measure the physical properties of SiC/SiC CMC high-temperature environments [3]-[6]. It is essential to perform accelerated heating tests, since actual aircraft engines are repeatedly heated and cooled many times for a long period of time. Therefore, accelerated heating tests have been performed using a high-power laser.

A heating test method using a CO_2 laser was reported by Appleby et al. in 2015 [7]. It is difficult to simulate the actual environment of an aircraft engine in a high temperature and humidity environment because CO_2 lasers are easily absorbed by water vapor.

Accelerated heating tests using a fiber laser have been reported by Whitlow et al. in 2019 [8]. A fiber laser with a wavelength of 1.07 μm of transmittance to water vapor is much higher than a CO_2 laser [9]. In this system, the laser beam is shaped into the shape of a target by optics before irradiation. Therefore, it is difficult to control a temperature distribution with this method dynamically.

We developed an accelerated heating test method named “Selective Laser Thermoregulation (SLT)” [10].

The SLT method can scan the laser by using a Galvano mirror. This system enables active control of a temperature distribution of a target. An optional temperature distribution can be achieved by changing many parameters such as laser power, irradiation coordinates, and scanning speed. However, it is difficult to determine the laser irradiation conditions that achieve the optional temperature distribution because the laser irradiation conditions have highly flexible and a large number of patterns are possible.
Therefore, we developed AI to estimate the laser power that reproduces the required temperature distribution [11]. In the previous study, we chose a fully connected neural network (FCNN) for AI which estimated the laser power from the required temperature distribution.
In this study, we developed AI to estimate the irradiation coordinates that reproduce the required temperature distribution. We chose and compared FCNN and a convolutional neural network (CNN) for AI which estimated the irradiation coordinates from the required temperature distribution because CNN is used for image recognition and classification and the temperature distribution is also 2-dimensional data.

2 GENERATING A DATASET BY NUMERICAL SIMULATION

We generated a dataset for the learning and evaluation of the AI. The dataset consists of a pair of the temperature distribution and the irradiation conditions which achieve the temperature distribution. We generated the temperature distribution by the same numerical simulation code used in the previous study [11]. It solved the heat balance of the input laser power, thermal conduction, heat transfer to the air, and radiation using the finite difference method in two dimensions.
The temperature distributions were obtained by irradiating a single point by the laser onto a $40\text{mm} \times 40\ \text{mm} \times 3\ \text{mm}$ (thickness) target. The calculated results are 2-dimensional data of 41 rows and 41 columns. The temperature distribution in the thickness direction is averaged. We performed a total of 16807 numerical simulations for generating the dataset, varying each of the five irradiation conditions as described below.

- x coordinate [mm] -12 mm to 12 mm in 4 mm increments
- y coordinate [mm] -12 mm to 12 mm in 4 mm increments
- Laser power [W] 400 W to 4000 W in 600 W increments
- Beam radius [mm] 2 mm to 8 mm in 1 mm increments
- Irradiation time [s] 0.2 mm to 0.8 mm in 0.1 mm increments

We divided 16807 data into three types: training data, validation data, and test data. 10757 data (64%) were used as the training data which is used in chapter 0, 2689 data (16%) were used as the validation data, and 3361 data (20%) were used as the test data which is used in chapter 0.
The AI learned the relationship between the laser irradiation conditions and the temperature distribution using the training data. During this learning process, the AI changed parameters such as weights and bias in the neural network to reduce the mean squared error (MSE) between the exact conditions and predicted conditions. Then the AI was able to predict the irradiation coordinates close to the exact irradiation coordinates.
The validation data was used to adjust hyperparameters. The hyperparameters were parameters determined by the AI developers, such as the number of neurons, the number of layers, and the learning rate.
After adjusting each parameter using training and validation data, we finally evaluated the generalization of the AI by using test data that is not used for learning. Even if AI can predict irradiation coordinates with high accuracy only for training data used for learning, it may not be able to predict them with the same high accuracy for unknown data. Therefore, we validate and evaluate the AI using validation and test data that are not used for learning.

3 NEURAL NETWORKS

We developed two types of the AI which uses FCNN and CNN. Table 1 shows the configuration of the AI with FCNN and Table 2 shows that of CNN.

Table 1: Configuration of the AI using FCNN.

Layer	Output Shape
Input	1681
Dense + ReLU	500
Dense + ReLU	200
Dense + ReLU	50
Dense + Liner	1

Table 2: Configuration of the AI using CNN.

Layer	Output Shape
Input	$41 \times 41 \times 1$
Conv + ReLU	$21 \times 21 \times 8$
Max pooling	$21 \times 21 \times 16$
Conv	$21 \times 21 \times 16$
Batch normalization	$21 \times 21 \times 16$
ReLU	$21 \times 21 \times 16$
Max pooling	$11 \times 11 \times 16$
Flatten	1936
Dense + ReLU	100
Dense + Liner	1

The input of the AI using FCNN should be the 1-dimensional vector. Therefore, the 2-dimensional temperature distribution of 41 rows and 41 columns was flattened to the 1-dimensional vector of 1681 elements. The hidden layers were three fully connected layers (Notated as "Dense" in Table 1 and Table 2). The output layer outputted the irradiation coordinate x or the irradiation coordinate y.

The input of the AI using CNN was the 2-dimensional temperature distribution of 41 rows and 41 columns. The hidden layers were two convolutional layers and pooling layers. The kernel size of the convolutional layers is (3, 3). The kernel size of the pooling layer is (2, 2). The stride is 1 and we used zero padding for both the convolutional and pooling layers. The second hidden layer used batch normalization. The third hidden layer used a fully connected layer. The output layer outputted the irradiation coordinate x or the irradiation coordinate y.

The following four items are common to the AI with FCNN and CNN. Regarding the learning rate, we tried different learning rates, such as 0.0001 and 0.001. When the learning rate was large, AI failed to learn. When the learning rate was small, AI learning took longer. In this study, we chose a learning rate of 0.0005, which allows for relatively fast learning.

- Activation function ReLU function (Output laser only identity function)
- Weight initialization He initialization
- Optimizer Adam (learning rate is 0.0005)
- Loss function Mean squared error (MSE)

A hardware specification of our computer is Intel Core i5 CPU and 8 GB of memory (RAM). GPU was not used.

4 LEARNING

The AI learned relationship between the laser irradiation conditions and the temperature distribution using the training data. Figure 1 and Figure 2 shows the MSE with epochs for x coordinate and y coordinate. One epoch means that the entire training data passed the neural network once.

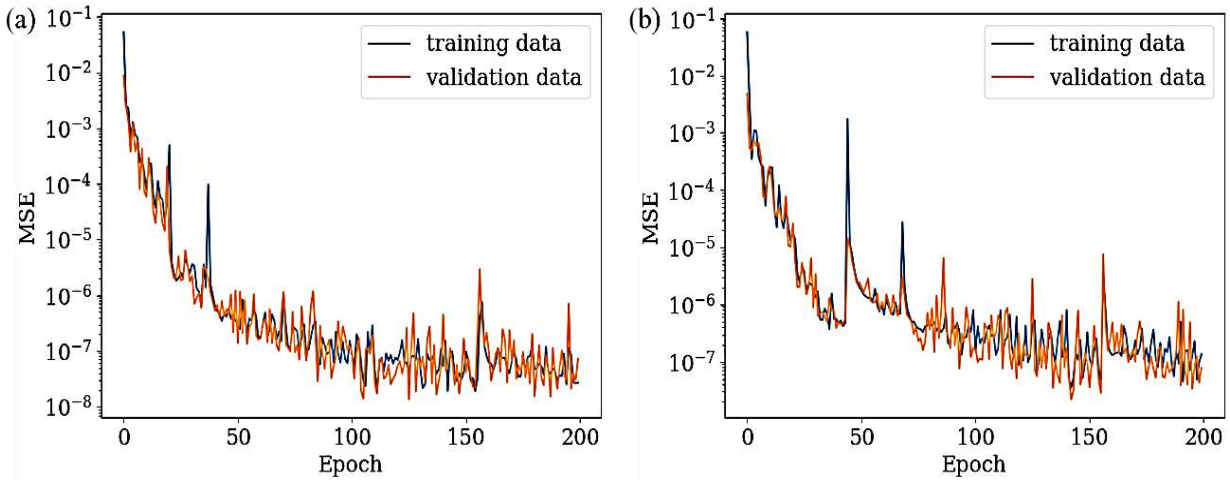


Figure 1: The MSE with epochs for the FCNN. (a) x coordinate, (b) y coordinate.

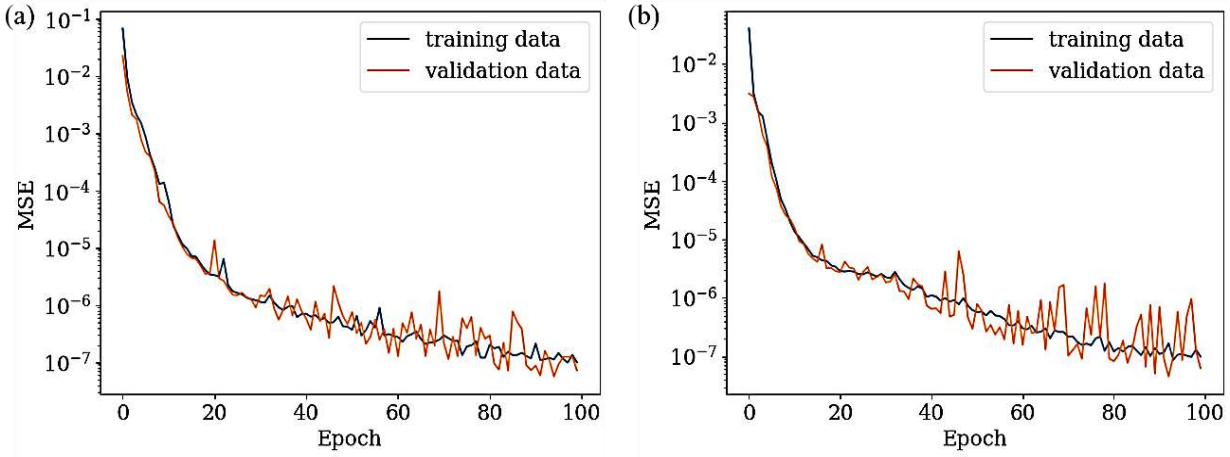


Figure 2: The MSE with epochs for the CNN. (a) x coordinate, (b) y coordinate.

During the learning process, AI is not overfitting because we confirmed that the MSE decreased in the validation data as well as in the training data as shown in Figure 1 and Figure 2.

Figure 1 shows the MSE with epochs for the FCNN. When the number of epochs were between 0 and 50, the MSE reduced significantly. The MSE changed slowly after 50 epochs. Since there is almost no change at around 200 epochs, the learning steps are completed at an epoch of 200. The learning time of the AI is about 11 minutes.

Figure 2 shows the MSE with epochs for the CNN. As same as shown in Figure 1, when the number of epochs was between 0 and 20, the MSE decreased significantly. The MSE changed slowly after 20 epochs. Since there is almost no change at around 100 epochs, the learning steps was completed at an epoch of 100. The learning time of the AI is about 18 minutes.

5 EVALUATIONS

We evaluated the AI using the test data that was not used for learning. We evaluated the relationship between the number of data and the prediction error which is the difference between the exact irradiation

coordinates and the irradiation coordinates predicted by the AI. Figure 3 and Figure 4 shows the evaluated results.

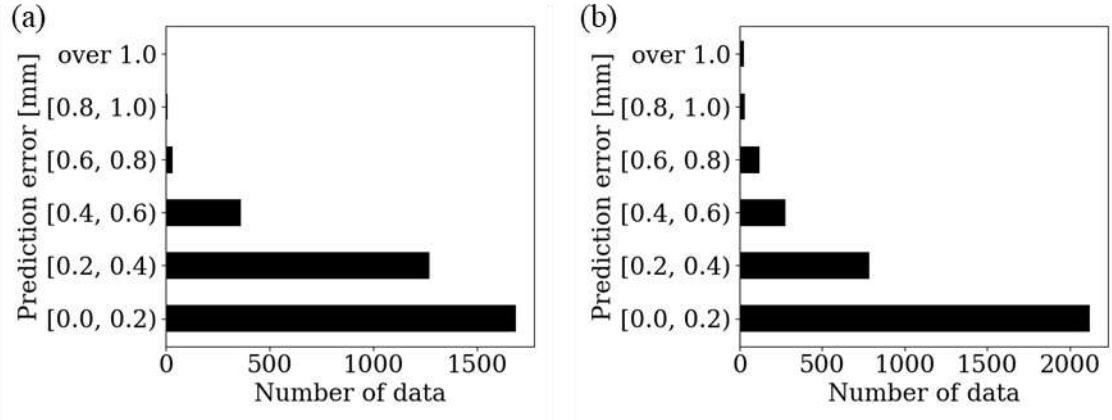


Figure 3: The relationship between the number of data and the prediction error by the AI using FCNN. (a) x coordinate prediction, (b) y coordinate prediction.

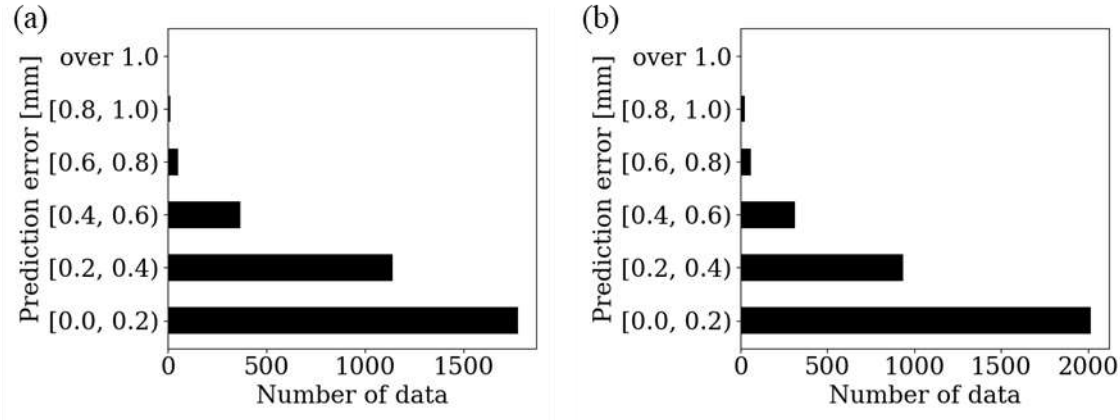


Figure 4: The relationship between the number of data and the prediction error by the AI using CNN. (a) x coordinate prediction, (b) y coordinate prediction.

In addition, we evaluated the AI using the mean absolute error (MAE). The definition of MAE is shown in Equation (1). n, α_i, and β_i are number of data, the i-th predicted value, and the i-th exact value, respectively. The MAE of the FCNN was 0.22 mm when the AI predicted the x coordinate and that of y coordinate was 0.19 mm. The MAE of the CNN was 0.22 mm when the AI predicted the x coordinate and that of y coordinatThee was 0.20 mm.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |\alpha_i - \beta_i| \tag{1}$$

6 DISCUSSIONS

In the chapter 0, learning time of the AI using FCNN is about 11 minutes and using CNN is about 18 minutes. Both AI was able to learn quickly. The FCNN was able to learn about 7 minutes faster than the CNN because FCNN is a simple neural network.

Figure 3 and Figure 4 show that both AI was able to predict the irradiation coordinates with the prediction error of 0.0 mm to 0.2 mm for about half of the test data. The MAE evaluation described in the previous chapter also showed not a large difference in prediction accuracy between the FCNN and the CNN.

The CNN processes the temperature distribution as 2-dimensional data directly, while the FCNN needs conversion of the temperature distribution from 2-dimensional to 1-dimensional. Therefore, we hypothesized that the FCNN are more difficult to predict the irradiation coordinates than the CNN because

of the loss of spatial information. However, the FCNN was able to predict the irradiation coordinates from the temperature distribution with sufficiently high accuracy.
In the case of laser irradiation to a single point, the AI is able to predict the irradiation coordinates from the highest temperature even if the input temperature distribution is converted from 2-dimensional data to 1-dimensional data. However, when the irradiation conditions become more flexible considering laser scanning, the CNN, which can learn spatial information, is expected to predict more accurate parameters than the FCNN.

7 CONCLUSIONS

We developed AI for estimating the laser irradiation coordinates that reproduce the required temperature distribution for laser heating. We chose FCNN and CNN for AI to estimate irradiation coordinates from the required temperature distribution. We clarified the followings.

- Both AI learned the relationship between the temperature distribution and the irradiation coordinates.
- The learning time of the FCNN was about 11 minutes and the CNN was about 18 minutes. The learning time of the FCNN was slightly shorter than that of the CNN.
- The MAE of the FCNN was 0.22 mm when the AI predicted the x coordinate and that of y coordinate was 0.19 mm.
- The MAE of the CNN was 0.22 mm when the AI predicted the x coordinate and that of y coordinate was 0.20 mm.

According to these results, the prediction accuracy was almost the same between the FCNN and the CNN and FCNN learns a little faster than CNN.
In this study, the AI estimated the irradiation coordinates from the temperature distribution generated by the laser irradiation to a single point. When a simple temperature distribution was the target of the prediction, even the FCNN was able to predict the irradiation coordinate with sufficient accuracy.
Although the learning time of the FCNN was slightly shorter than that of the CNN, we found that the FCNN are able to estimate as high accuracy as the CNN in the case of such a simple system.

ACKNOWLEDGMENT

This paper is based on results obtained from a project, JPNP20004, subsidized by the New Energy and Industrial Technology Development Organization (NEDO).

REFERENCES

[1] G.S. Corman and K.L. Luthra, *Handbook of Ceramic Composites*, ed. By N.P. Bansal, Boston, 2005, p99.

[2] M. Roode, J. Price, J. Kimmel, N. Miriyala, D. Leroux, A. Fahme, and K. Smith, Ceramic Matrix Composite Combustor Liners: A Summary of Field Evaluations, *J. Eng. Gas Turbines Power*, vol. 129, 1, pp. 21-30, 2007.

[3] R. J. Young, A. B. L. Broadbridge, and C.-L. So, Analysis of SiC fibres and composites using Raman microscopy, *J. Microscopy*, vol. 196, 2, pp. 257-265, 1999.

[4] Y. L. Dong, H. Kakisawa, and Y. Kagawa, Optical system for microscopic observation and strain measurement at high temperature, *Meas. Sci. Technol.*, vol. 25, no. 2, 025002, 2014.

[5] A. Haboub, H. A. Bale, J. R. Nasiatka, B. N. Cox, D.B. Marshall, R. O. Ritchie, and A. A. MacDowell, Tensile testing of materials at high temperatures above 1700 ℃ with in situ synchrotron X-ray micro-tomography, *Rev. Sci. Instrum.*, vol. 85, no. 8, 083702, 2014.

[6] T. Whitlow, E. Jones, and C. Przybyla, In-situ damage monitoring of a SiC/SiC ceramic matrix composite using acoustic emission and digital image correlation, *Composite Structures*, vol. 158, pp. 245-251, 2016.

[7] M.P. Appleby, D. Zhu, and G. Morscher, Mechanical properties and real-time damage evaluations of environmental barrier coated SiC/SiC CMCs subjected to tensile loading under thermal gradients, *surf Coat technol*, vol. 284, pp. 318-326, 2015.

[8] T. Whitlowa, J. Pitzb, J. Piercea, S. Hawkinsc, A. Samuelc, K. Kollinsd, G. jeffersond, E. Jonesd, J. Vernond, and C. Przybylad, Thermal-mechanical behavior of a SiC/SiC CMC subjected to laser heating, *Composite Structures*, vol. 210, pp. 179-188, 2019.

[9] W. M. Irvine and J. B. Pollack, Infrared optical properties of water and ice spheres, *Icarus*, vol. 8, pp. 324-360, 1968.

[10] H. Koshiji, T. Ohkubo, K. Azato, Y. Kameda, E. Matsunaga, T. Dobashi, N. Shichijo, K. Goto, M. Sato, C. Fujiwara, and Y. Kagawa, Selective Laser Thermoregulation System for Accelerated Degradation Test of SiC/SiC CMCs, *J. Laser Micro/Nanoengineering.*, vol. 15, no. 3, pp. 174-177, 2020.

[11] M. Nakaone, T. Ohkubo, Y. Ueno, K. Goto, and Y. Kagawa, Artificial Intelligence for Estimating Laser Power from Temperature Distribution, *J. Laser Micro/Nanoengineering.*, vol. 16, no. 2, pp. 84-87, 2021.

ICATSD2F.103

ADAPTIVE NONSINGULAR TERMINAL SLIDING MODE CONTROL FOR MANIPULATOR ROBOT

MAI THANG LONG*, TRAN HUU TOAN, TRAN VAN HUNG, TRAN NGOC ANH, NGUYEN HOANG HIEU, NGUYEN THI HONG HA

Faculty of Electronics Technology, Industrial University of Ho Chi Minh City, Vietnam

**maithanglong@iuh.edu.vn, tranhuutoan@iuh.edu.vn, tranvanhung@iuh.edu.vn, tranngocanh@iuh.edu.vn, nguyenhoanghieu@iuh.edu.vn, nguyenthihongha@iuh.edu.vn*

Abstract. This research proposed an improved control method for manipulator robot based nonsingular terminal sliding mode control technique. The improvement is implemented by designing the adaptive updating algorithms for constant control parameters of the traditional sliding mode controller. In addition, the proposed controllers require no prior knowledge of the manipulator robot dynamics based on the Neural approximator. The online updating strategies for control system are derived based on the Lyapunov theorem so the stability and robustness are guaranteed as well as the tracking errors converge to zero within finite time. Finally, the comparative numerical simulation results are addressed to prove the effectiveness of the proposed method.

Keywords. Nonsingular Terminal sliding mode control, manipulator robot, adaptive control, neural networks.

1 INTRODUCTION

For manipulator robot (MR) control, the nonsingular terminal sliding mode (NTSM) control techniques have been explored widely in recent years [1-8]. The NTSM control methods have many advantages for improve the tracking control performances, such as high precision control and states of the sliding surface converge to zero in finite-time [1]. When compared with the linear function-based sliding mode control or terminal sliding mode control, the NTSM control schemes have higher performances, better chattering phenomena decreasing, and more effectiveness in singularity elimination [2]. Based on the merits of the NTSM, the authors in [3] proposed the nonlinear sliding manifold including the proportional derivative (PD)-typed sign functions, without requirement of acceleration and dynamics uncertainties of the MR control system, to achieve higher fixed-time control performances. Unfortunately, this method [3] had some drawbacks as the requirement of nominal parameter of the MR dynamics [6] as well as the fixed control gains. In the [5], M. Van et al. improved the NTSM control by applying the self-turning proportional-derivative-integral (PID) technique into the sliding mode surface to obtain better performances as faster transient response with the finite time converge in considering to chattering phenomenon and tracking error reducing. This work [5] had a sound good reminder for improving the adaptability of the NTSM controller. However, the stability of the proposed control system had not been addressed fully with the fuzzy system and the proportional gain, the parameter has important role in reinforce the tracking errors to stable states, had not been updated adaptively [7, 8]. In this work, in order to improve for the NTSM controllers with the adaptability/robustness and tracking control performances, the proposed NTSM controller has some advanced features as follows. First, the main NTSM controller is proposed with the proportional gain that is online updated adaptively [9, 10]. This is to relax the drawbacks for fixed control gain for NTSM controllers [1-8]. Second, all unknown parameters of the MR dynamics will be estimated by the neural networks (NN) estimators [7] without assumptions of known nominal parameters. Third, the sub-controller, as the robust controller, is designed to compensate for estimating errors, other uncertainties of the MR control system. In addition, this robust controller is added a tanh(.) function, that is to replace for the sign or sat function [1-8], to smooth the control signals and reduce the chattering phenomena.

The next parts are organized as follows. Part 2 presents the design procedures and stability of the proposed control system. Part 3 shows the comparative simulation results. Conclusions are drawn in Part 4.

2 CONTROL ALGORITHMS

2.1 Design procedure

By considering the n-links MR dynamics model as presented in [10], it yields:

$$M(q)\ddot{q} + C(q,\dot{q})\dot{q} + G(q) + F(\dot{q}) + \tau_d = \tau \quad (1)$$

In this work, our goal is to design an adaptive control law based on the NTSM control technique to improve the achievable control performances for the RM system, in the considering of the time-convergence. Therefore, the NTSM surface may be defined as the following forms [1, 2]:

$$S = e + K_\alpha^{-1}\dot{e}^{a/b} \quad (2)$$

where $e(t) = q_d(t) - q(t) \in R^{n\times 1}$ is the tracking position errors, $q_d(t) \in R^{n\times 1}$is the desired positions. $K_\alpha = diag(K_{\alpha 1}, \dots, K_{\alpha n}) \in R^{n\times n}$ is a positive constant matrix. $a, b(a > b)$ are positive odd numbers. Moreover, the derivative of $S(t)$ can be obtained as:

$$\dot{S} = \dot{e} + \frac{a}{b}K_\alpha^{-1}diag(\dot{e}^{\frac{a}{b}-1})\ddot{e} = \dot{e} + \frac{a}{b}K_\alpha^{-1}diag(\dot{e}^{\frac{a}{b}-1})(f + \tau - \tau_d) \quad (3)$$

where $f = \ddot{q}_d - (I + M)\ddot{q} - C\dot{q} - G - F$. The conventional NTSM control input can be defined as:

$$\tau = \tau_d - f - \frac{b}{a}K_\alpha(\dot{e}^{2-\frac{a}{b}}) - Ksign(S) - KS \quad (4)$$

where $K = diag(K_1, \dots, K_n) \in R^{n\times n}$ also is a positive constant matrix. The input (4) can guarantee the control performance for the RM system, as well as the tracking errors can converge in finite time. However, it is difficult to apply due to the unknown/uncertain parameters (τ_d, f) and chattering problems (with the $sign(S)$ function). Therefore, in this research, the improved NTMS control input is designed as:

$$\tau = -\hat{f} - \frac{b}{a}K_\alpha(\dot{e}^{2-\frac{a}{b}}) - \widehat{K}S + \tau_r \quad (5)$$

where $\widehat{K}$ is an adaptive updating value for the K gain, which is the improved feature to increase the adaptability of the NTSM controller. The τ_r is considered as the robust controller (as the $Ksign(S)$ term in (4)) that is used to compensate for the RM control system's disturbances and other uncertainties. The$\hat{f}$ is the estimator for the unknown f function. This estimator is applied by the NN technique [10, 11]. We note that, as the detail in [10], the inevitable NN errors can be considered as:

$$\tilde{f} = f - \hat{f} = (W^{*T} - \widehat{W}^T)\phi + \varepsilon = \widetilde{W}^T\phi + \varepsilon \quad (6)$$

where the $W^{*T} = [w_1^*, w_2^*, \dots, w_h^*] \in R^{n\times h}$ is the optimal weight matrix of NN estimator. $h \in N$ is the number of NN hidden nodes. The $\widehat{W}^T \in R^{n\times h}$ is the approximation value of W^{*T}. The $\phi(x) = [\phi_1, \phi_2, \dots, \phi_h]^T \in R^{h\times 1}$ is the Gaussian function, $x = [\ddot{q}_d, \ddot{q}, \dot{q}, q]^T$ is the NN inputs. The $\varepsilon \in R^{n\times 1}$ is the NN approximator error. From the defined result in (6) and (5), the derivative of $S(t)$ can be rewritten as:

$$\dot{S} = -\frac{a}{b}K_\alpha^{-1}diag(\dot{e}^{\frac{a}{b}-1})KS + \frac{a}{b}K_\alpha^{-1}diag(\dot{e}^{\frac{a}{b}-1})(\widetilde{W}^T\phi + \widetilde{K}S + \varepsilon - \tau_d + \tau_r) \quad (7)$$

where $\widetilde{K} = K - \widehat{K}$, K can be considered as the optimal control gain. Therefore, by assuming that the ε, τ_d are bounded, and they satisfy with the constraint as:

$$\|\varepsilon - \tau_d\| \le R \quad (8)$$

where R is an uncertain positive parameter. With this analysis, to guarantee the control performance for RM control system with the close-loop control system in (7), as well as improving the tracking performances in the presence of uncertainties, the robust controller is designed as:

$$\tau_r = -\widehat{R}\frac{(K_\alpha^{-1})diag(\dot{e}^{\frac{a}{b}-1})S}{\|S\|\left\|(K_\alpha^{-1})diag(\dot{e}^{\frac{a}{b}-1})\right\|} - c\tanh(S) \quad (9)$$

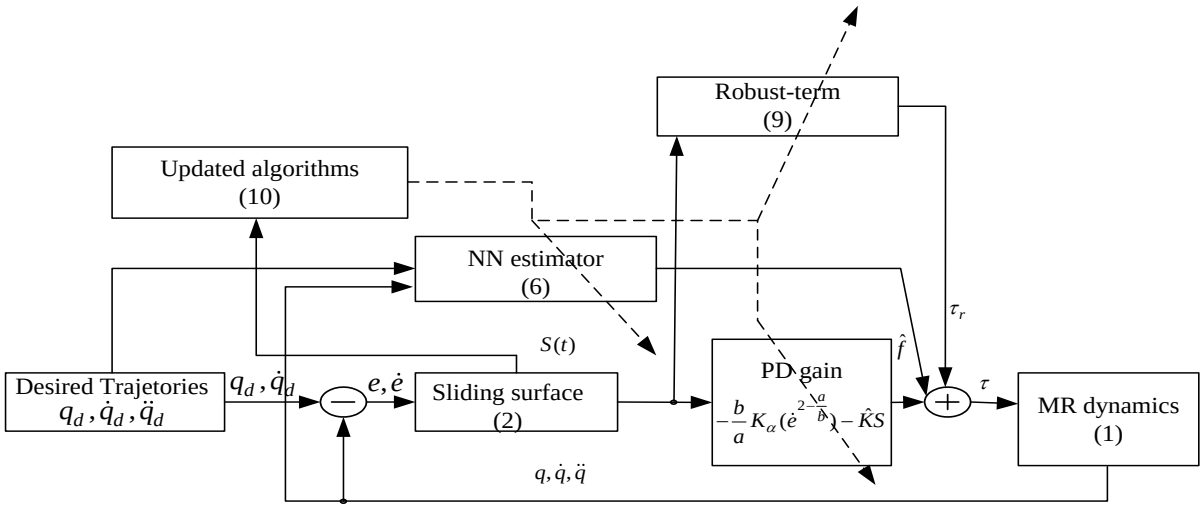


Figure 1: The proposed NTSM control system

where $\hat{R}$ will be adaptive updated to relax requirement of the uncertain R. The tanh (.) is the hyperbolic tangent sigmoid function, $tanh(S) = \frac{e^S - e^{-S}}{e^S + e^{-S}}$ [10], c is a designed positive constant.

This work focuses on the improvement of the NTSM controller for the RM system, such as the adaptability and tracking performances without requiring of the RM control system's information. The proposed NTSM control system is presented in the Fig. 1. Thus, the adaptive online updating algorithms for the proposed NTSM control system are proposed as:

$$\begin{cases} \dot{\hat{W}} = \delta_w \frac{a}{b} K_\alpha^{-1} diag(\dot{e}^{\frac{a}{b}-1}) \phi S^T \\ \dot{\hat{K}} = \delta_k \frac{a}{b} K_\alpha^{-1} diag(\dot{e}^{\frac{a}{b}-1}) S S^T \\ \dot{\hat{R}} = \delta_r \frac{a}{b} \|S\| \left\| (K_\alpha^{-1}) diag(\dot{e}^{\frac{a}{b}-1}) \right\| \end{cases} \tag{10}$$

where δ_w, δ_k are positive diagonal constant matrices, and $\delta_r$is positive constant parameter.

2.2 Stability analysis

Theorem: With the RM dynamics system in (1), by applying the proposed control laws and updating algorithms in (5), (9), and (10), all assumptions hold, the $S(t), \widetilde{W}, \widetilde{K}, \widetilde{R}$ are bounded.

Proof for theorem:

The following Lyapunov function candidate is considered as:

$$V(S(t), \widetilde{W}, \widetilde{K}, \tilde{R}) = \frac{1}{2} S^T S + \frac{1}{2} tr(\widetilde{W}^T \delta_w^{-1} \widetilde{W}) + \frac{1}{2} tr(\widetilde{K}^T \delta_k^{-1} \widetilde{K}) + \frac{1}{2} \tilde{R} \delta_r \tilde{R} \tag{11}$$

where$\tilde{R} = R - \hat{R}$. And, by differentiating (11) with respect to time, the following result can be obtained as:

$$\begin{aligned} \dot{V} &= -S^T K_s S + S^T \frac{a}{b} K_\alpha^{-1} diag(\dot{e}^{\frac{a}{b}-1})(\tilde{W}^T \phi + \tilde{K} S + \varepsilon - \tau_d + \tau_r) - tr(\tilde{W}^T \delta_w^{-1} \dot{\hat{W}}) - tr(\tilde{K}^T \delta_K^{-1} \dot{\hat{K}}) - \tilde{R} \delta_r^{-1} \dot{\hat{R}} \\ &= -S^T K_s S + S^T \frac{a}{b} K_\alpha^{-1} diag(\dot{e}^{\frac{a}{b}-1}) \tilde{W}^T \phi - tr(\tilde{W}^T \delta_w^{-1} \dot{\hat{W}}) + S^T \frac{a}{b} K_\alpha^{-1} diag(\dot{e}^{\frac{a}{b}-1}) \tilde{K} S - tr(\tilde{K}^T \delta_k^{-1} \dot{\hat{K}}) \\ &+ S^T \frac{a}{b} K_\alpha^{-1} diag(\dot{e}^{\frac{a}{b}-1})(\varepsilon - \tau_d) + S^T \frac{a}{b} K_\alpha^{-1} diag(\dot{e}^{\frac{a}{b}-1}) \tau_r - \tilde{R} \delta_r^{-1} \dot{\hat{R}} \end{aligned} \tag{12}$$

where $K_s = \frac{a}{b} K_\alpha^{-1} diag(\dot{e}^{\frac{a}{b}-1}) K \geq 0$. When the adaptive online updating algorithms in (10) are used, the (12) can be rewritten as:

$$\dot{V} = -S^T K_s S - \tilde{R} \frac{a}{b} \|S\| \left\| (K_\alpha^{-1}) diag\left(\dot{e}^{\frac{a}{b}-1}\right) \right\| + S^T \frac{a}{b} K_\alpha^{-1} diag\left(\dot{e}^{\frac{a}{b}-1}\right)(\varepsilon - \tau_d) +$$

$$S^T \frac{a}{b} K_\alpha^{-1} diag(\dot{e}^{\frac{a}{b}-1})\tau_r \tag{13}$$

With the robust control law in (9), from (13), yields:

$$\begin{aligned}
\dot{V} &= -S^T K_s S - \tilde{R}\frac{a}{b}\|S\|\left\|\left(K_\alpha^{-1}\right) diag(\dot{e}^{\frac{a}{b}-1})\right\| + S^T \frac{a}{b} K_\alpha^{-1} diag(\dot{e}^{\frac{a}{b}-1})(\varepsilon - \tau_d) \\
&\quad -\hat{R}\|S\|\frac{a}{b}\left\|\left(K_\alpha^{-1}\right) diag(\dot{e}^{\frac{a}{b}-1})\right\| - c\frac{a}{b} K_\alpha^{-1} diag(\dot{e}^{\frac{a}{b}-1}) S^T \tanh(S) \\
&\le -S^T K_s S - \tilde{R}\frac{a}{b}\|S\|\left\|\left(K_\alpha^{-1}\right) diag(\dot{e}^{\frac{a}{b}-1})\right\| + R\frac{a}{b}\|S\|\left\|\left(K_\alpha^{-1}\right) diag(\dot{e}^{\frac{a}{b}-1})\right\| \\
&\quad -\hat{R}\|S\|\frac{a}{b}\left\|\left(K_\alpha^{-1}\right) diag(\dot{e}^{\frac{a}{b}-1})\right\| - c\frac{a}{b} K_\alpha^{-1} diag(\dot{e}^{\frac{a}{b}-1}) S^T \tanh(S) \\
&\le -S^T K_s S - c\frac{a}{b} K_\alpha^{-1} diag(\dot{e}^{\frac{a}{b}-1}) S^T \tanh(S)
\end{aligned} \tag{14}$$

From the result in (14), then,$\dot{V}(S(t), \widetilde{W}, \widetilde{K}, \tilde{R}) \le 0, \forall t \ge 0$. Thus, the $\dot{V}(S(t), \widetilde{W}, \widetilde{K}, \tilde{R})$ is a negative semi-definite function, as a result, $V(S(t), \widetilde{W}, \widetilde{K}, \tilde{R}) \le V(S(0), \widetilde{W}, \widetilde{K}, \tilde{R})$. It can easily see that if $S(t), \widetilde{W}, \widetilde{K}, \tilde{R}$ are bounded at initial time$t = 0$, they will stay this bounded state for all$t \ge 0$. Therefore, $\widehat{W}, \widehat{K}, \hat{R}$ are also bounded. By applying the Lyapunov theorem [10, 12], the stability of the proposed control system is guaranteed. Next, the finite-time convergence of the $S(t), e(t)$ will be addressed by the following lemma.

Lemma: By assuming that$\left\|\widetilde{W}^T\phi - \widehat{K}S\right\| \le c_1, \left\|\tilde{R}\right\| \le c_2, (c_1, c_2 > 0)$, the $S(t), e(t)$ will approach to zero with finite time.

Proof for Lemma: Consider the second Lyapunov candidate function as:

$$V_s(S(t)) = \frac{1}{2} S^T S \tag{15}$$

By differentiating $V_s(t)$ in (15) with respect to time, yields:

$$\begin{aligned}
\dot{V}_s &= S^T \dot{S} = S^T \frac{a}{b} K_\alpha^{-1} diag(\dot{e}^{\frac{a}{b}-1})(\tilde{W}^T\phi - \hat{K}S + \varepsilon - \tau_d) \\
&\quad -\frac{a}{b}\|S\|\left\|\left(K_\alpha^{-1}\right) diag(\dot{e}^{\frac{a}{b}-1})\right\| \hat{R} - S^T \frac{a}{b} K_\alpha^{-1} diag(\dot{e}^{\frac{a}{b}-1}) c \tanh(S) \\
&\le \frac{a}{b}\|S\|\left\|\left(K_\alpha^{-1}\right) diag(\dot{e}^{\frac{a}{b}-1})\right\|\left\|\tilde{W}^T\phi - \hat{K}S\right\| + \frac{a}{b}\|S\|\left\|\left(K_\alpha^{-1}\right) diag(\dot{e}^{\frac{a}{b}-1})\right\|\|\varepsilon - \tau_d\| \\
&\quad -\frac{a}{b}\|S\|\left\|\left(K_\alpha^{-1}\right) diag(\dot{e}^{\frac{a}{b}-1})\right\| \hat{R} - S^T \frac{a}{b} K_\alpha^{-1} diag(\dot{e}^{\frac{a}{b}-1}) c \tanh(S) \\
&\le \frac{a}{b}\|S\|\left\|\left(K_\alpha^{-1}\right) diag(\dot{e}^{\frac{a}{b}-1})\right\|\left\|\tilde{W}^T\phi - \hat{K}S\right\| + \frac{a}{b}\|S\|\left\|\left(K_\alpha^{-1}\right) diag(\dot{e}^{\frac{a}{b}-1})\right\|\left\|\tilde{R}\right\| - S^T \frac{a}{b} K_\alpha^{-1} diag(\dot{e}^{\frac{a}{b}-1}) c \tanh(S) \\
&\le \frac{a}{b}\|S\|\left\|\left(K_\alpha^{-1}\right) diag(\dot{e}^{\frac{a}{b}-1})\right\|(c_1 + c_2 - c)
\end{aligned} \tag{16}$$

From the above Proof for theorem, it notes that $S(t), \widetilde{W}, \widetilde{K}, \hat{R}, \tau_r$ are bounded. If the designed constant parameter is selected as:$c_1 + c_2 < c$, then:

$$\dot{V}_s \le \frac{a}{b}\|S\|\left\|(K_\alpha^{-1}) diag(\dot{e}^{\frac{a}{b}-1})\right\|(c_1 + c_2 - c) \le 0 \tag{17}$$

From the result in (17), if $\dot{e} \ne 0, \dot{V}_s < 0$, it can conclude that $S(t)$ converges to zero within finite time [1, 2]. If $\dot{e} = 0, e \ne 0$ then $\ddot{e} = \widetilde{W}^T\phi - \widehat{K}S + \varepsilon - \tau_d - c\,tanh(S)$. This shows that [1, 2] if $S \ne 0, \dot{e} = 0$ then $\ddot{e} \le -c$ for $S > 0$ and $\ddot{e} \ge c$ for $S < 0$. As a result, $\dot{e} = 0$ is not an attractor and manifold $S(t)$ will converge to zero in finite time, as well as $e(t)$ will also converge to zero in finite time [1, 2]. This completes proof.

3 NUMERICAL SIMULATION RESULTS

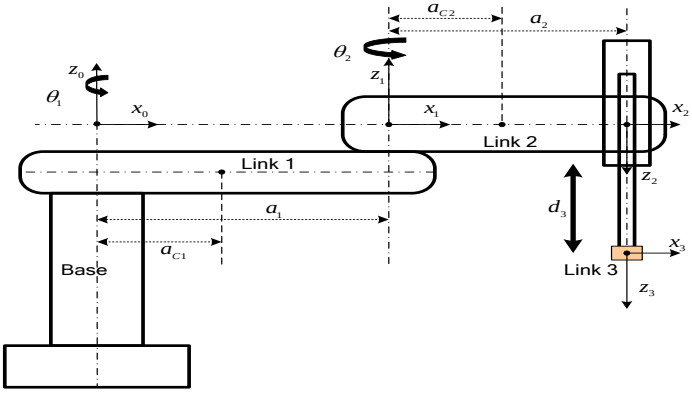


Figure 2: The MR 3 – DOF model

The MR 3 – DOF model, as shown in the Fig. 2, is used in the simulation process. All the dynamics parameters of the MR are provided by [10]. To verify the effectiveness of the proposed control method, the NTSM controller, call as NTSM1, with fixed proportional control gains and sat (.) function [1 – 8] and the proposed NTSM controller are applied to provide the compared simulation results. For more details, the control laws of the NTSM1 are: $\tau = -\hat{f} - \frac{b}{a}K_\alpha(\dot{e}^{2-\frac{a}{b}}) - KS + \tau_r$ and $\tau_r = -\hat{R}\frac{S(K_\alpha^{-1})diag(\dot{e}^{\frac{a}{b}-1})}{\|S\|\left\|(K_\alpha^{-1})diag(\dot{e}^{\frac{a}{b}-1})\right\|} - csat(S)$, its updating algorithms are same with the proposed NTSM controller. Therefore, the control parameters of both methods are provided as: $a = 11, b = 9, K_\alpha = diag(2), K = diag(50), c = 0.3, \delta_w = diag(50), \delta_k = diag(50), \delta_r = diag(50)$. The desired trajectories are defined as $q_d = [0.5\,cos(\,1.5t),0.5\,cos(\,2t),0.1\,cos(\,1.5t)](rad)$, the friction and disturbances as:
$F(\dot{q}) = 0.1[0.5\dot{q}_1 + 0.2sign(\dot{q}_1); 0.5\dot{q}_2 + 0.2sign(\dot{q}_2); 0.5\dot{q}_3 + 0.2sign(\dot{q}_3)]$,
$\tau_d = [0.5sin(2t); 0.5\,sin(\,2t); 0.5\,sin(\,2t)]$.
The Matlab/Simulink software with sample time is 0.001 (s) is done for the simulations. The simulation procedures are suggested as: when simulation time is 2.5s, $\tau_d = 2.0[\sin(20t)\,sin(\,20t); sin(\,20t)]$, the tip loads will add to the links 2 and 3 of the RM system (2 kg and 1.5 kg on links 1 and 2), and $F(\dot{q}) = 1.0[0.5\dot{q}_1 + 0.2sign(\dot{q}_1);; 0.5\dot{q}_2 + 0.2sign(\dot{q}_2); 0.5\dot{q}_3 + 0.2sign(\dot{q}_3)]$.
First, the Fig. 3 shows the compared simulation results for the tracking positions and errors of the proposed NTSM and NTSM1 methods. These results have shown that the tracking performances of both methods are good. However, the proposed NTSM scheme has the finite time-convergence and tracking error features that seem to be better. In this research, with the main purpose of improving for the NTSM with fixed main PD control parameters. Therefore, based on the adaptive updated main control parameters of the proposed NTSM method, as shown in the Fig. 4, the PD gains for NTSM1 method are selected in considering of the time-uncertainties in this simulation process to be able to obtain the acceptable best results. However, if the PD gains are increased to achieve faster finite time-convergences, the chattering phenomenon can occur. By observing with the Fig. 5, for the NTSM1 method, the chattering phenomenon has gradually appeared when the control conditions are impacted by the highly time-frequency uncertainties. In addition, when compared with the proposed NTSM method, the robust term of the NTSM1 method is not good and smooth, especially in the control condition variations for the amplitudes and continuous characteristics.

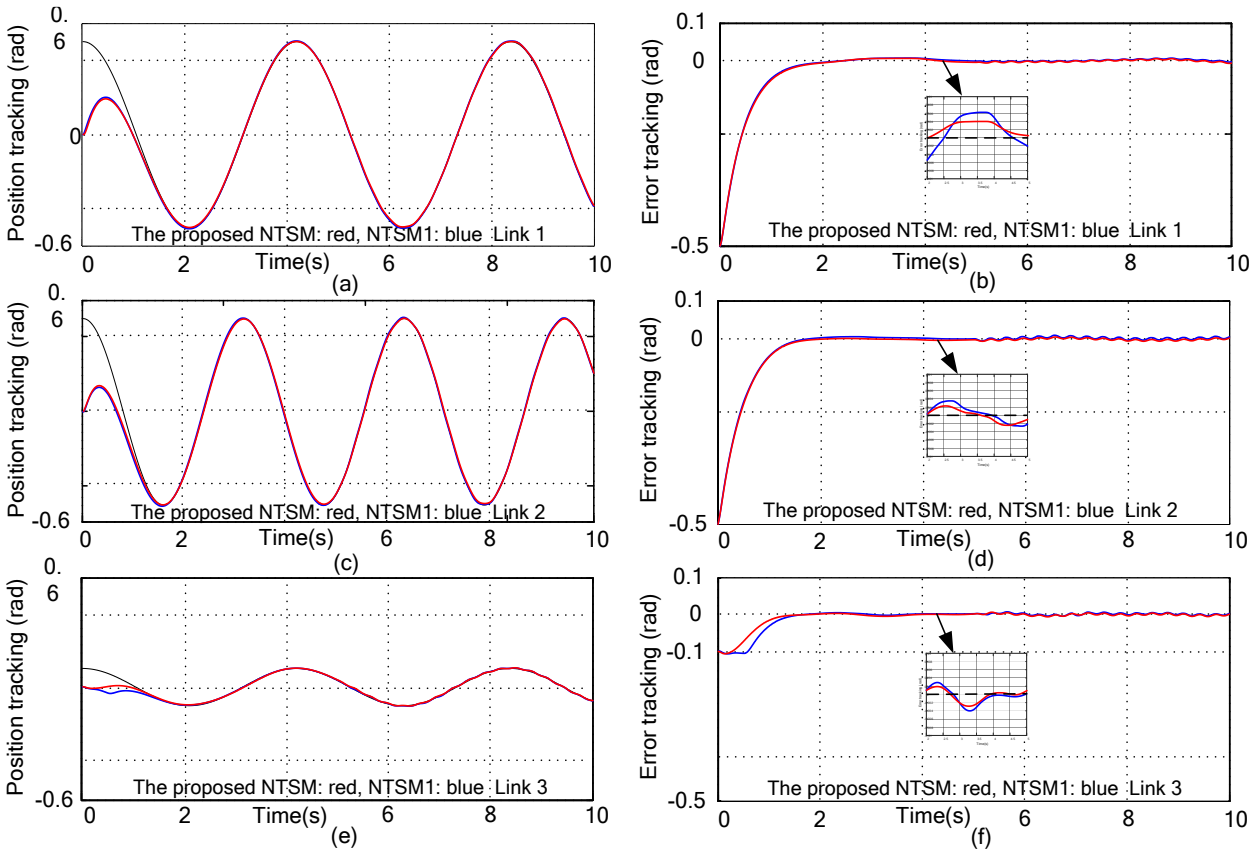


Figure 3: Tracking control simulation results

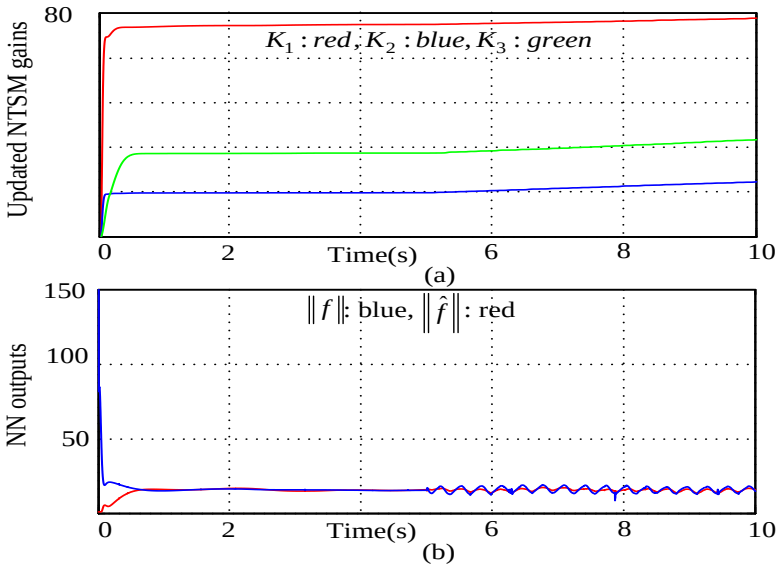


Figure 4: The simulation results: Adaptive updated parameters

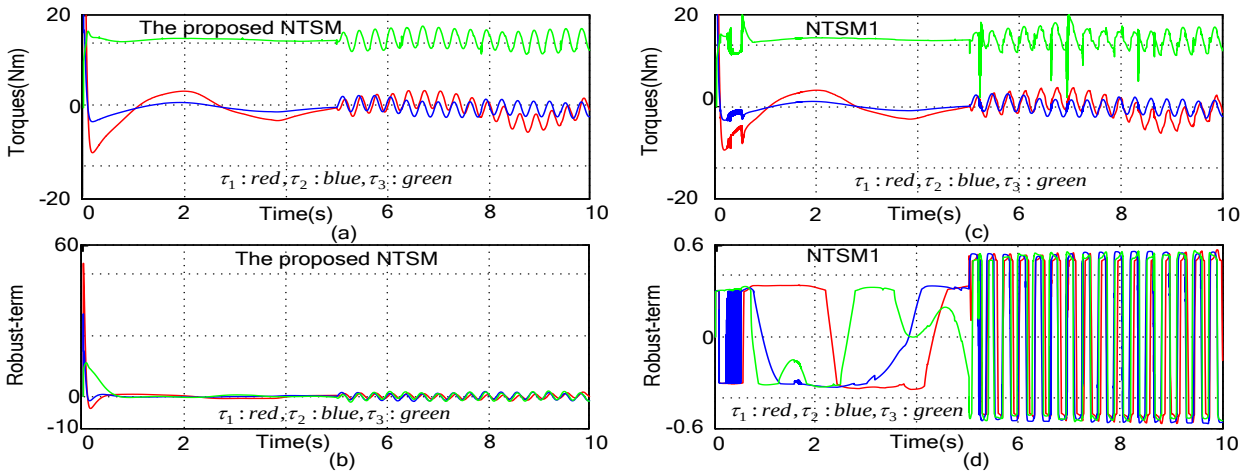


Figure 5: The simulation results: Control inputs

4 CONCLUSIONS

In this work, the improved NSTM scheme has been proposed and successfully addressed for the MR system. The improvement focuses on the adaptive self-updating for the PD gains of the traditional NTSM methods. In addition, by replacing the sign(.) of sat(.) functions with the nonlinear smooth function, tanh(.), the problems of the chattering phenomenon and discontinuous control signals have been explored. Based on the Lyapunov and the compared simulation results, the proposed NSTM strategy can be considered as a good alternative method in increasing the tracking and adaptability performances for the MR control.

ACKNOWLEDGMENT

This work was supported by the Industrial University of Ho Chi Minh City, Vietnam. The authors would like to thank the associate editor and the reviewers for their valuable comments.

REFERENCES

[1] C. Lin, Nonsingular Terminal Sliding Mode Control of Robot Manipulators Using Fuzzy Wavelet Networks, IEEE Transactions on Fuzzy Systems, vol. 14, no. 6, pp. 849-859, 2006.

[2] M. Jin, J. Lee, P. H. Chang and C. Choi, Practical Nonsingular Terminal Sliding-Mode Control of Robot Manipulators for High-Accuracy Tracking Control, IEEE Transactions on Industrial Electronics, vol. 56, no. 9, pp. 3593-3601, 2009.

[3] L. Zhang, Y. Wang, Y. Hou and H. Li, Fixed-Time Sliding Mode Control for Uncertain Robot Manipulators, IEEE Access, vol. 7, pp. 149750-149763, 2019

[4] M. Van, M. Mavrovouniotis and S. S. Ge, An Adaptive Backstepping Nonsingular Fast Terminal Sliding Mode Control for Robust Fault Tolerant Control of Robot Manipulators, IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 49, no. 7, pp. 1448-1458, July 2019.

[5] M. Van, X. P. Do and M. Mavrovouniotis, Self-tuning fuzzy PID-nonsingular fast terminal sliding mode control for robust fault tolerant control of robot manipulators, ISA Transactions, vol. 96, 2020.

[6] J. Zhai and G. Xu, A Novel Non-Singular Terminal Sliding Mode Trajectory Tracking Control for Robotic Manipulators, IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 68, no. 1, pp. 391-395, Jan. 2021.

[7] Y. Sun et al., Neural Network-Based Tracking Control of Uncertain Robotic Systems: Predefined-Time Nonsingular Terminal Sliding-Mode Approach, IEEE Transactions on Industrial Electronics, vol. 69, no. 10, pp. 10510-10520, 2022.

[8] Y. Liu, W. Yan, T. Zhang, C. Yu and H. Tu, Trajectory Tracking for a Dual-Arm Free-Floating Space Robot With a Class of General Nonsingular Predefined-Time Terminal Sliding Mode, IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 52, no. 5, pp. 3273-3286, 2022.

[9] T. L. Mai, Hybrid adaptive tracking control method for mobile manipulator robot based on Proportional–Integral–Derivative technique, Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, vol. 235, no. 22, pp. 6463-6480, 2021.

[10] T. L. Mai, H. T. Tran, V. H. Tran, N. A. Tran and H. H. Nguyen, An adaptive proportional-derivative control method for robot manipulator, Journal of Science and Technology - IUH, vol. 52B, no. 4, pp. 34-43, 2021.

[11] F. L. Lewis, K. Liu and A. Yesildirek, Neural net robot controller with guaranteed tracking performance, IEEE Transactions on Neural Networks, vol. 6, no. 3, pp. 703–713, 1996.

[12] J. J. E. Slotine, W. Li, Applied Nonlinear Control, Englewood Cliffs, NJ: Prentice-Hall, 1991.

ICATSD2F.104

BIOMECHANICS-BASED DEVELOPMENT OF AN UPPER LIMB REHABILITATION ROBOT

HUU-TOAN TRAN[1*], THANG-LONG MAI[1], VAN-HUNG TRAN[1], NGOC-ANH TRAN[1], THANH-HAI TRAN[2], HOANG-HIEU NGUYEN[1]

[1]*Faculty of Electronic Technology, Industrial University of Ho Chi Minh City, Vietnam*

[2]*Office of Facilities Management, Industrial University of Ho Chi Minh City, Vietnam*

**tranhuutoan@iuh.edu.vn*

Abstract. Upper limb rehabilitation robot is an exoskeletal mechanism whose joints and links are compatible with those of the human upper limb. Due to human factors, design and manufacturing of the robot requires a solid procedure of biomechanics analysis, mechanism evaluation and actuator collection. Inspired by the difficulties behind this procedure, this paper presents the design and development of a prototype of an Upper Limb Rehabilitation Robot (ULRob). The robot structure is designed through the analysis of human upper limb biomechanics and dynamics. Based on the requirements of human joint power, the solutions of mechanism and actuator for the exoskeleton are drawn. In our initial testing, a custom-build hardware and a position control were applied to all joints to evaluate the robot's performance. Normalized angular tracking human-robot errors that is $(27.2 \pm 3.15)\ x10^{-3}$ and the resulting human-robot interaction force that is maximum of 8.32 N showed the robot's applicable efficiency.

Keywords. Biomechanics Technology, Wearable Robot, Upper Limb Exoskeleton, Rehabilitation Robot, Physical Human-robot Interaction.

1 INTRODUCTION

Wearable robots, especially exoskeletons, are able to assist human in difficult and dangerous situations due to their unlimited power from external supplies. These robots bring a numerous practical application of human assistance in daily life. Especially, for patients with movement difficulties, the forgoing assistive robots play an important role to support them in physical training and rehabilitation [1, 2]. According to pathology breakdown that causes severe motor effects, stroke is thirdly ranked in popularity worldwide [3]. After entering a chronic phase of 6 months, 30% to 66% of stroke patients did not recover their upper limb function, meanwhile, only 5% to 20% of stroke patients is showed to complete the upper limb recovery according to Kwakkel *et al.* [4]. Therefore, the appropriate device-based therapies to enhance rehabilitation effectiveness for the upper limbs have been applied, such as standard multidisciplinary rehabilitation with one-to-one manual interactions with therapists (Van Peppen *et al.* [5]), practice with mirror and imagination (Langhorne *et al.* [6]), and robot-aided support (Prange *et al.* [7], Kwakkel *et al.* [8]). These endeavors demonstrated that the development of wearable assistive upper exoskeletons is a robot trend but pose challenges in rehabilitation application [9, 10]. Due to the human-robot cooperation, some key factors in the design of the exoskeletons need to be considered such as to tailor biomechanics individual users; to generate efficient assistive torques/forces; and to do not obstruct during human-robot interaction forces/torques. The exoskeleton is used to carry the human upper limb, so the anthropomorphic techniques and biomechanical problems need to be considered for the exoskeleton design. Furthermore, the choice of drive source and transmission should be analyzed appropriately to increase performance and reduce weight because the robot is a wearable device. Alexander Otten *et al.* introduced a prototype of four actuated DOF upper exoskeleton called LIMPACT using rotational series elastic hydraulic motors [11]. As a preamble,

this hydraulic system brings an advanced mechanism approach but makes the robot structure be more voluminous. Tomoyuki *et al.* proposed a pneumatic-electric actuation system for an upper body exoskeleton robot [12]. This exoskeleton used backdrivable actuators to make sufficiently large torques but the pneumatic source and related accessory are complex and not straightforward to control. Kiguchi *et al.* introduced a two DOF exoskeletal robotic system using direct DC motor including harmonic drive that can reduce robot structure but be high weight for users [13, 14].
Inspired by these challenges, this paper proposes the development of an Upper Limb Rehabilitation Robot, called ULRob, to support patients in recovery exercises, in which the mechanism solution and prototype of ULRob are highlighted. The remainder of this paper is organized as follows; firstly, the main design results consisting of mechanism and actuator solutions of ULRob are briefly introduced. Subsequently, ULRob structure and prototype are drawn. Then, hardware and control strategy are implemented on the robot. Finally, experiments will be implemented to evaluate the effectiveness of the robot on quasi-recovery exercises in rehabilitation.

2 BIOMECHANICS OF HUMAN ARM

In general, human arm is a complicated musculoskeletal system consisting of the coordinated movements among upper arm, forearm, and hand. These musculoskeletons linked to human body by three joints: shoulder, elbow, and wrist, respectively. In terms of several simplifications, the kinetics of the human arm includes seven degrees of freedom in the Cartesian coordinates. As depicted in Fig. 1, the core motions are performed at the three joints: adduction/abduction, flexion/extension, circumduction at the shoulder joint; flexion/extension at the elbow joint; supination/pronation, adduction/abduction, flexion/extension at the wrist joint. Based on statistics for the motor torque at the above seven main motions, the flexion/extension about the z axes at the shoulder (axis 2), the elbow (axis 4), and the wrist (axis 6) joints cost the most energy consumption at each joint during a daily activity [15]. Rosen *et al.* considered the arm action associated with reaching to head level in which the joint axis torques significantly change about axes 2 and 4. In particular, the total joint torque about axis 2 changes approximately from -4.5 to 5.2 Nm and about axis 4 changes approximately from 0 Nm to 3.6 Nm. The peak torques about the two axes 2 and 4 are around 12 to 15 time larger compared to that about the remaining axes [16]. In the scope of rehabilitation exercises, ULRob is designed so that it could assist the patients in flexion/extension movements at the shoulder and the elbow joints. The abduction/abduction movement at the shoulder and the flexion/extension movement at wrist joints are designed as passive ones. This is also the as-need fundamental to choose the type and location of the transmission power at the joints for the exoskeleton design.

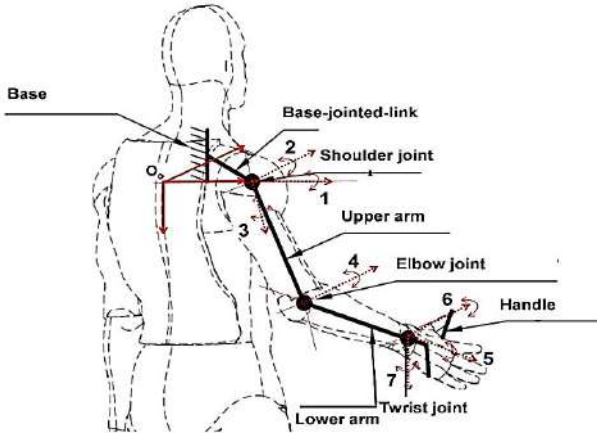


Fig. 1. Kinematic human arm consists of main movements: adduction/abduction, flexion/extension, circumduction around axis 1, 2, 3, respectively, at the shoulder joint; flexion/extension around axis 4 at the elbow joint; supination/pronation, adduction/abduction, flexion/extension around axis 5, 6, 7, respectively at the wrist joint.

3 MECHANICAL DESIGN

3.1 Mechanism solution

The primary design goals of ULRob are to ensure safety compact mechanism, motion range of rehabilitation exercises, lightweight for carrying during normal operation. These objectives affect the selection of a mechanical transmission structure for the ULRob system. There are three exoskeleton configurations proposed to compare and then select the most appropriate option for the design. The first option of robot mechanism is based on sliding-crank transmission to upper arm and forearm powered by electric motors. As seen in Fig.2a, this option can be backdriven, but makes the robot structure be voluminous, rendering higher weight.

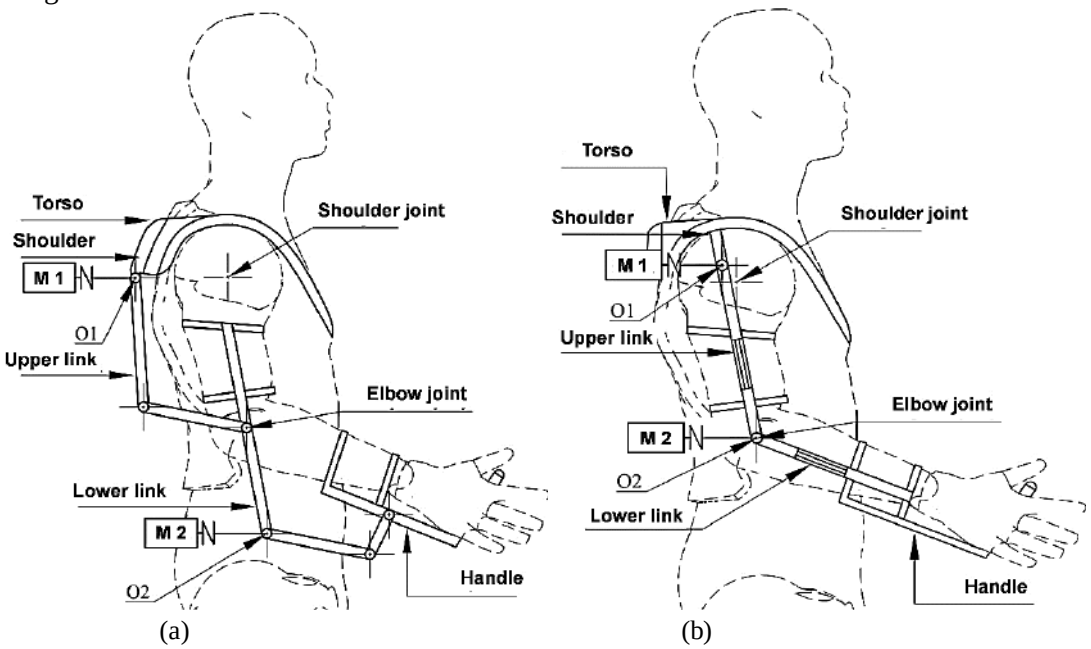


Fig. 2. The first option using sliding-crank transmission (a) and second option using direct harmonic drive (b) for the collection of mechanism for ULRob

The second option is based on harmonic drive powered directly by electric motor at each joint. Although this solution ensures the mechanical movement of the patient's skeletal structure and the robot be more compact than the first option, the transmission from the motor to center of rotation of the shoulder and the elbow joints makes the upper arm and forearm links be more complicated and low back-drive ability. Furthermore, the motors are allocated at the joints that make the link inertial be higher.

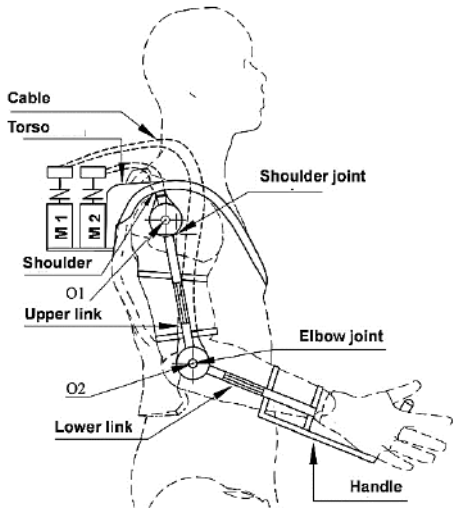


Fig.3. The third option using self-designed cable drive for the mechanism solution for *ULRob*

The third option takes advantages of both first options by transferring power from electric motors on the torso to the shoulder and the elbow joints through self-designed cable drive system, then be the solution for ULRob mechanism as seen in Fig. 3. This non-localized actuation system allows the exoskeleton to operate safely in a wide range of motion and attenuate the rigid link's weight. This option also brings the robot a lower link inertial due to the motor location on the torso. Moreover, a low–ratio gearbox is utilized to make ULRob a high back-drive ability

3.2 ULRob structure

Based on the analysis of biomechanical data and mechanic solutions, especially the dimensions and limits of the human arm movement, the computer-aided design (CAD) of *ULRob* model on human arm is designed using Inventors (AutoDesk Inc.) as shown in Fig. 4. The designed model includes five main assembly modules: the torso, the shoulder link, the upper arm, the forearm (lower) links, and the handle. Within the framework of the paper, the ULRob design model and prototype are summarized in table 1 and illustrated in Fig.4.

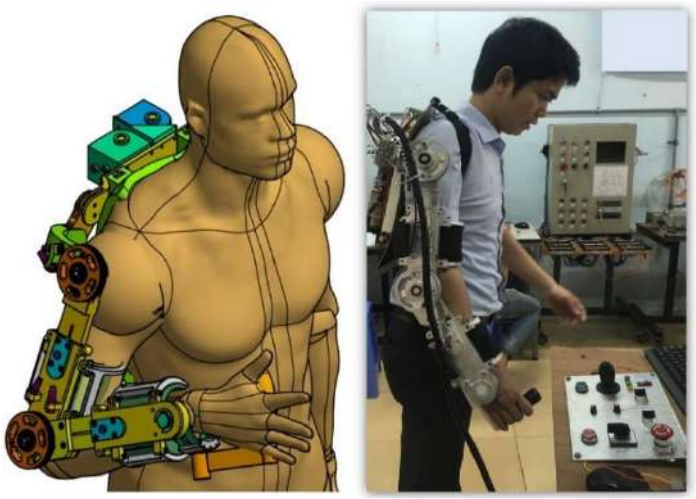


Fig.4. The computer aided design (CAD) and prototype of ULRob model

Table 1. Kinematic parameters for human model and ULRob (Max angle (◦), Max torque (Nm)). ULRob design specification was scaled for a 70 kg healthy person operating with load of 5 kg and the biomechanics profile was cited from [14, 15].

Joint	**Movement**	**Human model**		**ULRob**	
		Max Angle *(°)*	Max Torque *(Nm)*	Max Angle *(°)*	Max Torque *(Nm)*
Shoulder	Extension	-32	-4.5	-30	-44.8
	Flexion	178	10	135	44.8
	Adduction	-15	-1.8	-5	Passive
	Abduction	95	7.5	90	Passive
Elbow	Extension	0	-1.5	0	-32.5
	Flexion	152	2.2	135	32.5
Wrist	Extension	-60	-0.3	-45	Passive
	Flexion	60	0.3	45	Passive

4 PROTOTYPES

4.1 Hardware configuration

ULRob is a human-robot interaction system, so the control design has to be meet safety criteria including safety source and safety interrupt, in addition to a traditional real-time precise control. The ULRob control hardware is built as described in the Fig. 5. The control center is a computer using Matlab RealTimeWorkshop application to provide a consistent behavior. Two PCI 6221 NI (National Instrument) cards are utilized for the real-time control and communication via computer's PCI Express slots. Two digital servo drives (MSD) with 20 A of maximum continuous current are fed to drive the motors. A custom-built control panel is additionally built to select control modes and training exercises for users. The ULRob system utilizes several sensing modalities of position and interaction for control and evaluation.

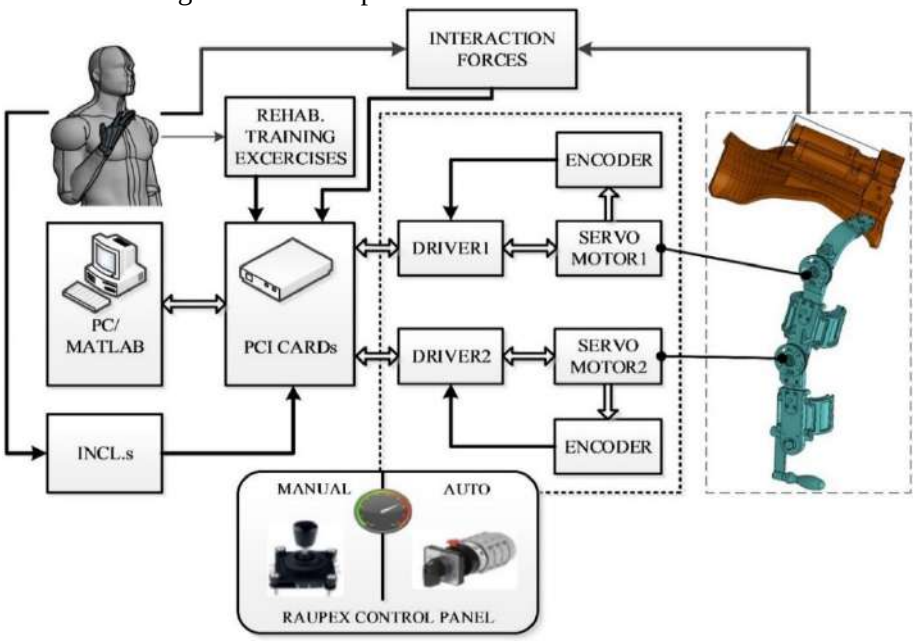


Fig.5. Hardware diagram of ULRob

4.2 Finite-State Machine Control

Ordinarily, the training exercise assigned by a therapist includes the trajectories of the shoulder and elbow joints incorporated together. In a training exercise, flexion/extension movements are performed at the shoulder and elbow joints while the transitions are triggered by limit points on the desired trajectories. Based on this principle, finite-state machine control method is fed for $Exercise_{ij}$ (Ex_{ij}) in the control panel where i is level of training complexity and j is level of operation velocity. These exercises are predefined corresponding to recovery progress of each individual patient. As an example, Fig. 6 shows the finite-state machine control model of $Exercise_{13}$ for programming. Assisted upper arm shift to up (State 1) and assisted lower arm shift to up (State 3) are defined for the active control of single joints, of which the signal collected from limit sensor is to lock transition conditions. State 2 and State 4 are defined for the training of both joints simultaneously until the joints return to the initial (zero) position of ULRob (State 5 and State 6).

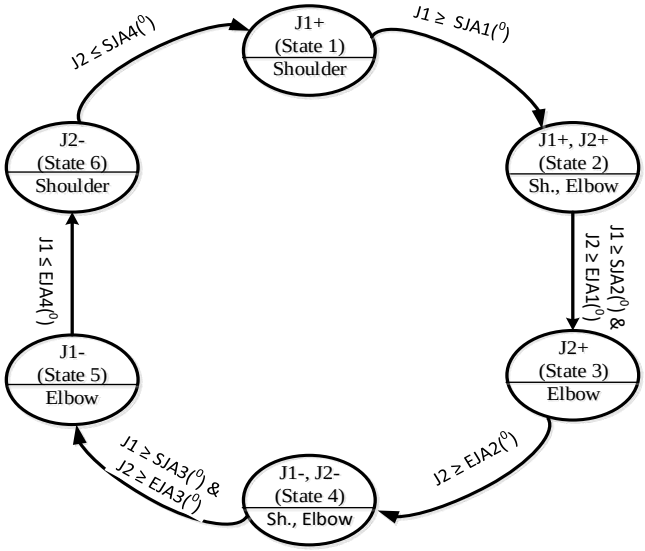


Fig.6. Finite-state machine control for one of the exercises programmed for ULRob ($Exercise_{31}$). In the figure, the abbreviations J1- Shoulder Joint; J2- Elbow Joint; SJA- Shoulder Joint Angle; EJA-Elbow Joint Angle are used

5 EXPERIMENTAL RESULTS AND DISCUSSION

In order to confirm the ability of ULRob to operate stably and safely, we first conducted swing exercises at each joint. Operation speed that varies from $\pi/4$ [rad/s] to $3\pi/2$ [rad/s] is divided into three speed levels for the finite-state machine control algorithm. The swing exercises were also utilized to check workspace of the shoulder and elbow joint angles. For performance evaluation of the finite-state machine control, five predefined exercises were implemented at the three training speeds for three subjects (A, B, C). Each exercise was repeated three times, in which data was collected and averaged to evaluate. As mentioned above, the collected data are the joint angles of both the operator and exoskeleton, and the resulting interaction forces at the upper and lower cuffs through each session.

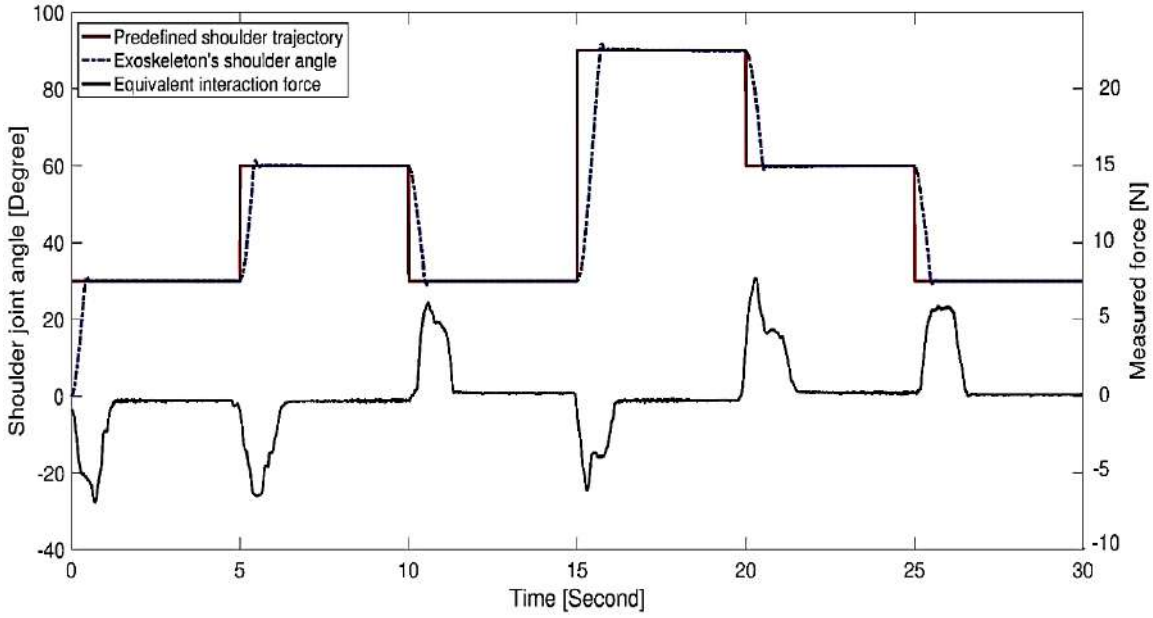


Fig. 7. Control performances at shoulder joint of the active finite-state machine control with $Exercise_{31}$

Fig. 7 shows the control performances of one of the training exercises, namely, $Exercise_{31}$ at the shoulder. As mentioned in section 4.2, $Exercise_{31}$ means the level of complexity is 3 and the level of movement

velocity is 1 ($\pi/4$ rad/s). It can be seen that the angular operator-ULRob tracking error is almost asymptotic to zero at every triggered set point. The normalized square sum of angular errors for the shoulder joint is slightly increased corresponding to higher training speeds. For example, as summarized in Table 1, the average tracking error of the shoulder joint at the speed of $\pi/4$ [rad/s] is increased by 2.5% and 4% of that at the speeds of $5\pi/8$ [rad/s] and $3\pi/2$ [rad/s], respectively. The performance indexes AE and IF are defined as follows:

$$AE_k = \frac{\int_0^{T_c} e_k^2(t)dt}{\int_o^{T_c} \theta_{Ok}^2(t)dt}; IF_k = \frac{\int_0^{T_c} f_{Ok}^2(t)dt}{\int_0^{T_c} \theta_{Ok}^2(t)dt}, \quad (1)$$

where AE_k (k=1,2) is the normalized square sum of joint errors (e_k) at the shoulder (AE_1) and the elbow (AE_2) in the interval Tc. Also, IF_k (k= 1, 2) is the normalized square sum of the interaction force (f_{Ok}) at the upper arm cuff (IF_1) and the forearm cuff (IF_2). Fig 11 also shows a slightly overshoot of angular tracking whose maximum percent value is around 5% at the shoulder joint at all speeds indicates relative stability of the system. The resulting interaction force at the upper arm is regular from -0.65 N to 0.65 N except transition moments. For example, when the shoulder joint shifts the state from 30 degree to 90 degree, the interaction force reaches a peak value of 7.32 N then reduce to approximately 0.44 N. This means the interaction force is significantly resulted now of transition while the robot does not impede the operator's motion. The peak value of the interaction force at the upper arm is increased by about 13.6 % compared that at the lower arm. Table 1. shows the performance indexes AE and IF with respect to different training exercises and operators at the three levels of training speed. Besides, the operator's feeling through three iterations of each exercise using Semantic Differential (SD) method is summarized in this table. The Semantic Differential (SD) evaluation approach is advocated for the subjective evaluation of the exoskeleton systems [17, 18].

Table 2. Experimental results of finite-state machine control on three operators (A-C): Mean values of AE and IF for the operator-exoskeleton control performance across different training exercises and training speeds ($Exercise_{ij}$); and evaluation of the operator's feeling recorded in the first and third sessions for every exercise

No.	Description (sex, weight, height)	Arm length (l_1/l_2 mm)	Exercise (Ex_{ij})	Performance indexes at shoulder		Performance indexes at elbow		(SDE) of the operator's feeling	
				AE1	IF1 ($x10^{-3}$)	AE2	IF2 ($x10^{-3}$)	First session	Third session
1	Subject A (man, 72 kg, 1,70 m)	292/240	Ex_{32}	0.0229	6.81	0.0432	3.18	1	1
			Ex_{22}	0.0243	6.72	0.0441	2.92	1	2
			Ex_{13}	0.0235	6.94	0.0458	3.02	2	1
2	Subject B (woman, 62 kg, 1,61 m)	272/220	Ex_{12}	0.0212	6.32	0.0397	2.83	2	2
			Ex_{31}	0.0205	6.17	0.0411	2.75	2	2
			Ex_{41}	0.0201	6.08	0.0426	3.06	1	2
3	Subject C (man, 68,5 kg, 1,66 m)	285/231	Ex_{23}	0.0219	6.46	0.0407	3.11	2	1
			Ex_{41}	0.0238	6.58	0.0415	3.15	-1	1
			Ex_{33}	0.0225	6.71	0.0426	2.89	2	1

6 CONCLUSION AND FUTURE WORKS

This paper summarizes the design and control of an Upper Limb Rehabilitation Robot, called ULRob, capable of supporting human arm muscle-skeleton. Here, the main issues related to a prototype of the exoskeleton consist of the analysis of human upper limb biomechanics, the solution of mechanism, the CAD design, and the control platform have been presented as a preamble for the development of the exoskeleton. In particular, a new prototype of the exoskeleton including a custom-built real-time hardware has been built and evaluated. In our initial testing, the finite-state machine position control was applied to

all joints to evaluate the robot's performance. Control performance of tracking human-robot errors and the resulting human-robot interaction force showed the robot's applicable efficiency. Our next step is to optimize the design parameters to the next version of ULRob such that the exoskeleton can provide more convenience to users. Furthermore, advance control algorithms will be applied with the adaptation ability for the robot.

REFERENCES

[1] Lum P, Reinkensmeyer D, Mahoney R, et al., Robotic Devices for Movement Therapy after Stroke: Current Status and Challenges to Clinical Acceptance, Topics in Stroke Rehabilitation, 8(4):40-53, 2002.

[2] Fasoli SE, Krebs HI, Stein J, Frontera WR, and Hogan N. Effects of Robotic Therapy on Motor Impairment and Recovery in Chronic Stroke. Archives of Physical Medicine and Rehabilitation, 84 (4):477–482, 2003.

[3] Mackay J, Mensah G, Mendis S, Greenlund K, The Atlas of Heart Disease and Stroke. World Health Organization, 2004.

[4] Kwakkel G, Kollen BJ; van der Grond J, et al. Probability Of Regaining Dexterity In The Flaccid Upper Limb. Stroke, 34(9): 2181-2186, 2003.

[5] Van Peppen RPS, Kwakkel G, Wood-Dauphinee S, et al. The Impact of Physical Therapy on Functional Outcomes after Stroke: What's the Evidence?, Clinical Rehabilitation, 18(8): 833-862, 2004.

[6] Langhorne P, Bernhardt J, Kwakkel G, Stroke rehabilitation, The Lancet, 377(9778): 1693-1702, 2011.

[7] Prange GB, Jannink MJA, Groothuis-Oudshoorn, CGM. et al. Systematic Review of The Effect of Robot-Aided Therapy on Recovery of The Hemiparetic Arm after Stroke, Journal of Rehabilitation Research and Development, 43(2):171, 2006.

[8] Kwakkel G, Kollen BJ and Krebs HI, Effects of Robot-Assisted Therapy on Upper Limb Recovery after Stroke: A Systematic Review, Neurorehabilitation and Neural Repair, 22(2): 111-121, 2008.

[9] Masiero S, Poli P, Rosati G et al., The Value of Robotic Systems in Stroke Rehabilitation, Expert Review of Medical Devices, 11(2): 187-198, 2014.

[10] Rehmat N, Zuo J, Meng W et al., Upper Limb Rehabilitation Using Robotic Exoskeleton Systems: a Systematic Review, International Journal of Intelligent Robotics and Applications, 2:283–295, 2018.

[11] Otten A, Voort C, Stienen A, et al., LIMPACT: A Hydraulically Powered Self-Aligning Upper Limb Exoskeleton, IEEE/ASME Transactions on Mechatronics, 20(5):2285-2298, 2015.

[12] Noda T, Teramae T, Ugurlu B, et al., Development of an Upper Limb Exoskeleton Powered via Pneumatic Electric Hybrid Actuators with Bowden Cable, 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp.3573-3578, 2014

[13] Kiguchi K, Active Exoskeletons for Upper-Limb Motion Assist, Journal of Humanoid Robotics, 4(3):607-624, 2007.

[14] Wu KY, Su YY, Yu YL, et al., A 5-Degrees-of-Freedom Lightweight Elbow-Wrist Exoskeleton for Forearm Fine-Motion Rehabilitation, IEEE/ASME Transactions on Mechatronics, 24(6):2684 2695, 2019.

[15] Winter DA, Biomechanics and Motor Control of Human Movement, 4th edition, New Jersey: Jonh Wiley and Sons Inc, 2009.

[16] Rosen J, Perry JC, Manning N et al., The Human Arm Kinematics and Dynamics during Daily Activities – Toward a 7 DOF Upper Limb Powered Exoskeleton, The 12th International Conference on Advanced Robotics-ICAR, pp. 532-539, 2005.

[17] Lee S and Sankai Y, Virtual Impedance Adjustment in Unconstrained Motion for an Exoskeletal Robot Assisting the Lower Limb, Advanced Robotics, 19(7):773-795, 2005

[18] Lee KH, Baek SG, Lee HJ, et al., Enhanced Transparency for Physical Human-robot Interaction using Human Hand Impedance Compensation, IEEE/ASME Transactions on Mechatronics, 23(6):2662-2670, 2018.

ICATSD2F.105

DEVELOPMENT OF A POSITIONING SYSTEM FOR MOBILE ROBOTS USING LIDAR SENSOR IN UNDETERMINED ENVIRONMENTS

NGUYEN VAN LANH

Faculty of Electronics Technology, Industrial University of Ho Chi Minh City

nguyenvanlanh@iuh.edu.vn

Abstract. Self-propelled robots have been used instead of humans. Mobile robots can complete many tasks in dangerous places that humans cannot access. However, there are challenges for mobile robots such as frequent changes in the environment, obstacles, and goal position. This paper proposes the development of positioning methods using a Lidar sensor for mobile robots in undetermined environments. To perform these tasks, the following steps are done. Firstly, a principle for distance measurement and orientation angle of the Lidar sensor is presented. Secondly, the positions of the landmarks are obtained using the Lidar sensor. Thirdly, the position of the Lidar sensor is obtained based on landmark positions. Fourthly, an Extended Kalman filter (EKF) is used to estimate the best position of the mobile robot. Finally, simulation results show the effectiveness of the proposed positioning system in an undetermined environment.

Keywords. Lidar sensor, mobile robot, positioning, undetermined environment.

1 INTRODUCTION

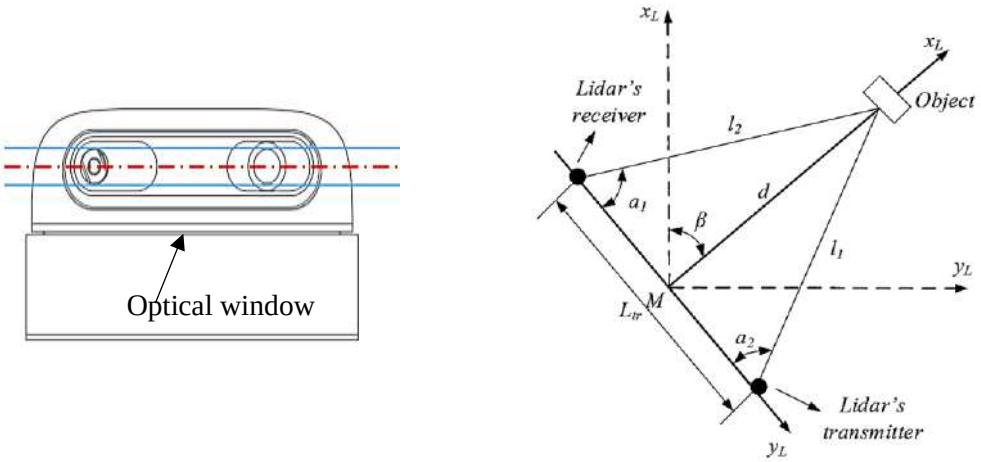
Applications of mobile robots have been increasing in various fields during recent decades. These mobile robots can complete various tasks in unsafe and dangerous places that workers cannot access. Typically, mobile robots with differential drive-wheeled configurations were chosen for these above tasks. The positioning design problem is one of the requisites and important tasks. Several methods are used to locate and control the tracking of mobile robots in [1-5]. To make the mobile robot move automatically, a positioning system design is needed. Several types of research have been proposed [6-11]. S. Y. Lee et al. [6] proposed a semi-guided navigation methodology by using magnetic tapes. Y. C. Yuan et al. [7] also proposed a navigation method based on a wall-following algorithm. P.T. Doan et al. [8] proposed tracking control for robots using a camera sensor. However, these methods depend on predefined lines and waypoints. Gatesichapakorn, S et al. [9] proposed an autonomous mobile robot navigation using 2D LiDAR and RGB-D camera. Asadi, K et al. [10] proposed vision-based integrated mobile robotic system for real-time applications in construction. Luyang Li, et al. [11] proposed an adaptive trajectory tracking of nonholonomic mobile robots using vision-based position and velocity estimation. Therefore, mobile robots can only operate in a determined environment. There are different types of location technologies that use wireless networks such as global positioning systems (GPS) in [12-16], WiFi navigation systems in [18], cellular phone navigation systems, RFID navigation systems in [19-20], localization systems based on the multiple sensors in [21-23]. Among these positioning systems, GPS is the most popular positioning system. However, GPS signals are easily intercepted by building materials, and it does not work well in indoor environments. There have been many types of research about positioning and navigation methods using laser sensors. T. L. Bui et al. [24] presented trajectory-tracking control based on laser sensor navigation using reflectors. P. S. Pratama et al. [25] presented a positioning for automated guided vehicles in a partially known environment. However, these positioning methods depend on predetermined reflectors and landmarks. Therefore, this paper proposes a positioning system for the mobile robot using a Lidar sensor. To perform these tasks, the following steps are done. Firstly, a principle for distance measurement and orientation angle of the Lidar sensor is presented. Secondly, the positions of the landmarks are obtained using the Lidar sensor. Thirdly, the position of the Lidar sensor is obtained based on landmark positions. Fourthly, an Extended Kalman filter (EKF) is used to estimate the best position of the mobile robot. To verify the effectiveness of the proposed positioning system, simulation results are shown. These results of

the proposed positioning system are compared with the laser sensor navigation using reflectors proposed by T. L. Bui in [24].

2 SYSTEM MODELING

2.1 Specification and construction of Lidar sensor

Fig. 1 shows the Lidar sensor. It adopts the low-cost laser triangulation measurement system.



a) Optical window b) Principle for distance measurement of Lidar sensor

Fig.1. Construction of Lidar sensor

The construction of the Lidar sensor is a 2D 360-degree type of laser sensor using a horizontal scanning line as shown in Fig. 1(a). In this paper, the Lidar sensor is used for detecting landmarks of positioning the mobile robot. It is also used for detecting obstacles in experimental environments. The basic principle of the Lidar sensor is described in Fig. 1(b). The Lidar sensor transmits modulated infrared laser signal, and the laser signal is then reflected by an object to be detected. The returning signal is then sampled by a vision acquisition system in the Lidar sensor. The data consists of an output distance value and angle value between a detected object and the Lidar sensor. The Lidar sensor performs the 360-degree scan for the experimental environment.

Fig. 1(b) shows the basic principle of measuring the distance between an object and the Lidar sensor. By using the cosine law of triangle from Fig. 1(b), the following can be obtained:

$$\begin{cases} cos\, a_1 = \frac{L_{tr}^2+l_2^2-l_1^2}{2L_{tr}l_2} \\ cos\, a_2 = \frac{L_{tr}^2+l_1^2-l_2^2}{2L_{tr}^2l_1} \end{cases} \Rightarrow \begin{cases} a_1 = arccos(\frac{L_{tr}^2+l_2^2-l_1^2}{2L_{tr}l_2}) \\ a_2 = arccos(\frac{L_{tr}^2+l_1^2-l_2^2}{2L_{tr}l_1}) \end{cases} \quad (1)$$

where $l_1 = ct_1,\ l_2 = ct_2$,L_{tr} is the distance between Lidar's transmitter and Lidar's receiver, $c = 3 \times 10^8 (m/s)$ is the speed of light, and $t_1,\ t_2$ are laser light transmission time and laser light reception time, respectively.

The distance *d* between the Lidar sensor and an object can be obtained as follows:

$$d = l_1\, sin\, a_2 = l_2\, sin\, a_1 \quad (2)$$

The orientation angle $\beta = 0 \sim 360^0$ is determined by the angle value between a detected object and the origin rotational angle of the Lidar sensor as follows:

$$\beta = -[\omega_l T_l - n_l(360^0) - \beta_0] \quad (3)$$

where $\omega_l = 360^0 \times f_l$ (degree/s) is a scanning angular velocity of the Lidar sensor, f_l (Hz) is a scanning frequency of the Lidar sensor, $T_l(s) = (t_1 + t_2)$ is a scanning time, n_l is a scanning revolution number of the Lidar sensor, and β_0 (degree) is an initial angle of Lidar sensor.

2.2 System modeling of the mobile robot

In this paper, it is assumed that the mobile robot is a differential drive-wheeled configuration. System modeling of the mobile robot and the coordinate frames are shown in Fig. 2.

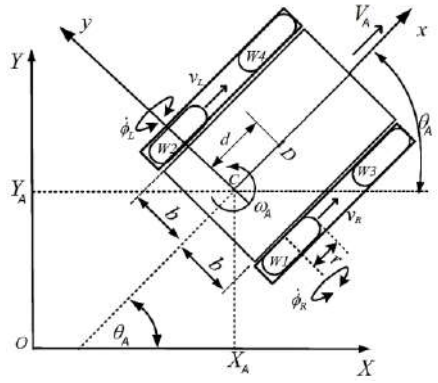


Fig. 2. System modeling of the mobile robot

In this figure, *OXY* is a global coordinate frame; *Cxy* is the moving coordinate frame mounted on the platform of the mobile robot. *W1* and *W2* are driving standard wheels with radius *r*. *W3* and *W4* are passive wheels. v_R and v_L are linear velocities for the right and left wheel, respectively. $\dot{\phi}_R$ and $\dot{\phi}_R$ are angular velocities for the right and left wheel, respectively. $x_v = [X_A \quad Y_A \quad \theta_A]^T$ is the posture vector of the mobile robot and $z = [V_A \ \omega_A]^T$ is a velocity vector of the mobile robot. *b* is the distance between *Cx axis* and wheels of the mobile robot. *D* is the mass center of the mobile robot.

The mathematical modeling of the mobile robot can be expressed as follows:

$$\dot{x}_v = \begin{bmatrix} \dot{X}_A \\ \dot{Y}_A \\ \dot{\theta}_A \end{bmatrix} = \begin{bmatrix} cos\,\theta_A & 0 \\ sin\,\theta_A & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} V_A \\ \omega_A \end{bmatrix} \text{ for } H(\dot{x}_v) = \begin{bmatrix} cos\,\theta_A & 0 \\ sin\,\theta_A & 0 \\ 0 & 1 \end{bmatrix} \tag{4}$$

The relation of the linear velocity V_A and angular velocity ω_A with the right and left wheel angular velocities ($\dot{\phi}_R$, $\dot{\phi}_L$) can be expressed as:

$$\begin{bmatrix} V_A \\ \omega_A \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2b} & \frac{-1}{2b} \end{bmatrix} \begin{bmatrix} v_R \\ v_L \end{bmatrix} = \begin{bmatrix} \frac{r}{2} & \frac{r}{2} \\ \frac{r}{2b} & \frac{-r}{2b} \end{bmatrix} \begin{bmatrix} \dot{\phi}_r \\ \dot{\phi}_l \end{bmatrix} \tag{5}$$

By substituting Eq. (5) into Eq. (4), the kinematic equation of the mobile robot can be expressed as follows:

$$\begin{bmatrix} \dot{X}_A \\ \dot{Y}_A \\ \dot{\theta}_A \end{bmatrix} = \begin{bmatrix} \frac{r}{2}\cos\theta_A & \frac{r}{2}\cos\theta_A \\ \frac{r}{2}\sin\theta_A & \frac{r}{2}\sin\theta_A \\ \frac{r}{2b} & \frac{-r}{2b} \end{bmatrix} \begin{bmatrix} \dot{\phi}_R \\ \dot{\phi}_L \end{bmatrix} \tag{6}$$

3 POSITIONING SYSTEM DESIGN

In this section, a Lidar sensor is used to detect landmarks. System modeling of the Lidar sensor is shown in Fig. 3. *OXY* is a global coordinate frame, Mx_Ly_L is a local coordinate frame of the Lidar sensor.

From Fig. 2, the position of i^{th} landmark in the local coordinate frame Mx_Ly_L can be obtained as follows:

$$\begin{cases} x_{iL} = d_i\,cos\,\beta_i \\ y_{iL} = d_i\,sin\,\beta_i \end{cases} \quad for\ i = 1,2,\cdots,n \tag{7}$$

where d_i is a distance from the Lidar sensor to the i^{th}landmark and $\beta_i$is a scanning angle between the Lidar sensor and the i^{th}landmark.

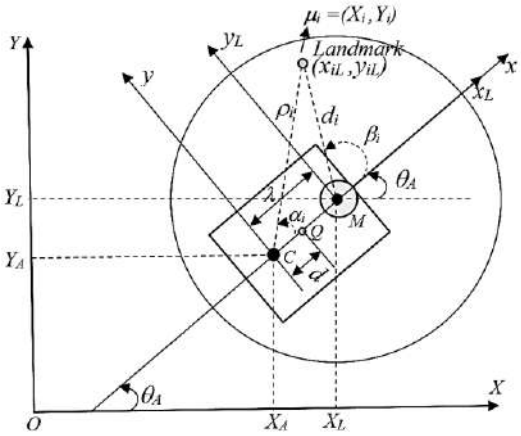


Fig. 3. Landmark detection using Lidar sensor

In global coordinate frame *OXY*, the position of the i^{th}landmark $\mu_{i,k}$ at time *k* is obtained as follows:

$$\mu_{i,k} = \begin{bmatrix} X_{i,k} \\ Y_{i,k} \end{bmatrix} = \begin{bmatrix} X_{A,k} \\ Y_{A,k} \end{bmatrix} + \begin{bmatrix} cos\,\theta_A & -sin\,\theta_A \\ sin\,\theta_A & cos\,\theta_A \end{bmatrix} \begin{bmatrix} x_{iL} \\ y_{iL} \end{bmatrix} + \lambda \begin{bmatrix} cos\,\theta_A \\ sin\,\theta_A \end{bmatrix} \quad (8)$$

where$(X_{A,k}, Y_{A,k})$ is the position of the mobile robot at time *k* and λ is the distance between *M* and *C*. From the landmark position obtained by measurement of Eq. (8), the position and orientation of the Lidar sensor can be determined as the followings.

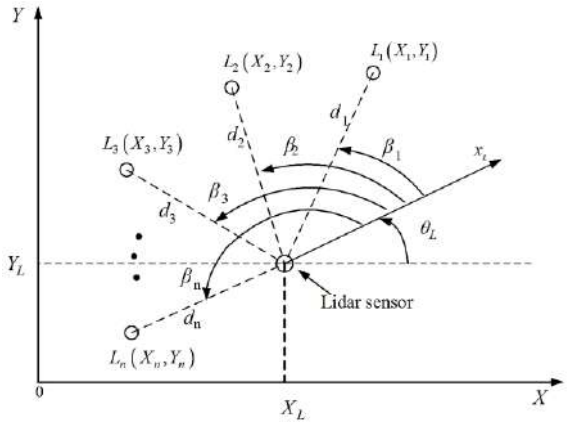


Fig.4. Principle for position measurement of the Lidar sensor

Fig. 4 shows the basic principle for position measurement of the Lidar sensor.
The coordinate of the Lidar sensor (X_L, Y_L) and its orientationθ_L concerning the global coordinate frame can be obtained as follows:

$$X_L = \frac{(p_2 p-1)[(p_1+p)X_1+(p_1 p-1)Y_1]}{(p_2-p_1)(1+p^2)} \quad - \frac{(p_1 p-1)[(p_2+p)X_2+(p_2 p-1)Y_2]}{(p_2-p_1)(1+p^2)} \quad (9)$$

$$Y_L = \frac{(p_1+p)[(p_2+p)X_2+(p_2 p-1)Y_2]}{(p_2-p_1)(1+p^2)} \quad - \frac{(p_2+p)[(p_1+p)X_1+(p_1 p-1)Y_1]}{(p_2-p_1)(1+p^2)} \quad (10)$$

$$\theta_L = \tan^{-1}\frac{Y_1-Y_L}{X_1-X_L} - \beta_1 \quad (11)$$

where $p_1 = tan\,\beta_1$, $p_2 = tan\,\beta_2$, $p_3 = tan\,\beta_3$, $p = tan\,\theta_L$, and p is given by:

$$p = \frac{(p_3-p_1)(Y_1-Y_2-p_1X_1+p_2X_2)-(p_2-p_1)(Y_1-Y_3-p_1X_1+p_3X_3)}{(p_3-p_1)(p_1Y_1+X_1-p_2Y_2-X_2)-(p_2-p_1)(p_1Y_1+X_1-p_3Y_3-X_3)} \quad (12)$$

where (X_i, Y_i) is a position of the i^{th} landmark in the global coordinate frame *OXY*.
The Lidar sensor is installed on the mobile robot as shown in Fig. 3. Therefore, the coordinate of the mobile robot and its orientation concerning the global coordinate frame can be obtained as follows:

$$x_v = \begin{bmatrix} X_A \\ Y_A \\ \theta_A \end{bmatrix} = \begin{bmatrix} X_L \\ Y_L \\ \theta_L \end{bmatrix} - \lambda \begin{bmatrix} cos\,\theta_A \\ sin\,\theta_A \\ 0 \end{bmatrix} = \begin{bmatrix} X_L - \lambda\, cos\,\theta_A \\ Y_L - \lambda\, sin\,\theta_A \\ \theta_L \end{bmatrix} \quad (13)$$

Finally, an Extended Kalman Filter (EKF) is used to get the best-estimated position of the mobile robot by combining the current position of the mobile robot and positioning results obtained from the Lidar sensor. Block diagram of the positioning system can be described as shown in Fig. 5.

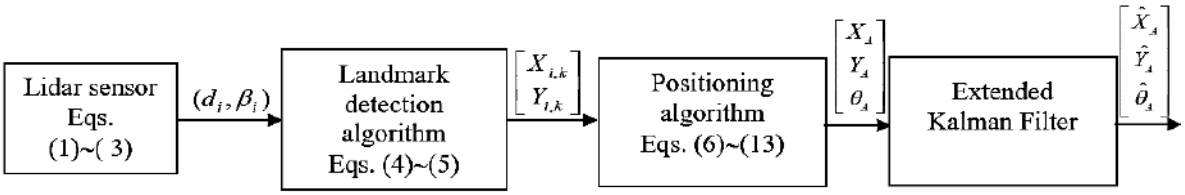


Fig.5. Block diagram of the positioning system

4 SIMULATION RESULTS

The specification and parameters of the Lidar sensor in this simulation is shown in Table 1.

Table 1. Specification of Lidar sensor

No.	Parameters	Values
1	Light source	Infrared (785nm)
2	Laser class	FDA Class1
3	Field of view (angular range)	0^0~360^0
4	Scanning frequency (rate)	10Hz(600rpm)
	Sample frequency	8000Hz
5	Angular resolution	0.9^0 for 10 Hz scan rate
6	Distance resolution	< 0.5mm for distance of 0.2 m ~ 1.5m
		< 1% of the distance for all distance range
7	Operating range	0.15m ~ 8m
8	Data communication	Serial (RS-232)
9	Sample Duration	0.125ms
10	Operating voltage/current	5 VDC/1.5A
11	Laser power	3mW
12	Weight	0.35kg
13	Dimensions (DxH)	75.70mm x 40.80mm

In this simulation, the mobile robot tracks the reference trajectory from START point *S* (*X*=-6m, *Y*=8m) to GOAL point *G*(*X*=6m, *Y*=-7m). These simulation results of the proposed positioning system are compared with the laser sensor navigation using reflectors proposed by T. L. Bui in [24].
Fig. 6 shows a green line for the desired reference trajectory and a magenta line for the trajectory result by simulation. The posture vector of the mobile robot $x_v = [X_A \quad Y_A \quad \theta_A]^T$ is obtained based on the positioning system using a Lidar sensor in section 3. The positions of the mobile robot during the simulation are shown in Fig. 6. The mobile robot tracks the desired reference trajectory well in Figs. 6(a-d) and the mobile robot finally reaches the goal position of Fig. 6(d).
The positioning results on the global coordinate frame *OXY* are shown in Figs. 7(a-c). Fig7. (a) shows a blue continuous line for the desired reference trajectory *on the X axis Xr(m)*, and a magenta dashed line for

the trajectory result *on the X axis XA(m)* by simulation. Fig. 7(b) shows a blue continuous line for the desired reference trajectory *on the Y axis Yr(m)*, and a magenta dashed line for the trajectory result *on the Y axis YA(m)* by simulation. Fig. 7(c) shows a blue center line for the desired reference orientation θr, and a magenta dashed line for the orientation result θA by the simulation.

These figures show that the proposed positioning system successfully makes the mobile robot tracks the desired reference trajectory with an acceptable small error.

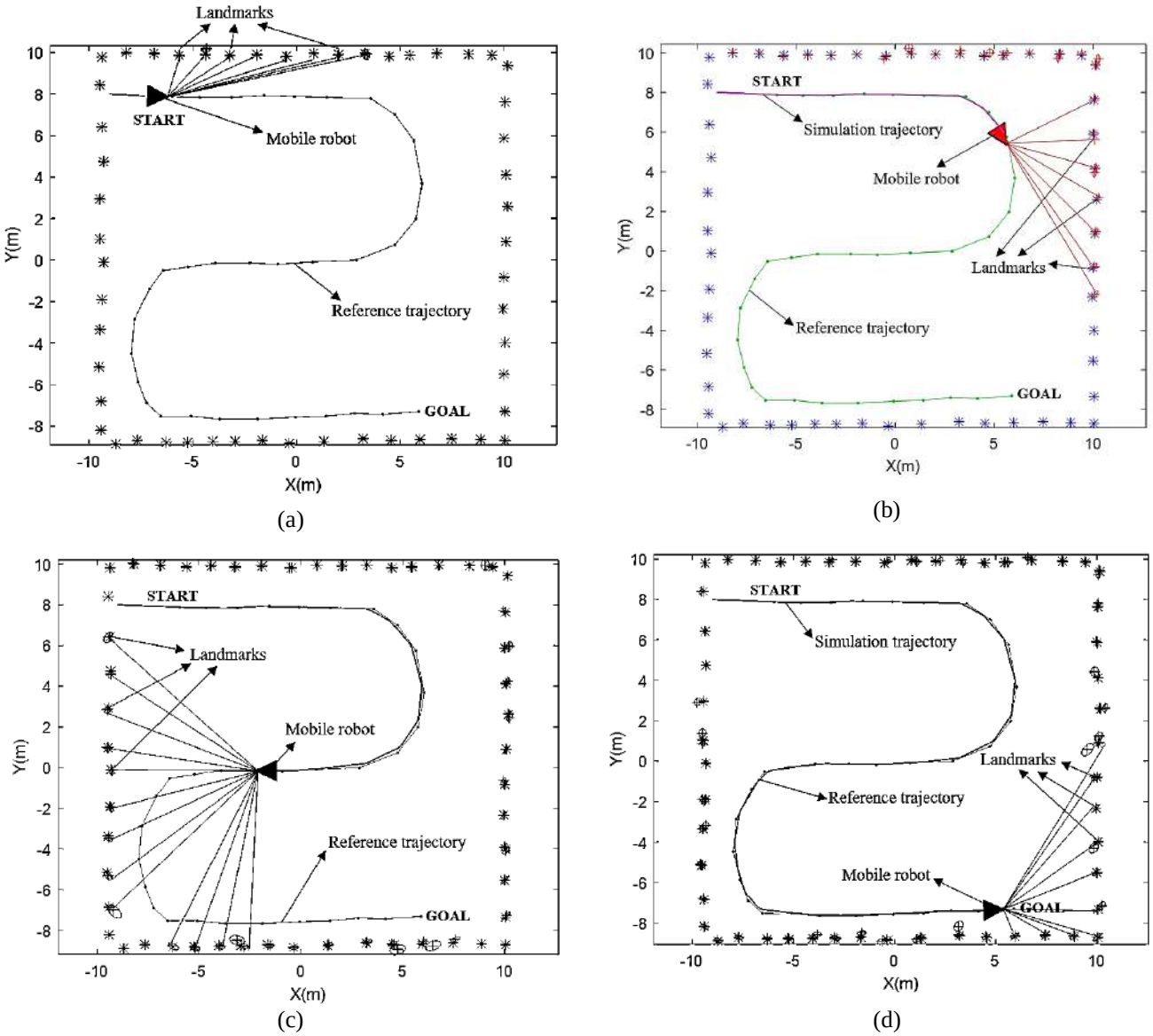
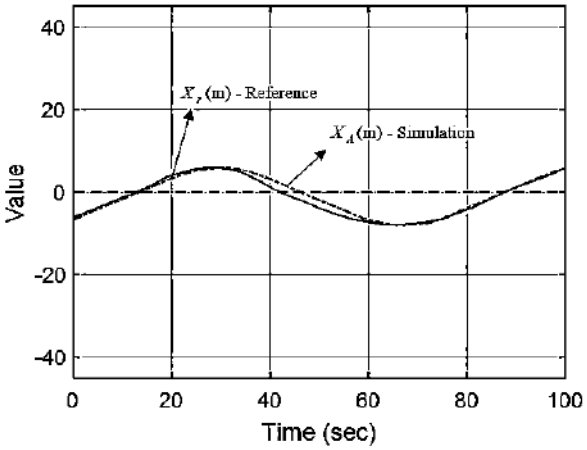
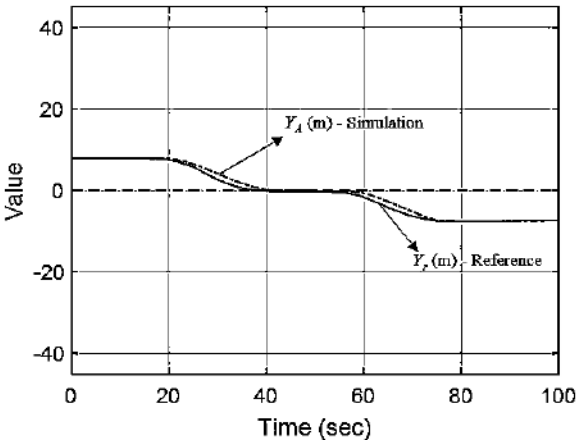


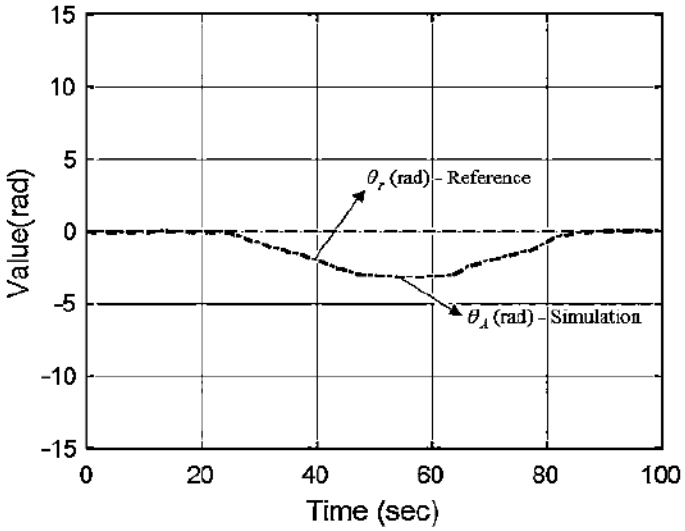
Fig. 6. Positioning results of mobile robot



(a) Positioning result of the mobile robot on *X axis*



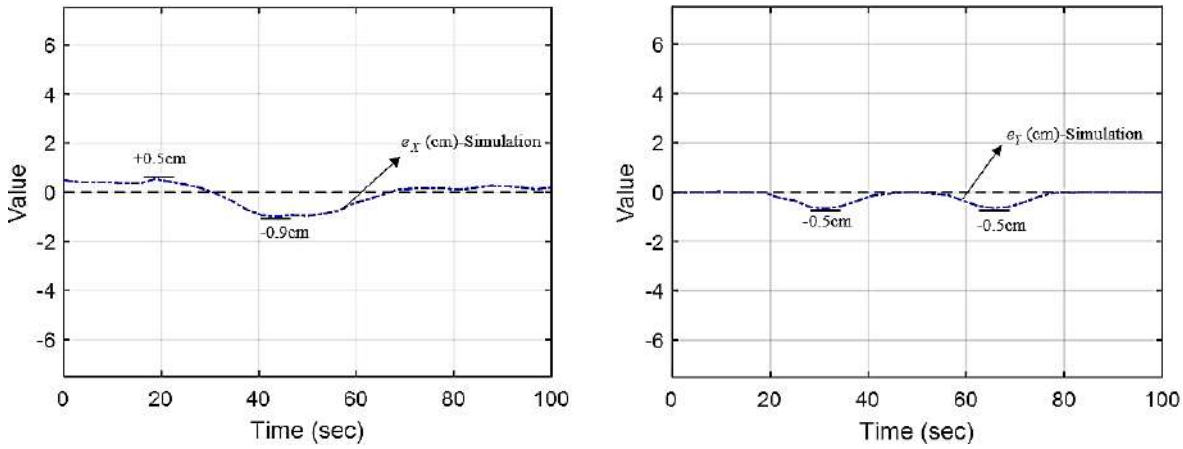
(b) Positioning result of the mobile robot on *Y axis*



(c) Orientation angle of mobile robot

Fig. 7. Positioning results of the mobile robot: (a) on *X axis*, (b) on *Y axis*, and (c) orientation angle

The tracking errors on the global coordinate frame *OXY* are shown in Figs. 8(a-b). In Fig. 8(a), the tracking error $e_X = X_r - X_A$ in the simulation converges to zero and it is bounded from −0.9cm to +0.5cm around zero, while the tracking error *e1* in [24] bounded from −0.25cm to +0.25cm around zero. In Fig. 8(b), the tracking error $e_Y = Y_r - Y_A$ in the simulation converges to zero and it is bounded from −0.5cm to 0cm around zero, while the tracking error *e2* in [24] bounded from −0.5cm to +2.5cm around zero. These results show the tracking performance of the proposed method. However, tracking errors have a small value at the straight line and a big value at the curved line of the reference trajectory.



(a) Global tracking error on *X axis* e_X (b) Global tracking error on *Y axis* e_Y

Fig. 8. Global tracking error on *X axis* e_X and on *Y axis* e_Y

5 CONCLUSIONS

A positioning system design for a mobile robot using a Lidar sensor was proposed. The simulation results showed good positioning performance for the mobile robot. The tracking error *on the X axis* converged to zero and it was bounded from −0.9cm to +0.5cm around zero, the tracking error *on the Y axis* converged to zero and it was bounded from −0.5cm to 0cm around zero. During the simulation, the mobile robot tracked the desired trajectory well and finally reached the goal position. Experiment and prove the role of the EKF filter will be developed for the mobile robots in the undetermined environment in future work.

ACKNOWLEDGEMENT

This paper is supported by the Industrial University of Ho Chi Minh City, Ho Chi Minh City, Vietnam.

REFERENCES

[1] H. Hu and D. Gu, Landmark-based navigation of nndustrial mobile robots, *International Journal of Industry Robot*, vol. 27, no. 6, pp. 458–467, 2000.

[2] A. Diosi and L. Kleeman, Laser scan matching in polar coordinates with application to SLAM, *Proc. of IEEE International Conference on IRS*, 2005, pp. 1439–1444.

[3] N. M. Kwok, Q. P. Ha, and G. Fang, Data association in bearing-only SLAM using a cost function-based approach, *Proc. of IEEE- ICRA*, 2007, pp. 4108–4113.

[4] T. Bailey, J. Nieto, J. Guivant, M. Stevens, and E. Nebot, "Consistency of the EKF-SLAM algorithm, *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006, pp. 3562–3568.

[5] W. H. Durrant and T. Bailey, Simultaneous Localisation and Mapping (SLAM): Part I the essential algorithms, *IEEE Robotics and Automation Magazine*, vol. 13, no. 2, pp.99–110, 2006.

[6] S. Y. Lee and H.W. Yang, Navigation of automated guided vehicles using magnet spot guidance method, *JRCIM*, vol. 28, no. 2, pp. 425-436, 2012.

[7] S. C. Yuan and L. Yao, Robust type-2 fuzzy control of an automatic guided vehicle for wall-following, *Proc. Of ICSCPR*, 2009, pp. 172-177.

[8] P. T. Doan, T. T. Nguyen, V. T. Dinh, H. K. Kim, and S. B. Kim, Path tracking control of automatic guided vehicle using camera sensor, *Proc. of ISACE*, 2011, pp. 20-26.

[9] Gatesichapakorn, S., Takamatsu, J., and Ruchanurucks, M, ROS based autonomous mobile robot navigation using 2D LiDAR and RGB-D camera, *first international symposium on instrumentation, control, artificial intelligence, and robotics, IEEE,* 2019, pp. 151-154.
[10] Asadi, K., Ramshankar, H., Pullagurla, H., Bhandare, A., Shanbhag, S., Mehta, P., and Wu, T, Vision-based integrated mobile robotic system for real-time applications in construction, *Automation in Construction*, pp.470-482, 2018.
[11] Luyang Li, Yun-Hui Liu, Tianjiao Jiang, Kai Wang, and Mu Fang, Adaptive trajectory tracking of monholonomic mobile robots using vision-based position and velocity estimation, *IEEE Transactions on Cybernetics*, 2018.
[12] Y. Zhang and K. T. Chong, A GPS / DR data fusion method based on the GPS characteristics for mobile robot navigation, *International Journal of Control and Automation*, vol. 7, no. 10, pp. 119–132, 2014.
[13] Zhang, Yuanliang, and Dong Pyo Hong, Navigation of mobile robot using Low-cost GPS, *International Journal of Precision Engineering and Manufacturing*, pp. 847-850, 2015.
[14] Bao, J., Yao, X., Tang, H., and Song, A., Outdoor navigation of a mobile robot by following GPS waypoints and local pedestrian lane, *8th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), IEEE*, 2018, pp. 198-203.
[15] Zhang, Yuanliang, and Kil to Chong, An GPS/DR navigation system using neural network for mobile robot, *International journal of precision engineering and manufacturing*, pp. 2513-2519, 2014.
[16] Hewawasam, H. S., Ibrahim, M. Y., Kahandawa, G., and Choudhury, T. A., Comparative study on object tracking algorithms for mobile robot navigation in GPS-denied environment, *IEEE International Conference on Industrial Technology (ICIT)*, 2019, pp. 19-26.
[17] Iqbal, J., Xu, R., Sun, S., and Li, C., Simulation of an autonomous mobile robot for LiDAR-based in-field phenotyping and navigation, *Robotics*, vol. 9, no. 2, 46, 2020.
[18] J. Biswas and M. Veloso, WiFi Localization and navigation for autonomous indoor mobile robots, Proc. of IEEE-ICRA, 2010, pp. 4379–4384.
[19] Peng, J., Qin, Y., Wei, Q., He, Q., Wan, Z., and Jiang, RFID-based indoor mobile robot navigation, *International Journal of RF Technologies*, pp.1-8, 2018.
[20] Teo, J. H., Loganathan, A., Goh, P., and Ahmad, N. S., Autonomous mobile robot navigation via RFID signal strength sensing, *International Journal of Mechanical Engineering and Robotics Research*, vol.9, no. 8, pp.1140-1144, 2020.
[21] Nguyen, L. A., Dung, P. T., Ngo, T. D., and Truong, X. T., Improving the accuracy of the autonomous mobile robot localization systems based on the multiple sensor fusion methods, *3rd International Conference on Recent Advances in Signal Processing, Telecommunications & Computing (SigTelCom), IEEE*, 2019, pp. 33-37.
[22] Deng, Y., Shan, Y., Gong, Z., & Chen, L, Large-scale navigation method for autonomous mobile robot based on fusion of GPS and lidar SLAM, *Chinese Automation Congress (CAC), IEEE*, pp. 3145-3148, 2018.
[23] Guo-Sheng, Huei-Yung Lin, and Shih-Fen Kao, Mobile robot localization using GPS, IMU and visual odometry, *International Automatic Control Conference (CACS), IEEE*, 2019.
[24] T. L. Bui, P. T. Doan, S. S. Park, H. K. Kim, and S. B. Kim, AGV trajectory control based on laser sensor navigation, *International Journal of Science and Engineering*, vol. 4, no. 1, pp. 16–20, 2013.
[25] S. Pratama, T. H. Nguyen, H. K. Kim, D. H. Kim, and S. B. Kim, Positioning, and obstacle avoidance of automatic guided vehicle in partially known environment," *International Journal of Control, Automation and Systems*, vol. 14, no. 6, pp. 1572–1581, 2016.

ICATSD2F.106

A STUDY ON DIRECT AND SIMULTANEOUS ANALYSIS OF ^{238}U AND ^{226}Ra NUCLIDES

VO XUAN AN

Industrial University of Ho Chi Minh City, Vietnam

voxuanan@iuh.edu.vn

Abstract. In the ^{238}U decay series, there are two interesting nuclides, which are ^{238}U and ^{226}Ra. The direct analysis of the activities of these two nuclides by gamma spectrometry meets some difficulties. Therefore, the indirect method by analyzing the daughter nuclides ^{214}Pb and ^{214}Bi is used. The disadvantage of this method is that the sample confinement technique must be applied for reaching the equilibrium between the activities of ^{214}Pb and ^{214}Bi and those of ^{226}Ra. We suggest a directly and simultaneously analytical method of the activities of ^{238}U and ^{226}Ra by experimental measurements and calculations using the Monte Carlo MCNP5 code. The activities of ^{238}U are determined by the 63.3 keV peak emitted from ^{234}Th, while those of ^{226}Ra nuclide are determined by the 186.2 keV peak. However, the 186.2 keV peak is interfered by the 185.7 keV peak emitted from ^{235}U. These two peaks cannot be separated by modern gamma spectrometries. The elimination of 185.7 keV peak from the 186 keV sumpeak is carried out using the 63.3 keV peak. As a result, the use of the three peaks 63.3 keV, 185.7 keV and 186.2 keV makes it possible to determine the activities of 238U directly and simultaneously and ^{226}Ra. In this work, the above-mentioned analytical method applied for the IAEA RGU-1 reference radioactive material is carried out.

Keywords. gamma ray spectrometry, HPGe detector, activity, ^{238}U, ^{226}Ra.

1 INTRODUCTION

For many years, environmental radiation research has always been a topical issue of interest to physicists because it directly affects human health and life. It is very necessary to research and propose methods of measuring accurate radioactivity. There have been scientific research works related to this field, from the experimental methods of the radioactive determination on gamma ray and alpha particle spectrometries [1, 2] to the calculation methods using the Monte Carlo simulation codes [3, 4]. In nature, there are three radioactive decay series starting from the nuclides ^{235}U, ^{238}U and ^{232}Th. Since the abundance of ^{235}U and ^{238}U is kept in balance with the mass relation m(^{235}U)/m(^{238}U) = 0.714%, or the radioactive relation a(^{235}U)/a(^{238}U) = 0.0462, the radioactive determination of ^{235}U can be done through the ^{238}U nuclide. Of the two remaining decay series ^{238}U and ^{232}Th, the ^{232}Th decay series achieves an equilibrium between their progeny nuclides, so this decay series has no interesting features. For the ^{238}U decay series, the equilibrium between their progeny nuclides is disrupted for several reasons. Therefore, the study of the properties of the ^{238}U decay series is an interesting problem for the nuclear physicists. The ^{238}U decay series occurs as follows:

$$_{92}U^{238} \rightarrow {}_{90}Th^{234} \rightarrow {}_{91}Pa^{234m} \rightarrow {}_{91}Pa^{234} \rightarrow {}_{92}U^{234} \rightarrow {}_{90}Th^{230} \rightarrow {}_{88}Ra^{226} \rightarrow {}_{86}Rn^{222} \rightarrow {}_{84}Po^{218} \rightarrow {}_{85}At^{218} \rightarrow {}_{86}Rn^{218} \rightarrow {}_{82}Pb^{214} \rightarrow {}_{83}Bi^{214} \rightarrow {}_{84}Po^{214} \rightarrow {}_{81}Tl^{210} \rightarrow {}_{82}Pb^{210} \rightarrow {}_{83}Bi^{210} \rightarrow {}_{84}Po^{210} \rightarrow {}_{82}Pb^{206}$$

In this decay series, the ^{238}U nuclide has a very large half-life ($T_{1/2}$ = 4.56×10^9 years) compared to its daughter nuclides, so there is a secular equilibrium between all nuclides. Thus, in principle, the radioactivity of all nuclides must be the same, and to measure the activity of ^{238}U, one can measure the radioactivity of any daughter nuclide, the most common being the radio activities of ^{214}Pb ($T_{1/2}$ = 26.8 minutes) and ^{214}Bi ($T_{1/2}$ = 19.9 minutes) due to their large yield. However, there are two disequilibria between the nuclides in this decay series. The first is a geochemical disequilibrium between ^{238}U and ^{226}Ra ($T_{1/2}$ = 1620 years) because these two elements have different chemical properties, and the second is a disequilibrium between ^{226}Ra and its daughter nuclides because ^{222}Rn ($T_{1/2}$ = 3.825 days) is an inert gas that is highly emanated

from the sample. Thus, the radioactive measurement of the nuclides ^{214}Pb and ^{214}Bi has not yet resulted in the activity of ^{226}Ra. The sample confinement technique is used to create an equilibrium between ^{222}Rn and ^{226}Ra, and through the activities of the ^{214}Pb and ^{214}Bi nuclides, the ^{226}Ra radioactivity can be inferred. However, the sample sealing is a major disadvantage of this indirect measurement method. In addition, the geochemical disequilibrium between ^{226}Ra and ^{238}U does not allow inferring the activity of ^{238}U from that of ^{226}Ra. Thus, the first interesting problem is to find a method to analyze two nuclides ^{238}U and ^{226}Ra, independently, to study the disequilibrium between the activities of these two nuclides. The second problem is to find a method for directly analyzing the ^{226}Ra nuclide to avoid the disadvantage of indirect analysis due to sample confinement. The above two problems were solved in the works [5, 6]. In this work, we suggest a new method to analyze directly and simultaneously two nuclides ^{238}U and ^{226}Ra through experimental measurements on gamma spectrometry and calculations using the Monte Carlo MCNP5 code.

2 METHODOLOGY

2.1 Principles of the analytical method

In the method of direct and simultaneous analysis of two nuclides ^{238}U and ^{226}Ra, three photopeaks (or peaks) 63.3 keV, 185.7 keV, and 186.2 keV and a sumpeak of 186 keV are considered as shown in Figure 1. The gamma ray of 63.3 keV with a yield of 4.49% is emitted from the ^{234}Th nuclide ($T_{1/2}$ = 24.1 days), which is a daughter nuclide directly emitted from the ^{238}U nuclide. The gamma ray of 185.7 keV is emitted from the ^{235}U nuclide ($T_{1/2} = 7.1\times10^{8}$ years) with a yield of 57.5%. The gamma ray of 186.2 keV with a yield of 3.5% is directly emitted from the ^{226}Ra nuclide. The two peaks 185.7 keV and 186.2 keV are superposed without separation on the modern gamma spectrometries using the HPGe detectors. These two peaks convolve into a 186 keV sumpeak on the gamma spectrum.

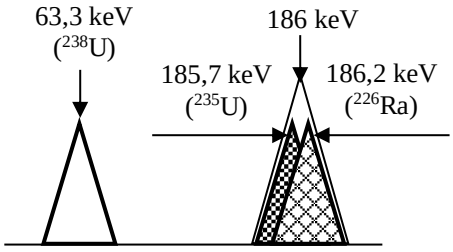


Figure 1. The photopeaks used in the radioactive measurement of the ^{238}U and ^{226}Ra nuclides.

In the three peaks mentioned above, the peak of 63.3 keV is used to analyze the activity of the ^{238}U nuclide. The gamma rays of 63.3 keV are emitted from the ^{234}Th nuclide, but it is very short-lived, so its activity can be considered the activity of ^{238}U. The activity of ^{226}Ra is determined by the peak area of 186.2 keV, which is the difference between the sumpeak area of 186 keV and the peak area of 185.7 keV. The areas of the two peaks 63.3 keV and 186 keV are experimentally determined on the gamma spectrometry while the peak area of 185.7 keV is calculated from the peak area of 63.3 keV. The Monte Carlo MCNP5 code is used to calculate the photopeak efficiencies, thereby inferring the activities of the respective nuclides.

2.2 Carrying out the analysis

Based on the above principle, the radioactive determination of ^{238}U and ^{226}Ra is carried out according to the following steps:

a) Determining the 63.3 keV peak area via the gamma spectrum measured by the gamma spectrometry. Calculating this peak efficiency using the MCNP5 code and determining the activity of ^{238}U.

b) Determining the 186 keV sumpeak area via the gamma spectrum measured by the gamma spectrometry. This sumpeak is a superposition of two peaks 185.7 keV of ^{235}U and 186.2 keV of ^{226}Ra.

c) Calculating peak efficiency of 185.7 keV using the MCNP5 code. Combining this efficiency with the activity of ^{238}U in the step (a) to determine the peak area of 185.7 keV.

d) Calculating the peak area of 186.2 keV as the difference between the sumpeak area of 186 keV in the step (b) and the peak area of 185.7 keV in the step (c).

e) Calculating the peak efficiency of 186.2 keV. Combining this efficiency with the peak area of 186.2 keV in the step (d) to determine the activity of ^{226}Ra.

3 RESULTS OF RADIOACTIVE ANALYSIS OF THE ^{238}U AND ^{226}Ra NUCLIDES IN THE IAEA RGU-1 STANDARD SAMPLE

The determination of the ^{238}U activity using the gamma ray of 63.3 keV by experiments was presented in the work [7]. The main difficulty of using the gamma ray of 63.3 keV is that it is strongly absorbed in the sample. In the work [7], the correction for self-absorption effect, geometric effect and density effect is completely carried out by experiments. In this work, those effects will be calculated by the Monte Carlo MCNP5 code. Calculating these effects results in the photopeak efficiency for a given gamma ray energy emitted from a nuclide with known yield. The experimental measurements were carried out on a gamma spectrometry using the HPGe GC1518 detector located at the Center for Nuclear Techniques in Ho Chi Minh City. The input data of this spectrometry system for the MCNP5 code has been presented in the works [8, 9, 10] and shows in the appendix below. Among the input data, the most important parameter is the dead layer thickness of the detector. According to the works [9, 10], the dead layer had a thickness of 0.35 mm in 1996, 0.65 mm in 1999, 1.15 mm in 2005, 1.46 mm in 2009, and 1.52 mm at present.

This section shows results of radioactivity analysis of ^{238}U and ^{226}Ra nuclides in the standard sample IAEA RGU-1 [11] via the peaks of 63.3 keV and 186 keV by the experimental measurements combined with the MCNP5 calculations. This sample contains the ^{238}U nuclide with specific activity 4940 Bq/kg (4920 - 4970 Bq/kg) and the ^{235}U nuclide with specific activity 228 Bq/kg (226 - 230 Bq/kg). Other nuclides with very small activity can be ignored. The layout of the lead shielding, the HPGe GC1518 detector and the sample container is shown in Figure 2. Figure 3 illustrates the sample container shaped a Marinelli beaker. The standard sample IAEA RGU-1 with a mass of 0.740 kg was kept in a hermatically sealed Marinelli beaker over one month and measured on the gamma spectrometry using the HPGe GC1518 detector. Thus, the radioactivities of ^{214}Pb and ^{214}Bi have reached equilibrium with the activity of ^{226}Ra. On the other hand, the IAEA RGU-1 standard sample reached geochemical equilibrium, so the activity of ^{226}Ra being also the same activity of ^{238}U is equal to 4940 Bq/kg.

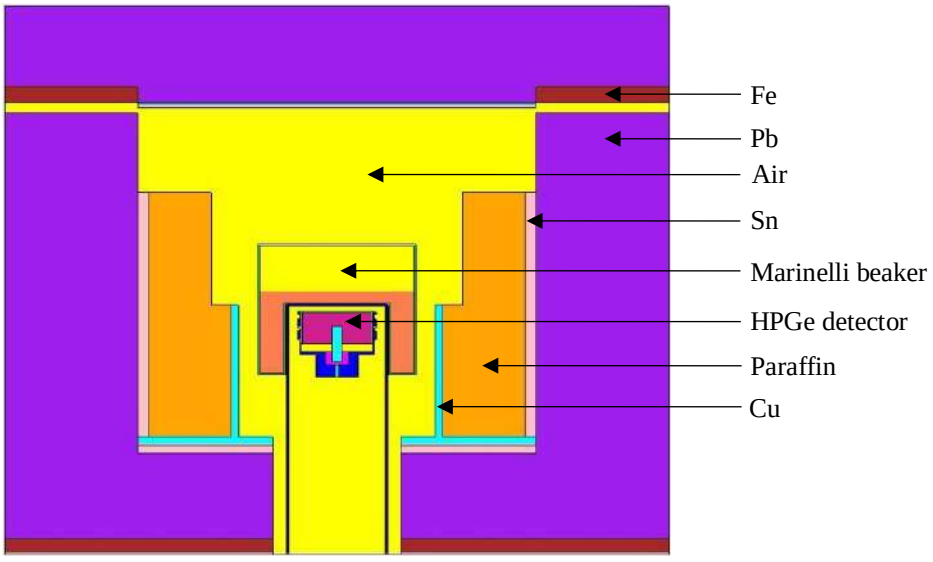


Figure 2. A vertical cross-sectional view of MCNP5 modeled lead shielding geometry with lead, tin, copper, and paraffin shieldings, a Marinelli beaker, and a HPGe detector.

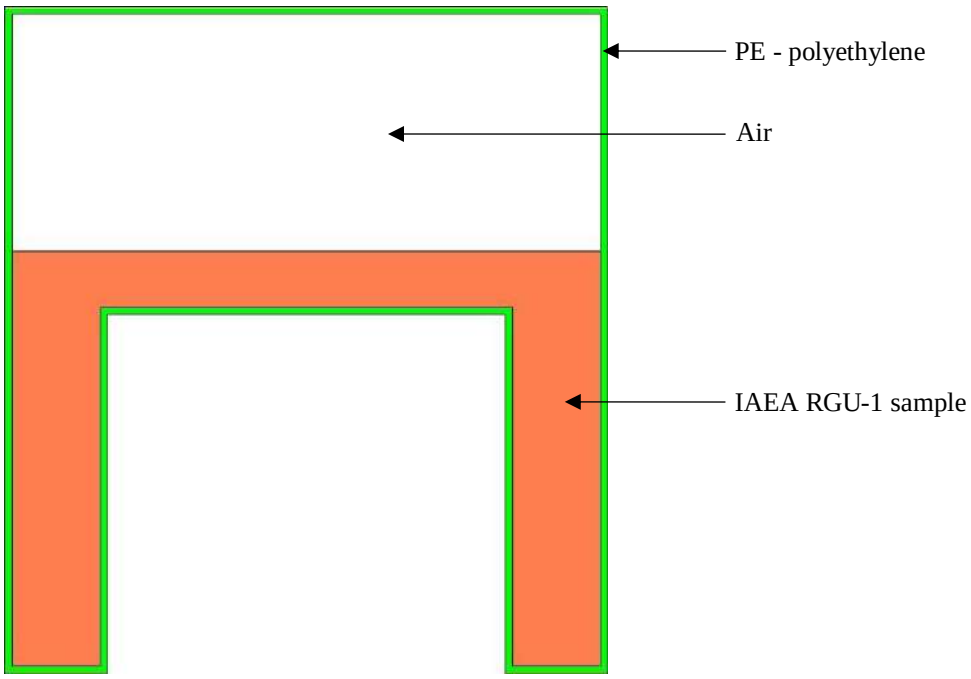


Figure 3. A vertical cross section of MCNP5 modeled Marinelli beaker geometry with a plastic container made of polyethylene and IAEA RGU-1 sample.

The full gamma spectrum of the standard sample IAEA RGU-1 measured on a gamma spectrometry using the HPGe GC1518 detector is shown in Figure 4, in which the full gamma spectrum contains the main photopeaks of the nuclides ^{238}U, ^{235}U, ^{226}Ra, ^{214}Pb, and ^{214}Bi. Figure 5 extracts a part of the spectrum containing the 63.3 keV peak of ^{238}U and the 186 keV sumpeak of ^{235}U and ^{226}Ra. The measurement and calculation data for the photopeaks of 63.3 keV, 185.7 keV, 186.2 keV, 186 keV (including the peak areas of 185.7 keV and 186.2 keV), 295 keV, 352 keV, 609 keV, and 1121 keV are presented in Table 1. The areas of the full energy peaks $S_{fep,\,exp}$ for the above photopeaks are shown in column 2 of Table 1 with their relative errors. Column 3 of Table 1 shows the photopeak efficiencies calculated by the MCNP5 code for the IAEA RGU-1 sample contained in the Marinelli beaker (Figure 3) with the sample density of 1.5353 g/cm^3 and the detector dead layer thickness of 1.52 mm at present. From these data, the specific activity a (Bq/kg) of the nuclides can be calculated by the formula [12]:

$$a\ (\text{Bq/kg}) = \frac{S_{fep,\,exp}}{T\varepsilon_{cal}m} \qquad (1)$$

where T = 3600 s and m = 0.740 kg. The calculation results of the activity are quoted in column 4 of Table 1. According to the measurement and calculation results shown in Table 1, it can be recognized:
a) For the 63.3 keV photopeak of ^{238}U, calculating according to the formula (1) results in the specific activity of 5002 ± 300 Bq/kg are consistent with the reference activity of 4940 Bq/kg within the margin of error.
b) For the 186 keV sumpeak of ^{235}U and ^{226}Ra, it has the area of S (186 keV) = 23360, including the two peak areas of S (^{235}U,185.7 keV) and S (^{226}Ra,186.2 keV).
The peak area of S (^{235}U,185.7 keV) is calculated by the following formula:

$$S\ (^{235}\text{U},185.7\ \text{keV}) = \varepsilon_{cal}\,(185.7\ \text{keV}) \times a(^{238}\text{U}) \times T \times m \qquad (2)$$

$$= 0.0007666 \times 5002 \times 3600 \times 0.74 \approx 10215$$

The peak area of S (^{226}Ra,186.2 keV) is equal to:

$$S\ (^{226}\text{Ra},186.2\ \text{keV}) = S\ (186\ \text{keV}) - S\ (^{235}\text{U},185.7\ \text{keV}) \qquad (3)$$

= 23360 - 10215 = 13145

The activity of the ^{226}Ra nuclide (186.2 keV) is calculated by the formula (1) and equal to 4913 ± 295 Bq/kg, coinciding with the reference activity of 4940 Bq/kg within the margin of error.

c) The photopeaks 295 keV (^{214}Pb), 352 keV (^{214}Pb), 609 keV (^{214}Bi), and 1121 keV (^{214}Bi) calculated through the formula (1) have the specific activities of 5056 ± 303 Bq/kg, 5208 ± 312 Bq/kg, 4757 ± 285 Bq/kg and 4600 ± 276 Bq/kg, respectively, also agreeing with the reference activity of 4940 Bq/kg within the margin of error.

Table 1. The measured peak area - $S_{fep, exp}$, the MCNP5 calculation efficiency - ε_{cal}, the specific activity - a (Bq/kg) and the reference activity (Bq/kg) [7] for the photopeaks 63.3 keV (^{238}U), 185.7 keV (^{235}U), 186.2 keV (^{226}Ra), 295 keV (^{214}Pb), 352 keV (^{214}Pb), 609 keV (^{214}Bi), and 1121 keV (^{214}Bi).

Energy (keV)	$S_{fep, exp}$ (relative error)	ε_{cal}	a (Bq/kg)	Reference activity (Bq/kg)
63.3 (^{238}U)	5429 (5.79%)	0.0004070	5002 ± 300	4940
186 (^{235}U+^{226}Ra)	23360 (1.48%)			
185.7 (^{235}U)	10215	0.0007666	231 ± 14	228
186.2 (^{226}Ra)	13145	0.0010041	4913 ± 295	4940
295 (^{214}Pb)	44902 (0.66%)	0.0033336	5056 ± 303	4940
352 (^{214}Pb)	74241 (0.49%)	0.0053505	5208 ± 312	4940
609 (^{214}Bi)	50317 (0.57%)	0.0039698	4757 ± 285	4940
1121 (^{214}Bi)	9604 (1.9%)	0.0007836	4600 ± 276	4940

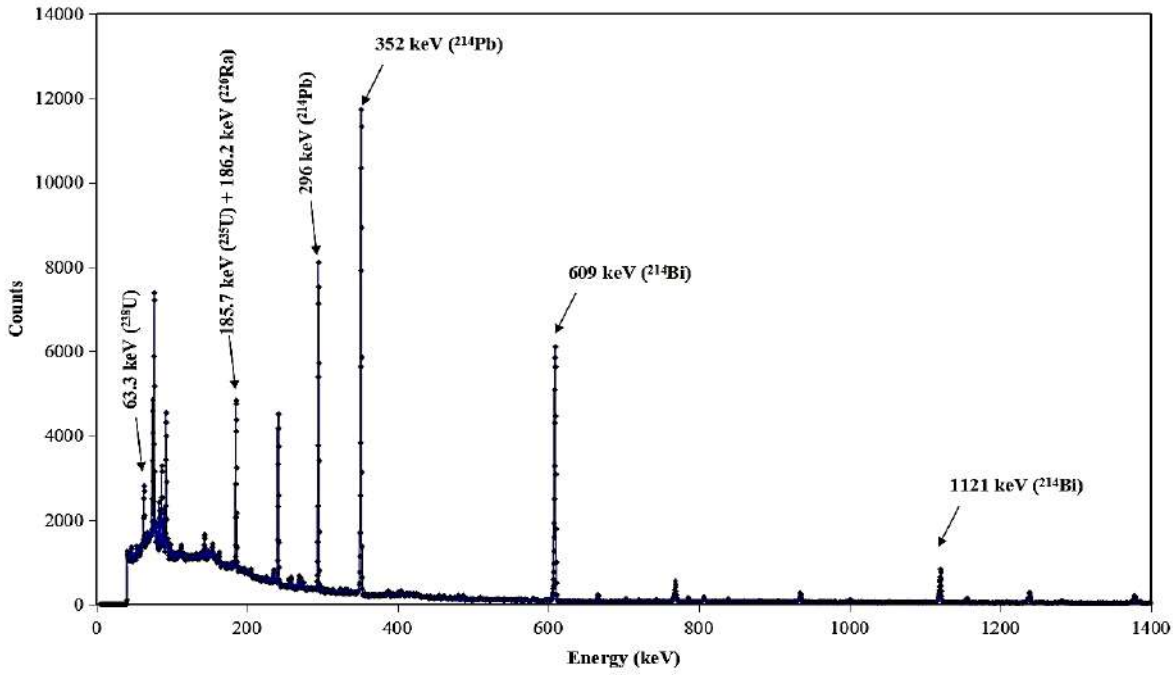


Figure 4: The full gamma spectrum of the IAEA RGU-1 standard sample measured on a gamma spectrometry using the HPGe GC1518 detector, which consists of the photopeaks of ^{238}U (63.3 keV), ^{235}U+^{226}Ra (186 keV), ^{214}Pb (295 keV and 352 keV), and ^{214}Bi (609 keV and 1121 keV).

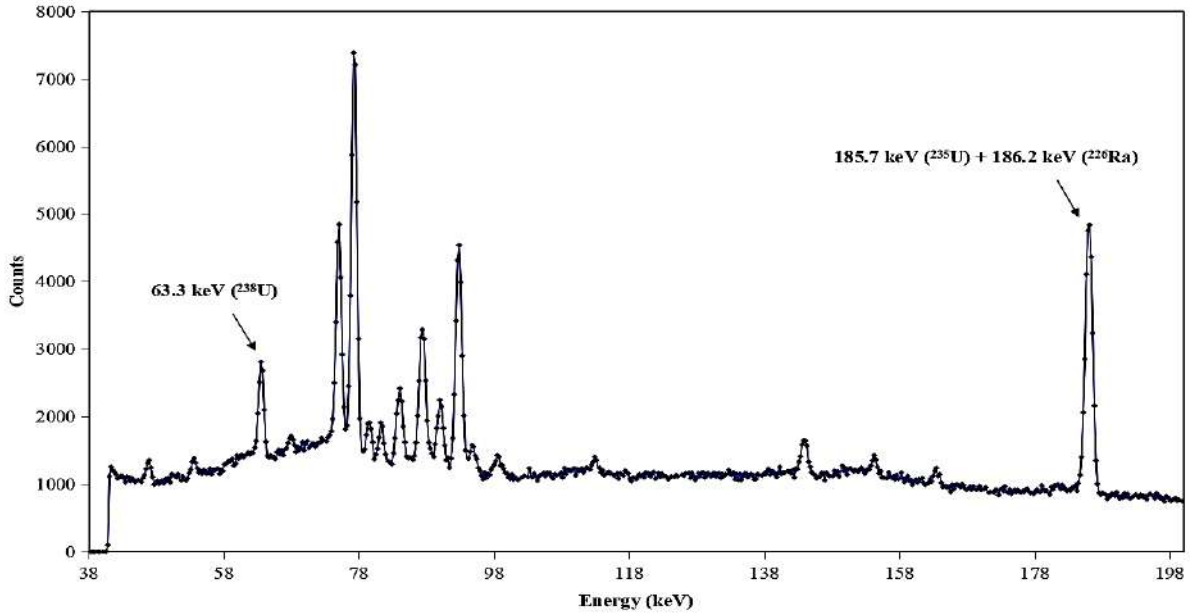


Figure 5: Part of the gamma spectrum of the IAEA sample RGU-1, which only consists of the 63.3 keV photopeak of ^{238}U and the 186 keV sumpeak of ^{235}U and ^{226}Ra.

4 CONCLUSIONS

This work shows the results of the method of direct and simultaneous analysis of the ^{238}U and ^{226}Ra nuclides by experimental measurements of the gamma ray spectrometry using the HPGe GC1518 detector combined with calculations using the Monte Carlo MCNP5 code. The Monte Carlo simulations with the carefully prepared MCNP5 input have remarkably contributed to correct the effects of self-absorption, geometry, and density, and allowing to calculate the predicted peak areas. This method is applicable to the reference standard sample IAEA RGU-1 containing nuclides ^{238}U, ^{235}U and ^{226}Ra. The specific activities of these nuclides, obtained by this method, are quite consistent with the reference activities.

APPENDIX

A standard input file of the MCNP5 code is as follows:
Problem - A study on direct and simultaneous analysis of ^{238}U and ^{226}Ra nuclides
c Cell cards
1 6 -8.94 (-2 -34 27) imp:p,e=1 $ loi Cu dan tin hieu
2 15 -2.31 (-3 -35 34): (2 -3 -34 30) imp:p,e=1 $ lop boron khuech tan
3 1 -5.35 (-6 -38 35): (3 -6 -35 30) imp:p=1 imp:e=0 $ tinh the HPGe
4 17 -5.05 (-7 -39 38): (6 -7 -38 30) imp:p,e=1 $ lop Li khuech tan
5 2 -2.6989 ((7 -9 -39 37): (7 -8 -37 36):(7 -9 -36 33) &
:(7 -8 -33 32): (7 -9 -32 31): (7 -8 -31 29): (4 -8 -29 28) &
:(4 -5 -28 26): (1 -5 -26 25) # (-52 -29 28) imp:p,e=1 $ holder bang Al
6 3 -0.00129 ((-10 -42 41): (9 -10 -41 37): (8 -10 -37 36) &
:(9 -10 -36 33): (8 -10 -33 32): (9 -10 -32 31):(8 -10 -31 28) &
:(5 -10 -28 27): (5 -10 -27 25))#(-52 -28 27) &
:(-10 -25 20) imp:p,e=1 $ khoang kk detector
7 2 -2.6989 (-11 -43 42): (10 -11 -42 20) imp:p,e=1 $ vo Al
8 20 -1.11 -9 -41 40 imp:p,e=1 $ lop kapton
9 18 -1.435 -9 -40 39 imp:p,e=1 $ lop mylar
10 4 -0.92 (51 -53 -30 29): (51 -52 -29 28) &
:(51 -52 -28 27) imp:p,e=1 $ vo cach dien in/out
11 3 -0.00129 -51 -30 27 imp:p,e=1 $ loi day dan in/out

```
12   3 -0.00129 (2 -7 -30 29) #1# (-53 -30 29) imp:p,e=1 $ khoang chan khong
13   21 -2.2 (2 -4 -29 27): (-4 -27 26) imp:p,e=1 $ coc teflon
14   6 -8.94 -1 -26 25 imp:p,e=1 $ que dan lanh bang Cu
c ==> nguon Marinelli
15   3 -0.00129 ((17 -18 -4801 44): (-17 -4401 4301): (-1502 -4301 43) &
:(11 -15 -43 24): (11 -14 -24 20) #27#28#29#30 imp:p,e=1 $ khoang kk buong chi
16   6 -8.94 (15 -1501 -43 24): (14 -17 -24 23) imp:p,e=1 $ lop Cu
1601 7 -0.88 (1502 -16 -4301 43): (1501 -16 -43 24) imp:p,e=1 $ lop paraffin
17   8 -7.28 16 -17 -4301 24 imp:p,e=1 $ lop Sn
1701 8 -7.28 14 -17 -23 22 imp:p,e=1 $ lop Sn
18   9 -11.34 (17 -18 -44 22): (14 -18 -22 21) imp:p,e=1 $ lop Pb
19   10 -7.86 17 -18 -50 4801 imp:p,e=1 $ lop Fe
1901 10 -7.86 14 -18 -21 20 imp:p,e=1 $ lop Fe
20   9 -11.34 (-18 -5001 50): (-17 -50 4801) imp:p,e=1 $ lop Pb
22   6 -8.94 -17 -4801 48 imp:p,e=1 $ lop Cu
2201   8 -7.28 -17 -48 4401 imp:p,e=1 $ lop Sn
24   0 18:5001: -20 imp:p,e=0
c ==> nguon Marinelli
27   4 -0.92 (-62 -68 67): (61 -62 -67 64):(59 -62 -64 63) &
:(59 -60 -43 64): (-60 -65 43) imp:p,e=1 $ thanh hop M
28   3 -0.00129 -61 -67 66 imp:p,e=1 $ khoang kk trong hop M
29   26 -1.535327045 -61 -66 65 imp:p,e=1 $ phan chat phong xa upper trong hop M
30   26 -1.535327045 60 -61 -65 64 imp:p,e=1 $ phan chat phong xa lower trong hop M

c Surface cards
c ==> mo ta tu trong ra ngoai
1    cz 0.13 $ que dan lanh bang Cu
2    cz 0.35 $ dien cuc bang Cu
3    cz 0.3503 $ lop boron khuech tan
4    cz 0.88 $ mat ngoai lop teflon
5    cz 1.546 $ mat ngoai lop Al bac 1 cua holder
6    cz 2.545 $ lop Li khuech tan
7    cz 2.7 $ ban kinh tinh the Ge
8    cz 2.776 $ mat ngoai lop Al bac 2 cua holder
9    cz 2.97 $ mat ngoai lop Al bac 3 cua holder
10   cz 3.66 $ mat trong vo Al
11   cz 3.81 $ mat ngoai vo Al
14   cz 4.81 $ mat trong lop Pb 1
15   cz 7.35 $ mat trong lop Cu 2
1501 cz 7.95 $ mat trong lop paraffin 1
1502 cz 9.45 $ mat trong lop paraffin 2
16   cz 14.2 $ mat trong lop Sn 2
17   cz 15.0 $ mat trong lop Pb 2
18   cz 25.0 $ mat ngoai lop Pb 2
c ==> mo ta tu duoi len tren
20   pz 0 $ mat duoi lop Fe 1
21   pz 1.6 $ mat tren lop Fe 1
22   pz 10.0 $ mat tren lop Pb 1
23   pz 10.8 $ mat tren lop Sn 1
24   pz 11.6 $ mat tren lop Cu 1
c ==> ben trong detector
25   pz 17.536 $ mat duoi holder Al 2
26   pz 18.696 $ mat duoi coc teflon
```



```
27   pz 19.042 $ mat duoi loi Cu
28   pz 19.742 $ mat duoi holder Al 1
29   pz 20.062 $ mat duoi lop chan khong
30   pz 20.815 $ mat duoi lop Ge
31   pz 21.255 $ mat duoi vanh holder 3
32   pz 22.115 $ mat tren vanh holder 3
33   pz 22.415 $ mat duoi vanh holder 2
34   pz 22.515 $ mat tren loi Cu
35   pz 22.5153 $ mat tren lop boron khuech tan
36   pz 23.275 $ mat tren vanh holder 2
37   pz 23.715 $ mat duoi vanh holder 1
38   pz 23.86 $ mat tren lop Ge
39   pz 24.015 $ mat tren lop Li khuech tan
40   pz 24.016 $ mat tren mylar
41   pz 24.026 $ mat tren kapton
42   pz 24.55 $ mat tren khong khi
43   pz 24.7 $ mat tren vo Al detector
c ==> phan tren buong chi
4301 pz 35.8 $ mat tren lop Sn 2 TTHN
44   pz 43.6 $ mat tren lop Pb 2
4401 pz 44.1 $ mat duoi lop Sn 2
48   pz 44.5 $ mat tren lop Sn 3
4801 pz 44.6 $ mat tren lop Cu 3 TTHN
50   pz 46.2 $ mat tren lop Fe 2
5001 pz 54.2 $ mat tren lop Pb 3 TTHN
c ==> day dan in/out
51   c/z 0 -2.1 0.1 $ loi day dan in/out
52   c/z 0 -2.1 0.2 $ lop cach dien day dan in/out 1
53   c/z 0 -2.1 0.49 $ lop cach dien day dan in/out 2
c ==> nguon Marinelli
59   cz 3.9051 $ mat ben trong hop M in
60   cz 4.0451 $ mat ben trong hop M out
61   cz 5.7751 $ mat ben ngoai hop M in
62   cz 5.9151 $ mat ben ngoai hop M out
63   pz 17.77 $ mat day duoi hop M lower
64   pz 17.91 $ mat day duoi hop M upper
65   pz 24.84 $ mat day tren hop M upper
66   pz 25.91 $ mat chat phong xa long
67   pz 30.49 $ mat nap hop M lower
68   pz 30.63 $ mat nap hop M upper

c Data cards
mode p
m1   32000 -1.0 $ Ge
m2   13000 -1.0 $ Al
m3   7000 -0.755 8000 -0.232 18000 -0.013 $ atmosphere
m4   1000 -0.14372 6000 -0.85628 $ polyethylene
m6   29000 -1.0 $ Cu
m7   1000 -0.1549 6000 -0.8451 $ paraffin C9H20
m8   50000 -1.0 $ Sn
m9   82000 -1.0 $ Pb
m10 26000 -1.0 $ Fe
m15 5000 -1.0 $ B
```

```
m17 32000 -0.9999 3000 -0.0001 $ Ge Li
m18 1000 -0.053 6000 -0.526 8000 -0.421 $ mylar C10H12O6
m20 1000 -0.028 6000 -0.720 7000 -0.077 8000 -0.175 $ kapton C22H10N2O4
m21 6000 -0.24 9000 -0.76 $ teflon (C2H4) n
m26 13000 -0.02593000 33000 -0.00000001 56000 -0.00000001 &
    35000 -0.00000001 20000 -0.00486000 48000 -0.00000001 &
    17000 -0.00000001 27000 -0.00000001 24000 -0.00000001 &
    55000 -0.00000001 29000 -0.00000001 63000 -0.00000001 &
    26000 -0.04739000 80000 -0.00000001 19000 -0.00863000 &
    57000 -0.00000001 12000 -0.00296000 25000 -0.00039000 &
    42000 -0.00000001 11000 -0.00356000 28000 -0.00000001 &
    15000 -0.00044000 82000 -0.00000001 37000 -0.00000001 &
    16000 -0.00000001 51000 -0.00000001 21000 -0.00000001 &
    34000 -0.00000001 38000 -0.00000001 30000 -0.00000001 &
    6000 -0.00515000 1000 -0.00000001 8000 -0.50714977 &
    14000 -0.38940000 22000 -0.00414000 $ IAEA RGU-1
c ==> nguon Marinelli
sdef cel=d1 pos=0 0 0 axs=0 0 1 ext fcel d2 rad fcel d5 erg=d8 par=2 wgt=10
si1 l 30 29
sp1 v
ds2 s d3 d4
si3 h 17.91 24.84
sp3 d -21 0
si4 h 24.84 26.41
sp4 d -21 0
ds5 s d6 d7
si6 h 4.0451 5.7751
sp6 -21 1
si7 h 0.0000 5.7751
sp7 -21 1
si8 s d9 d10 d11
sp8 d 1.0 1.0 1.0
#    si9         sp9       $ ra-226
     l           d
     0.186101    0.0350
     0.002       0.9650
#    si10        sp10      $ pb-214
     l           d
     0.241982    0.0750
     0.295214    0.1850
     0.351922    0.3580
     0.002       0.3820
#    si11        sp11      $ bi-214
     l           d
     0.609313    0.4480
     0.768356    0.0480
     0.934061    0.0303
     1.120287    0.1480
     1.238110    0.0586
     1.377669    0.0392
     1.407984    0.0280
     1.729595    0.0288
     1.764494    0.1536
```



```
    0.002       0.0175
f8:p 3
e8   0 .0001 .005616 8190i 1.941030
phys:p $ produce bremsstrahlung radiations
phys:e
cut:p 2j 0 0 $ because of taking a tally of pulse height distributions
cut:e
nps 9000000
ctme 60
```

REFERENCES

[1] I. Vukanac, B. Seslak, A. Kandic, M. Durasevic, A comparison of alpha-particle and gamma-ray spectrometry methods for determination of 235U, 238U and 226Ra activity concentration in samples of coal, slag and fly-ash, *Radiation Physics and Chemistry*, vol. 193, 109933, 2022.

[2] A. Angileri, P. Sardini, D. Beaufort, M. Descostes, Mobility of daughter elements of 238U decay chain during leaching by In Situ Recovery (ISR): New insights from digital autoradiography, *Journal of Environmental Radioactivity*, vol. 221, 106274, 2020.

[3] N. Mezerreg, A. Azbouche, and M. Haddad, Study of coincidence summing effect using Monte Carlo simulation to improve large samples measurement for environmental applications, *Journal of Environmental Radioactivity*, vol. 232, 106573, 2021.

[4] S. Landsbergera, A. Sharpa, S. Wanga, Y. Pontikesb, and A.H. Tkaczyk, Characterization of bauxite residue (red mud) for ^{235}U, ^{238}U, ^{232}Th and ^{40}K using neutron activation analysis and the radiation dose levels as modeled by MCNP, *Journal of Environmental Radioactivity*, vol. 173, pp. 97-101, 2017.

[5] M. Dowdall, O.G. Senas, J.P. Gwynn, and C. Davids, Simultaneous determination of ^{226}Ra and ^{238}U in soil and environmental materials by gamma-spectrometry in the absence of radium progeny equilibrium, *Journal of Radioanalytical and Nuclear Chemistry*, vol. 261, pp. 513-521, 2004.

[6] F.D. Corte, H. Umans, D. Vandergerghe, A.D. Wispelaere, and P.V. Haute, Direct gamma-spectrometric measurement of the ^{226}Ra 186.2 keV line detecting ^{238}U/^{226}Ra disequilibrium in determining the environmental dose rate for the luminescence dating of sediments, *Applied Radiation and Isotopes*, vol. 63, pp. 589-598, 2005.

[7] N.Q. Huy and T.V. Luyen, A method to determine ^{238}U activity in environmental soil samples by using 63.3 keV-photopeak-gamma HPGe spectrometer, *Applied Radiation and Isotopes*, vol. 61, pp. 1419-1424, 2004.

[8] N.Q. Huy, D.Q. Binh, and V.X. An, Modeling the gamma spectrometry using the hyper pure germanium by the Monte Carlo MCNP code (in Vietnamese), *Science & Technology Development Journal, VNUHCM*, vol. 8, no. 8, pp. 17-25, 2005.

[9] N.Q. Huy, D.Q. Binh and V.X. An, Study on the increase of inactive germanium layer in a high purity germanium detector after a long-time operation applying MCNP code, *Nuclear Instruments and Methods in Physics Research A*, vol. 573, pp. 384-388, 2007.

[10] N.Q. Huy, The influence of dead layer thickness increase on efficiency decrease for a coaxial HPGe p-type detector, *Nuclear Instruments and Methods in Physics Research A*, vol. 621, pp. 390-394, 2010.

[11] IAEA, Analytical Quality Control Services Reference Materials Catalogue 2004-2005, IAEA: Vienna, Austria, 2004.

[12] N.N. Jibiri, I.P. Farai and S.K. Alausa, Estimation of annual effective dose due to natural radioactive elements in ingestion of foodstuffs in tin mining area of Jos-Plateau, Nigeria, *Journal of Environmental Radioactivity*, vol. 94, pp. 31-40, 2007.

ICATSD2F.107

TELEOPERATION OF A CAR-LIKE MOBILE ROBOT WITH HAPTIC JOYSTICK USING POTENTIAL FIELD BASED FORCE FEEDBACK

ANH KHOA TRAN, HUNG HOANG, DUC THIEN TRAN

Department of Automatic Control, Ho Chi Minh City University of Technology and Education, Ho Chi Minh, Vietnam

anhkhoatran.289@gmail.com, hunghoang.1806@gmail.com, thientd@hcmute.edu.vn

Abstract. This paper presents research about the teleoperation of a wheeled mobile robot using a haptic joystick with force feedback. Conventionally, through direct visual observation or an integrated camera, the teleoperation system allows the operator to control the mobile robot in the remote environment and perform complex tasks. Therefore, the control accuracy depends on how the operator perceives the environment's information, which may be insufficient for the suddenly appearing obstacles. The use of a two degree of freedom (DOF) haptic joystick is proposed to enhance the control performance by producing force feedback for the operator. First, the feedback force from the nearest obstacle is computed by using the potential field algorithm. Then, the force is executed by adjusting the stiffness of a virtual spring on each motor's axis, making the joystick back to its equilibrium position, and alerting the obstacle's proximity to the operator. The proposed teleoperation system has been verified on simulation and experiment to demonstrate the superiority of haptic control using force feedback in advancing the control performance.

Keywords. Haptic teleoperation, haptic device, potential field, mobile robot.

1 INTRODUCTION

In recent years, the robot field has developed strongly and rapidly, in which mobile robots occupy a significant market share. There are numbers of research on mobile robots' application in different fields like industrial, military, medical, etc. In these applications, mobile robots carry out tasks from simple to complex, such as transporting goods [1], rescue [2], assisting in rehabilitation for patients [3], building maps [4], etc. Therefore, it is needed to consider the mobile robot control methods' problem to enhance task performance. Three typical ways to control a mobile robot are automatic, direct, and remote control. Automatic robots can complete tasks without the operator's control using the control algorithms processing the environment information taken from sensors [5][6]. However, for the high-speed automatic requirement, the control algorithms must be advanced as the hardware are also powerful and reliable, which makes the investment cost higher. Direct control robots have a lower cost than automatic robots. Still, it requires the operator to stand nearby the robot, which may be unfeasible when operating in a hazardous environment.

The remote control uses wireless communication to transmit the control signals like steering angle, velocity from the operator to the robots, and send back the robot environment information to the operator. Therefore, the tasks can still be carried out while the robot is working in a hazardous and unstructured environment. However, the traditional remote-control method still faces many limitations due to the operator can only control the robot with a subjective opinion based on the information collected by the visual observation. Another disadvantage is the time delay in communication between the operator and the mobile robot. The operator cannot control flexibly and makes mistakes when performing tasks that require high accuracy. Several studies have been conducted to solve the above problems, such as predicting the environment to handle the problem of time delay in communication [7][8], using different techniques in sensing to increase environmental awareness [9], etc. Another outstanding method to solve the above problem is to combine Haptic control with the traditional remote-control method.

Haptic control is a tactile control method that lets the operator perceive the environment intuitively. The most basic and standard way to create a Haptic control is to use a linear function for generating the feedback force proportional to the distance between the mobile robot and the obstacle [10], … Besides that, there are

disadvantages, such as when using a linear function, the impact force will increase steadily with distance, and it will be difficult for the operator to precisely recognize the changing distance from the robot to the obstacle. Because of such disadvantages, in recent times, Haptic research based on the Potential Filed method is getting the attention of many researchers worldwide.

From the above analysis, this paper proposes a Haptic control solution using potential field-based force feedback integrated into the joystick with the traditional remote-control method for the Car-like mobile robot model containing nonholonomic constraints. A two-degree of freedom haptic joystick is built to facilitate experimentation. The virtual feedback force is calculated from the environmental information based on the potential field algorithm. From the virtual feedback force, the stiffness of the virtual spring is adjusted to create a real feedback force acting on the two axes of the joystick, helping the joystick to return to an equilibrium position, and then affecting the tactile sensation of the operator. This method will solve the problem of limited information from the environment and helps to increase the interaction between the operator and the robot. The control becomes more flexible, solving unexpected problems or complex tasks of mobile robots. The algorithm's effectiveness is proven by MATLAB simulation, experiments, and evaluations in practice.

The structure of this paper is the following. In section II, the system overview structure including joystick and mobile robot is presented. The proposed algorithm for Haptic control based on the Potential Field theory for the handle is described in section III. Section IV deals with the simulation model, experimental procedure and the algorithm test results in practice. Conclusions are presented in section V.

2 SYSTEM DESCRIPTION

The overview diagram of the proposed system is shown in Figure . The system consists of two main parts, the joystick is the master and the Car-like mobile robot is the slave. The operator acts on the joystick to create a control signal sent to the Car-like mobile robot. The robot moves according to the signal, and during operation, the sensor mounted on the robot's body will measure the distance to the nearest obstacle. The control and sensor signals are circulated between master and slave through wireless communication. The feedback sensor signals, and the joystick terminal’s position are used to create a feedback force on the joystick, thereby helping the operator perceive the distance to the obstacle. The structure and signal in Joystick and Car-like mobile robot are described below.

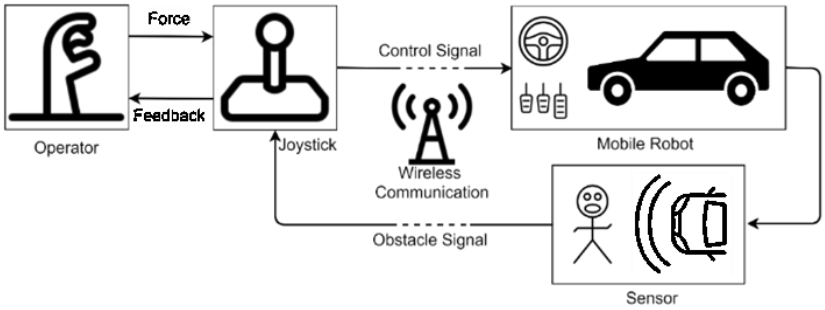
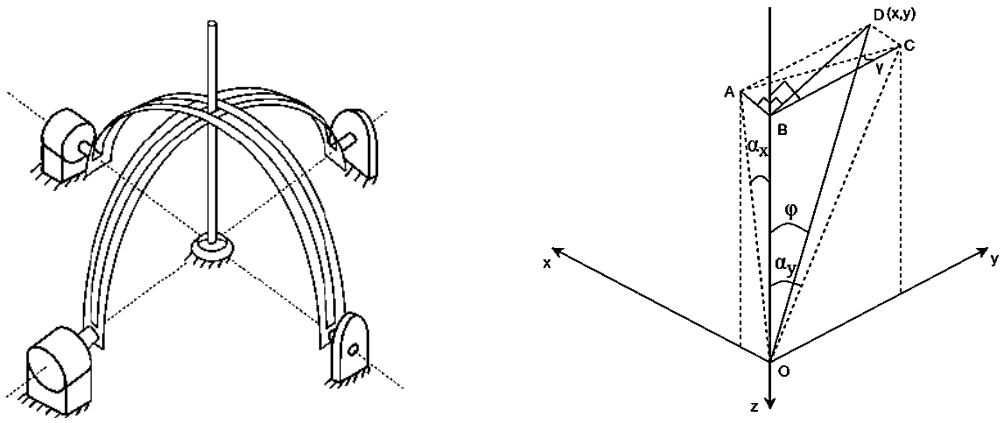


Figure 1. Haptic control system

2.1 Joystick model

The applicability of joysticks depends mainly on their design; the simpler the design, the lesser DOF of the joystick, leading to the control dimension limitation. The low DOF joystick, however, has wide applications than the high DOF joystick. Like gamepad used in the game industry usually contains 2-DOF joysticks for controlling, especially for racing games. Therefore, Figure presents a simple design of 2-DOF joystick for controlling a mobile robot model in a two-dimensions environment.



(a) (b)

Figure 2. Structure (a) and Coordinate system (b) of the 2-DOF joystick

The joystick model has two degrees of freedom controlled by two motors and attached to the base via a spherical joint. The concentric circular mechanical structure connects two motor shafts and the joystick. The intersection of the two motor shafts coincides with the center of the spherical joint. Therefore, the spherical joint rotates when the joystick terminal moves, causing the two mechanical links to move and rotate the motor shaft. The two motor axes represent the x-axis, controlling the velocity, and the y-axis controlling the steering angle of the Car-like mobile robot. This control signal is inferred from the joystick terminal in the O_{xy} plane. From the joystick's coordinate system is selected in Figure b, the deflection angle at the two motors will be converted to the position of the joystick through the forward kinematics formulas calculated from:

$$\begin{aligned} x &= L\,sin(\varphi)\,sin(\gamma) \\ y &= L\,sin(\varphi)\,cos(\gamma) \end{aligned} \tag{1}$$

where L is the joystick handle's length, φ is the joystick's angle relative to the equilibrium position, γ is the angle between AC and BC; A, C, B are the projection of the joystick terminal onto the O_{xz}, O_{yz} plane and O_z axis respectively. φ and γ are calculated as the following equation:

$$\begin{aligned} \varphi &= \arctan\left(-\sqrt{\tan^2(\alpha_x)+\tan^2(\alpha_y)}\right) \\ \gamma &= \arctan\left(\frac{\tan(\alpha_x)}{\tan(\alpha_y)}\right) \end{aligned} \tag{2}$$

where α_x is the deflection angle in the x-axis motor, α_y is the deflection angle in the y-axis motor.

2.2 Car-like mobile robot model

According to the joystick's design, a Car-like mobile robot model suitable for the joystick is described in Figure . The Car-like mobile robot is driven by the velocity of the rear wheels v_1 and navigated through the front wheels steering angle ψ. The rear wheels are paralleled and fixed to the rigid body, and the front wheels navigate left and right but are always kept parallel to each other. Because of those constraints, which is nonholonomic constraints, the Car-like mobile robot can only move and steer at a limited angle and cannot change direction immediately or rotate in one place.

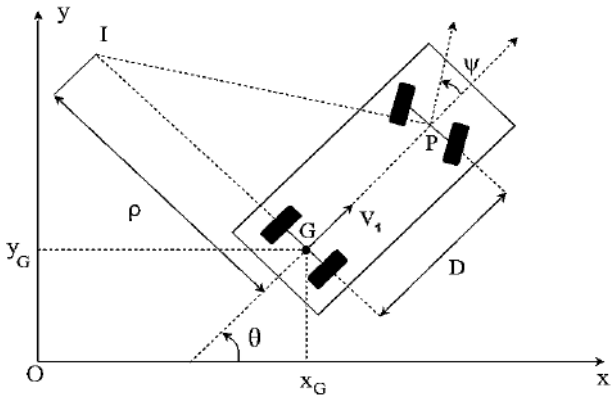


Figure 3. Car-like mobile robot model

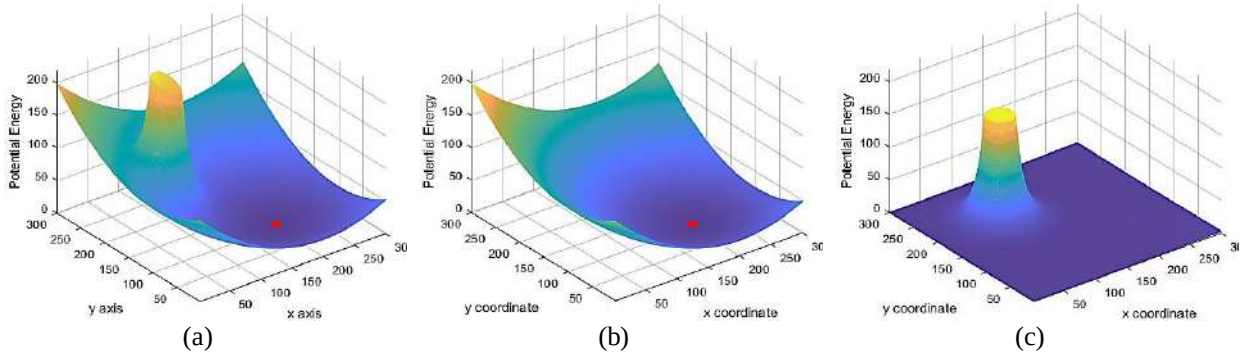
3 DESIGN HAPTIC CONTROLLER BASED ON POTENTIAL FIELD

From the overview diagram and the system model presented in section 2, the use of the Car-like mobile robot model that existing motion constraints can affect the precise control of the operator. When obstacles appear suddenly, the operator can have a slow reaction due to the lack of environmental information leading to collisions and damage to the robot. Since then, it has been found that applying the Haptic controller to feedback information about the robot's surroundings to the operator is essential. Potential Field is a potential method being considered by researchers around the world to optimize the feedback force in Haptic controllers. The following section discusses the theory of Potential Field and how to create a Haptic control.

3.1 Potential Field

The Potential Field algorithm has been widely applied to path planning [11] and solve obstacle avoidance problems [12]. The algorithm, as its name, creates a potential field formed with the repulsive and attractive fields. The repulsive field appears around the obstacle, the closer to the obstacle, the stronger the repulsive force. In contrast with the repulsive field, the attraction field is created from the target point, and the further from the target point, the stronger the attractive force. An intuitive description of the attractive and repulsive fields is presented in Figure . The Potential Field energy of Z, the considering point, is the sum of the repulsive field and the attractive field, described by the equation:

$$U(Z) = U_{att}(Z) + U_{rep}(Z) \tag{3}$$



(a) (b) (c)

Figure 4. The Potential Field energy (a) created from Attractive Field (b) and Repulsive Field (c)

In the proposed system, the operator controls the Car-like mobile robot through the joystick. Therefore, the attractive field does not need to be considered, and the operator only needs to be aware of obstacle information. The repulsive field around the obstacle will generate force to push the joystick return to the equilibrium position. The equation of the repulsive field is described as follows:

$$U_{rep}(Z) = \begin{cases} \frac{1}{2} k_{rep} \left(\frac{1}{\rho(Z)} - \frac{1}{\rho_0} \right)^2 & , \rho(Z) \le \rho_0 \\ 0 & , \rho(Z) > \rho_0 \end{cases} \quad (4)$$

where k_{rep} is the amplification coefficient of $U_{rep}(Z)$, $\rho(Z)$ is the distance between the point under consideration Z and the nearest obstacle,$\rho_0(Z)$ is the effective distance of the obstacle.

The repulsive force from the obstacle to Z is calculated using the gradient descent of the repulsive field:

$$F_{rep}(Z) = -\nabla U_{rep}(Z) = \begin{cases} k_{rep} \left(\frac{1}{\rho(Z)} - \frac{1}{\rho_0} \right) \frac{1}{\rho^2(Z)} \frac{\partial \rho(Z)}{\partial Z} & , \rho(Z) \le \rho_0 \\ 0 & , \rho(Z) > \rho_0 \end{cases} \quad (5)$$

Only when the robot moves inside the affected region of the obstacle, the feedback force will be generated.

3.2 Haptic controller based on Potential Field

For the joystick to return to its equilibrium position when the operator stops acting force, a virtual torsion spring is attached between the motor's shaft and the soles, as illustrated in Figure . The elastic force to pull the joystick back is described as:

$$F = ke \quad (6)$$

where k is the virtual spring's stiffness, e is the error between the motor's current and its equilibrium position.

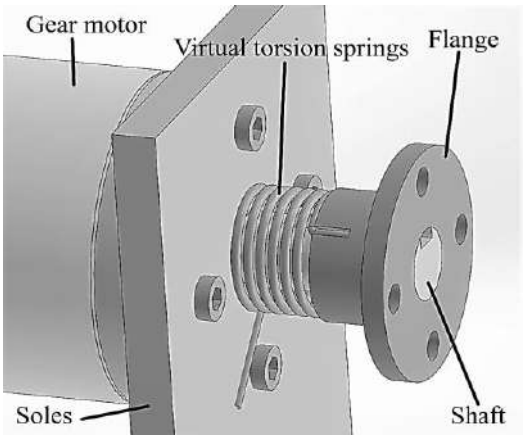


Figure 5. A virtual spring attached to motor flange and soles

The haptic controller's purpose is to create tactile feedback when the mobile robot moves close to an obstacle. Therefore, the virtual spring's stiffness should be adjustable to enhance stronger force feedback for the same e, the stiffness is consisted of:

$$k = k_o + \Delta k \quad (7)$$

where k_o is the natural stiffness of the virtual spring and Δk is the adjustable stiffness based on the distance to an obstacle.

Typically, the adjustable stiffness is based on a linear function of the distance between the obstacle and the mobile robot. However, with the haptic feedback force increasing steadily, the operator cannot precisely perceive how close the obstacle is to the mobile robot and could misjudge the collision rate's severity. Therefore, the nonlinear function from the repulsive force of the potential field algorithm is used to create the haptic force feedback from the nearest obstacle. In addition, virtual damping is also used to reduce the sudden feedback force from injuring the operator. The proposed haptic controller formula is:

$$F_h = k_P e + k_D \frac{de}{dt} \quad (8)$$

The feedback force from the controller is applied for each motor axis, where the total feedback force is divided into the joystick's two axes, x and y, thereby calculating the component force on each axis. From the equation (5), the adjustable stiffness of each axis's virtual spring can be calculated as:

$$\Delta k_{Px}(Z) = k_{rep}\left|cos\,\alpha\left(\frac{1}{\rho(Z)}-\frac{1}{\rho_0}\right)\frac{1}{\rho^2(Z)}\right|$$
$$\Delta k_{Py}(Z) = k_{rep}\left|sin\,\alpha\left(\frac{1}{\rho(Z)}-\frac{1}{\rho_0}\right)\frac{1}{\rho^2(Z)}\right| \quad (9)$$

where Δk_{Px}, Δk_{Py} are the adjusted stiffness of the virtual spring along the x-axis and y-axis respectively, α is the deviation angle between the obstacle and the vehicle direction.

When the mobile robot moves closer to the obstacle, depending on α, the virtual spring stiffness in the two axes will be different. The absolute value sign makes the stiffness of the virtual springs in the two axes always positive increases when the robot approaches an obstacle in any direction. The purpose is to create a tactile force against the operator's control force whenever an obstacle is approaching, increasing the operator's perception rather than taking control of the joystick. Therefore, both Δk_{Px} and Δk_{Py} have upper limits as shown in Figure .

From equations (7) and (9), the equation for the haptic controller's virtual spring stiffness is described as follows:

$$k_{Px}(Z) = \begin{cases} k_{Pox} + \Delta k_{Px}(Z) & ,\rho(Z) \le \rho_0 \\ k_{Pox} & ,\rho(Z) > \rho_0 \end{cases}$$
$$k_{Py}(Z) = \begin{cases} k_{Poy} + \Delta k_{Py}(Z) & ,\rho(Z) \le \rho_0 \\ k_{Poy} & ,\rho(Z) > \rho_0 \end{cases} \quad (10)$$

where k_{Pox}, k_{Poy} is the natural stiffness of virtual spring for the x-axis and y-axis, respectively.

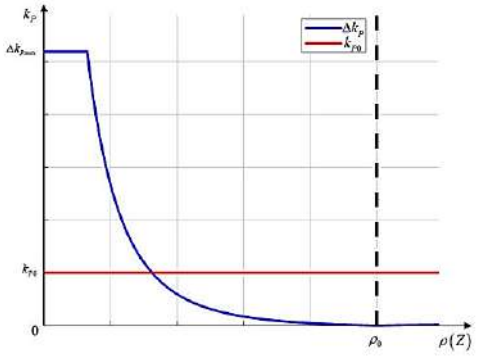


Figure 6. The virtual spring's component stiffnesses change with different distances to obstacle.

The equation (8) now can be rewritten for each joystick axis as follows:

$$F_{hx} = k_{Px}(Z)e_x + k_{Dx}\frac{de_x}{dt}$$
$$F_{hy} = k_{Py}(Z)e_y + k_{Dy}\frac{de_y}{dt} \quad (11)$$

In addition, to reduce the effect of unwanted control signals when the joystick is at its equilibrium position, a dead-zones are applied for both axes in Figure 7.

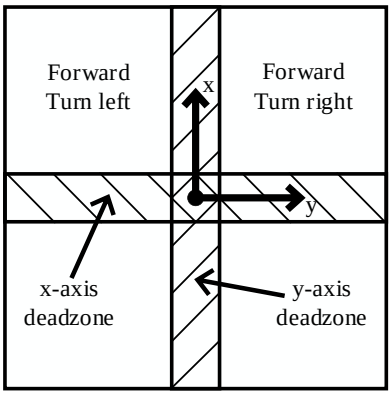


Figure 7. Dead-zones for joystick control signal

4 SIMULATIONS AND EXPERIMENTAL RESULTS

4.1 Simulation

In this section, two reference signals for both joystick axes are used to create an assuming trajectory through the repulsive field of an obstacle. A simple controller is used to help joystick follow those signals. The haptic controller adjusts the virtual spring's stiffness to change the reference control signals. Although the haptic controller aims to enhance the operator's perception, it is difficult to simulate. Therefore, in this simulation, the haptic controller will adjust the control signals, and no change for the reference signals despite of obstacle.

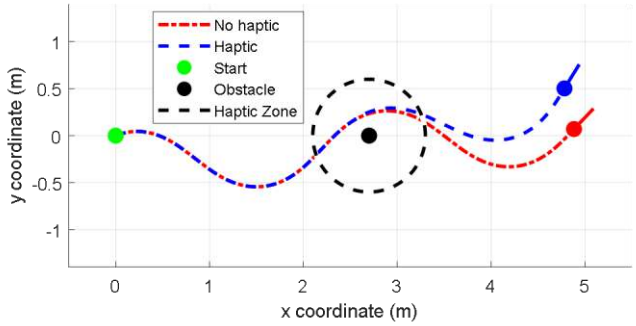
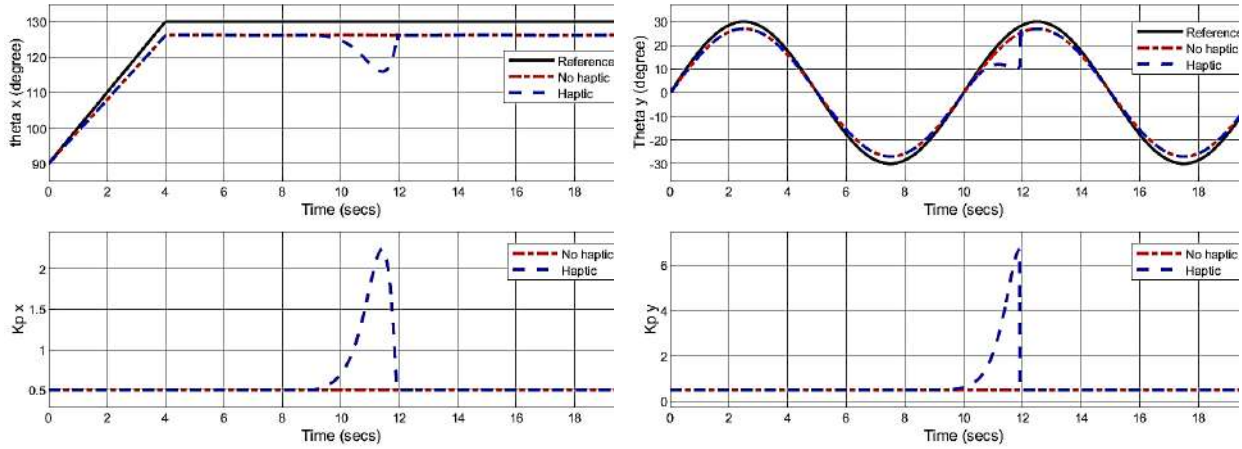


Figure 8. Comparison trajectory of Car-like mobile robot between haptic and non-haptic

Figure 8 shows the difference in the trajectory between haptic and non-haptic controllers. Initially, the mobile robot starts moving from the origin and has a waveform steering angle signal in Figure 5. When the robot turns right and moves close to the obstacle, 9 seconds from the start, a repulsive force lightly risen by the haptic controller affects both axes of the joystick. Due to the kinematic constraint mentioned in section 2.2, the Car-like mobile robot needs time to change its direction. Therefore, while the repulsive force keeps increasing, the Car-like mobile robot keeps reducing the velocity and steering angle to change direction until it bypasses the obstacle at 12 seconds. Then the haptic control stops creating repulsive force, and the joystick can keep following the reference signal of both axes. From that onward, the trajectory of both situations significantly differs because of the different control signals when the Car-like mobile robot moves close to the obstacle.



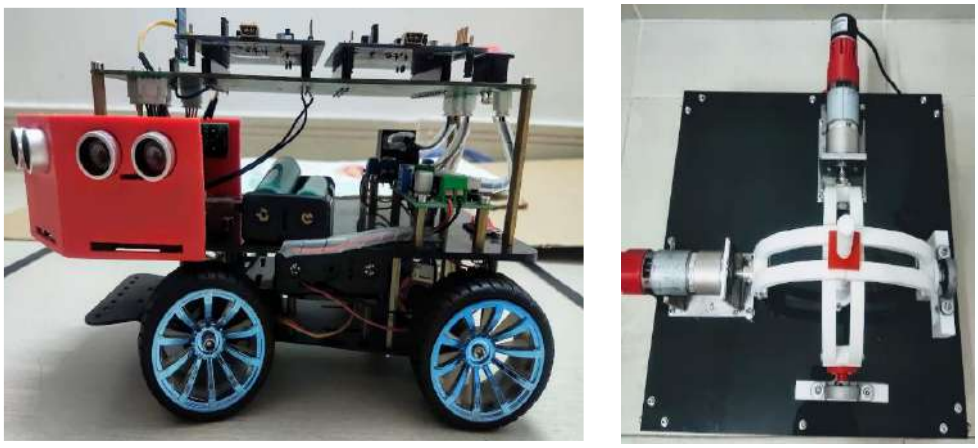
(a) (b)

Figure 5. Joystick's velocity signals on x-axis (a) and steering angle on y-axis (b)

The result shows that haptic control can help improve the control signal when mobile robot is close to obstacle, although there is no actual operator in the simulation. The haptic controller is also tested in experiments to evaluate the efficiency of improving the operator's obstacle perception.

4.2 Experimental

In Figure 6a, the experimental model Car-like mobile robot uses two STM32F407 boards; one for receiving signals from five sonars; one for controlling rear wheels velocity, front wheels steering angle and communicating with joystick model in Figure 6b. The joystick also uses two STM32F407 boards to control two axes and communicate with Car-like mobile robot. The joystick's length is $L = 18.3\ cm$ and the deflection angle range for each axis is $[-40° \sim 40°]$.



(a) (b)

Figure 6. Experiment model

The experimental map is demonstrated in Figure 7. Four operators will take part in the experiment with no experience in using a haptic controller before. The task is to control the Car-like mobile robot one circle around the wall at fast as possible and avoid collision. The operator stands with the joystick near the robot's start point. An obstacle higher than the wall is placed behind at the end of the first turn. The purpose is to limit the operator's vision and evaluate the performance when using haptic and non-haptic. Operators perform each test five times with no pre-arranged order.

Parameters for the experiment are: $k_{rep} = 200$, $\rho_o = 30\ cm$, $k_{Pox} = 0.02$, $k_{Poy} = 0.02$, $k_{Px} \leq 0.2$, $k_{Py} \leq 0.2$, $\alpha \in \{0°, \pm 22.5°, \pm 45°, \pm 77.5°, \pm 90°\}$.

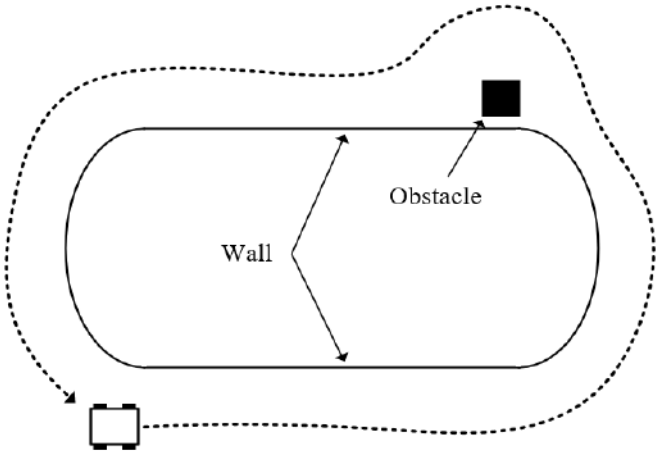


Figure 7. Experiment map

The execution time for all tests is shown in Figure 8. All operators have significant change between using only vision and vision combined with haptic controller. Figure 9 shows that the collision number also reduces when using haptic controller but still have some collisions from operator 2 and 4. The reasons are the kinematics constraint from the Car-like mobile robot, unable to change direction immediately, and the haptic controller only enhances the operator's perception of distance to the obstacle, not taking control. Therefore, when the Car-like mobile robot is going to collide and it takes too much time to change the robot's direction, the operator intends to accept the collision in exchange for a shorter moving time.

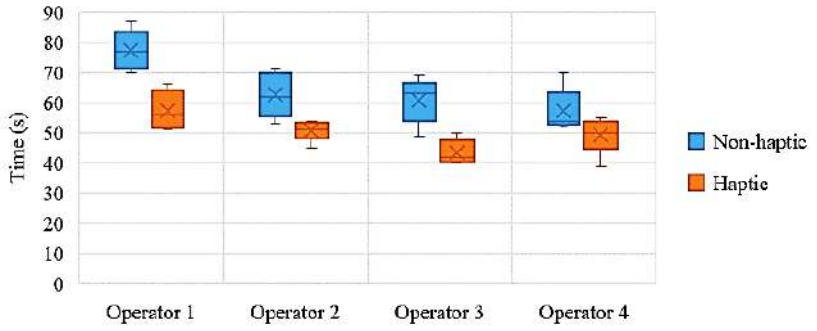


Figure 8. Execution time

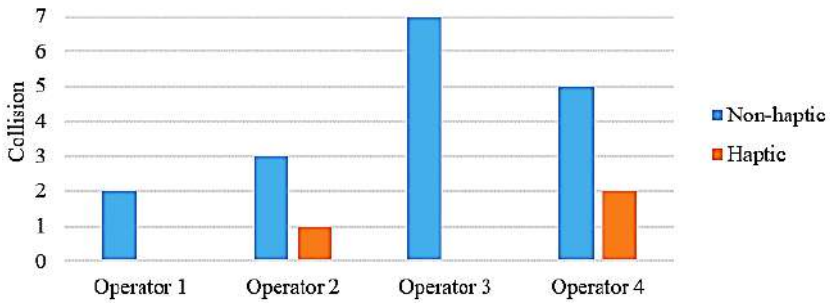


Figure 9. Number of collision

5 CONCLUSION AND FUTURE WORK

In this paper, a haptic controller is designed based on repulsive force from the Potential Field theory to enhance the operator's perception of obstacle distance. Models of a 2-DOF joystick and the Car-like mobile robot have been analyzed and implemented in simulations and experiments. The nonholonomic constraints of the Car-like mobile robot model have been considered due to the haptic controller's feedback force is only for a warning instead of taking control. The results show the feedback force of the haptic controller

helps the operator to recognize more precisely the obstacle distance and thereby give a more suitable control signal to avoid the collision. Future work will aim to feedback the combined repulsive force from multi obstacles and reduce the time delay of wireless communication in the experimental model.

ACKNOWLEDGEMENT

This research was implemented at Robotics and Intelligent Control Laboratory (RIC Lab), Faculty of Electrical and Electronics Engineering, Ho Chi Minh City University of Technology and Education, Vietnam.

REFERENCES

[1] Melonee Wise, M. Ferguson, D. King, E. Diehr, and D. Dymesich, "Fetch & Freight: Standard Platforms for Service Robot Applications," *Work. Auton. Mob. Serv. Robot. held 2016 Int. Jt. Conf. Artif. Intell.*, pp. 2–7, 2016.

[2] T. Linder et al., "Rescue robots at the Collapse of the municipal archive of Cologne City: A field report," *8th IEEE Int. Work. Safety, Secur. Rescue Robot*. SSRR-2010, 2010.

[3] M. J. Matarić, J. Eriksson, D. J. Feil-Seifer, and C. J. Winstein, "Socially assistive robotics for post-stroke rehabilitation," *Journal of NeuroEngineering and Rehabilitation*, vol. 4. 2007.

[4] S. Y. Harmon, "The Ground Surveillance Robot (GSR): An Autonomous Vehicle Designed to Transit Unknown Terrain," *IEEE J. Robot. Autom.*, vol. 3, no. 3, pp. 266–279, 1987.

[5] L. Yin, Y. Yin, and C. J. Lin, "A new potential field method for mobile robot path planning in the dynamic environments," Asian J. Control, vol. 11, no. 2, pp. 214–225, 2009.

[6] R. Nikoukhah, "An autonomous vision-based mobile robot," IEEE Trans. Automat. Contr., vol. 39, no. 3, pp. 493–502, 1994.

[7] M. Schwartz, "Haptic perception of the distance walked when blindfolded.," *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 25, no. 3, pp. 852–865, 1999.

[8] O. Martinez-Palafox, D. Lee, M. W. Spong, I. Lopez, and C. T. Abdallah, "Bilateral teleoperation of mobile robot over delayed communication network: Implementation," IEEE Int. Conf. Intell. Robot. Syst., no. May, pp. 4193–4198, 2006.

[9] J. Borenstein, H. R. Everett, L. Feng, and D. Wehe, "Mobile robot positioning: Sensors and techniques," *J. Robot. Syst.*, vol. 14, no. 4, pp. 231–249, 1997.

[10] F. Künhe, J. Gomes, and W. Fetter, "Mobile Robot Trajectory Tracking Using Model Predictive Control," *II IEEE latin-american Robot. …*, pp. 1–7, 2005.

[11] Y. K. Hwang and N. Ahuja, "A Potential Field Approach to Path Planning," IEEE Trans. Robot. Autom., vol. 8, no. 1, pp. 23–32, 1992.

[12] S. Mohammad, H. Rostami, A. K. Sangaiah, and J. Wang, "Obstacle avoidance of mobile robots using modified potential field algorithm," Eurasip J. Wirel. Commun. Netw., vol. 2019, no. 1, pp. 1–19, 2019.

ICATSD2F.108

ADAPTIVE SLIDING MODE CONTROL FOR THE 6 DOF MANIPULATOR WITH TIME-VARYING PAYLOAD

HA THANH BINH[1], TONG HAI NINH[1], TRAN MINH PHUC[1],
TRAN DUC THIEN[1], LE HOANG LAM[2]

[1]*Department of Automatic Control, Ho Chi Minh City University of Technology and Education, Ho Chi Minh, Vietnam*

[2]*Department of Electrical and Electronic Engineering, Ho Chi Minh University of Technology and Education, Ho Chi Minh, Vietnam*

htb.nckh@gmail.com, tongninh7@gmail.com, phuctm012@gmail.com, thientd@hcmute.edu.vn, lamlh@hcmute.edu.vn

Abstract. In this paper, an adaptive sliding mode control (ASMC) for the 6-DOF manipulator with time-varying payload is presented. In practice, the error of parameters, the mistakes in the connection between actuators and drivers, and the unknown control directions can make serious damage and chattering phenomenon for the manipulator. To deal with above problem, a proposed control is estimated based on the combination of sliding mode control (SMC) and adaptive law. SMC in the proposed control guarantees the tracking problem and fast responses for the manipulator. Adaptive laws are conducted to adjust the robust gain how to guarantee that the amplitude of the uncertainties is always bounded by them. To investigate the stability of the controller, the Lyapunov approach is used to analyze the whole system in theory. Finally, the proposed control is applied to 6 DOF manipulator in simulation on MATLAB Simulink to prove the effectiveness.

Keywords. ASMC, adaptive control, Lyapunov stability, Sliding mode control, 6-DOF manipulator.

1 INTRODUCTION

For the past few decades, manipulators have supported humans in complex works, iterate quickly and accurately in many industrial automation processes such as assembly, weld, clevis, and load lifting. Motion control for a robotic arm is always a difficult task because of its high non-linear, time-varying parameters, model error, uncertain parameters, and the influence of unknown noise in the model. In order to obtain reasonable control performance, several robust advanced control methods have been developed, such as adaptive control [1, 2], back-stepping control [3-5], sliding mode control [6], and neural control [7]. Although all these controllers have been successfully used in several situations, they do have significant drawbacks.

In Getachew (2018) [8], the authors aimed at sliding mode control for a two degree of freedom (2-DOF) robotic manipulator. Sliding mode control is a characteristic control of robust control for non-linear systems. The outstanding feature of the SMC controller is its robustness to the model containing uncertain parameters and immeasurable noise. The principle of SMC is based on discontinuous control to force the control system's error state variable toward zero. Based on simulation result, when the system is subject to a disturbance of 25% of the input amplitude, the performance response is excellent. The issue with this was that the sliding surface and the derivative of the sliding surface of the system generated a chattering, which caused a high-frequency vibration in the systems when the disturbance of the system climbed to 40% of the input amplitude. This phenomenon can excite high-frequency dynamics and cause instability in the controlled system. In order to handle this issue, some methods, such as boundary layer methods [9], Fuzzy Logic systems [10] have been integrated into SMC. In [9], the presentation of SMC with a boundary layer approach to reduce or eliminate chattering. The fundamental concept behind the boundary layer technique is to replace the saturation function in a short area around the switching surface with a discontinuous function. However, this substitution increased the rising time and the performance of errors. Fuzzy logic

system [10] show the advantages of approximating the undefined nonlinear functions in the system and change the robust component in SMC to reduce the chattering in system, the limitation of these methods is the need for qualified experts with experience in algorithm testing. In other approaches, adaptive laws [11] have been used to adjust the switching gain online. The advantage of this method is simply structure and less computation. However, in the simulation, to highlight the control algorithm's accuracy, the models used only revolve around the 2 to 3 degrees of freedom robotic arms with simple mathematical models, some complex mathematical models are not mentioned such as 6-DOF manipulator.

This work proposes a novel adaptive sliding mode control for the 6 DOF manipulator with time-varying payload and the impact of the uncertainties parameters based on the aforementioned research. The proposed control is designed by using the sliding mode control and adaptive laws. In the suggested control, sliding mode control ensures that the tracking issue and quick responses for the manipulators. The adaptive laws are used to adjust the switching gains in the sliding mode control. Thereby, when combines the sliding controller and the adaptive law will help the system to track the set signal more stably and accurately, and the influencing factors external effects when the load changes over time. In addition, the stability and sustainability of the entire system with the proposed controller are proven according to Lyapunov criteria. Finally, some simulations of the controller were performed on a six-degree-of-freedom robotic arm and the simulation results were compared with conventional SMC to show that the suggested controller is effective.

The rest of this study is as follows: In section 2, the dynamic of the 6 DOF manipulator with a time-varying load is shown. In section 3, the design procedures of an adaptive sliding mode control for the manipulator are presented and analyzed, besides, to demonstrate the stability and robustness of the entire system, a Lyapunov technique is applied. Some simulations will be performed in section 4. Finally, several conclusions are discussed in section 5.

2 PROBLEM FORMULAION

Consider a six-degree-of-freedom manipulator with a time-varying load, unknown friction, and outside disturbance. Its dynamics are depicted in the join space by [12].

$$\boldsymbol{M}(\boldsymbol{\theta})\ddot{\boldsymbol{\theta}} + \boldsymbol{C}\left(\boldsymbol{\theta},\dot{\boldsymbol{\theta}}\right)\dot{\boldsymbol{\theta}} + \boldsymbol{G}(\boldsymbol{\theta}) + \boldsymbol{J}^{T}(\boldsymbol{\theta})\boldsymbol{f} + \boldsymbol{\tau}_{fric} = \boldsymbol{\tau} \tag{1}$$

where $\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}, \ddot{\boldsymbol{\theta}} \in \Re^{6\times1}$ display, in turn, the angular position, velocity, and acceleration vectors in the manipulator's joint space; $\boldsymbol{M}(\boldsymbol{\theta}) \in \Re^{6\times6}$ presents the inertia matrix; $\boldsymbol{C}\left(\boldsymbol{\theta},\dot{\boldsymbol{\theta}}\right)\dot{\boldsymbol{\theta}} \in \Re^{6\times6}$ expresses the Coriolis and centrifugal term matrix; $\boldsymbol{G}(\boldsymbol{\theta}) \in \Re^{6\times1}$ calculates the gravity vector; $\boldsymbol{J}(\boldsymbol{\theta}) \in \Re^{6\times6}$ is a nonsingular Jacobian matrix; $\boldsymbol{f} \in \Re^{6\times1}$ indicates the external disturbance vector; $\boldsymbol{\tau}_{fric} \in \Re^{6\times1}$ is the unknown friction vector; $\boldsymbol{\tau} \in \Re^{6\times1}$ is the torque vector affects on joints.

The friction model, $\boldsymbol{\tau}_{fric}$ is exhibited as

$$\boldsymbol{\tau}_{fric} = \boldsymbol{F}_{v}\dot{\boldsymbol{\theta}} + \boldsymbol{F}_{c}sign\left(\dot{\boldsymbol{\theta}}\right) \tag{2}$$

where $\boldsymbol{F}_{v} \in \Re^{6\times6}$ and $\boldsymbol{F}_{c} \in \Re^{6\times6}$ are positive diagonal matrices

Property 1 [13]. $\dot{\boldsymbol{M}}(\boldsymbol{\theta}) - \boldsymbol{2C}\left(\boldsymbol{\theta},\dot{\boldsymbol{\theta}}\right)$ is a skew-symmetric matrix, that provided as $\boldsymbol{z}^{T}\left[\dot{\boldsymbol{M}}(\boldsymbol{\theta}) - \boldsymbol{2C}\left(\boldsymbol{\theta},\dot{\boldsymbol{\theta}}\right)\right]\boldsymbol{z} = 0$, where $\boldsymbol{z} \in \Re^{6\times1}$is a any vector.

Let $\boldsymbol{x}_1 = \boldsymbol{\theta} \in \Re^{6\times1}$ and $\boldsymbol{x}_2 = \dot{\boldsymbol{\theta}} \in \Re^{6\times1}$; then the robotic dynamics (1) can be presented by the equation of state as:

$$\begin{aligned} \dot{\boldsymbol{x}}_1 &= \boldsymbol{x}_2 \\ \dot{\boldsymbol{x}}_2 &= \boldsymbol{M}^{-1}(\boldsymbol{x}_1)(\boldsymbol{u} - \boldsymbol{C}(\boldsymbol{x}_1,\boldsymbol{x}_2)\boldsymbol{x}_2 - \boldsymbol{G}(\boldsymbol{x}_1) - \boldsymbol{\Delta}) \end{aligned} \tag{3}$$

where $\boldsymbol{u} = \boldsymbol{\tau}$depicts the control input vector; $\boldsymbol{\Delta} = \boldsymbol{J}^{T}(\boldsymbol{x}_1)\boldsymbol{f} + \boldsymbol{\tau}_{fric} \in \Re^{6\times1}$ produces a lumped disturbance that includes the external disturbance and unknown friction.

Assumption 1: *The external disturbance and unknown friction,$\boldsymbol{\Delta}$, are bounded function. This assumption is presented as* $\|\boldsymbol{\Delta}\|_{\infty} \leq \delta$.

Assumption 2: *The control goal of this study is to design an adaptive controller for the manipulator how the output responses,$\boldsymbol{\theta}$, in each joint track to the desired response,* $\boldsymbol{\theta}_{d}$.

3 CONTROL DESIGN

3.1 Conventional sliding mode control design

Tracking error is defined as follows:

$$\boldsymbol{e} = \boldsymbol{x}_1 - \boldsymbol{x}_{1\boldsymbol{d}} \in \Re^{6\times 1} \tag{3}$$

The sliding mode manifold is derived as follows:

$$\boldsymbol{s} = \dot{\boldsymbol{e}} + \boldsymbol{\lambda e} \in \Re^{6\times 1} \tag{4}$$

where $\boldsymbol{\lambda} = diag([\lambda_1 \quad \lambda_2 \quad \lambda_3 \quad \lambda_4 \quad \lambda_5 \quad \lambda_6]) \in \Re^{6\times 6}$ are positive definite matrices.
The time derivative of the sliding surface is computed as:

$$\dot{\boldsymbol{s}} = \ddot{\boldsymbol{e}} + \boldsymbol{\lambda}\dot{\boldsymbol{e}} \tag{5}$$

Replace to (5), it is rewritten as follows:

$$\dot{\boldsymbol{s}} = \boldsymbol{\lambda}\dot{\boldsymbol{e}} - \ddot{\boldsymbol{x}}_{1\boldsymbol{d}} + \boldsymbol{M^{-1}}(\boldsymbol{x}_1)(\boldsymbol{u} - \boldsymbol{C}(\boldsymbol{x}_1, \boldsymbol{x}_2)\boldsymbol{x}_2 - \boldsymbol{G}(\boldsymbol{x}_1) - \boldsymbol{\Delta}) \tag{6}$$

Based on assumption 2 and the sliding mode design procedure, a SMC can be designed consist of two terms which are equilibrium and robust terms:

$$\boldsymbol{u} = \boldsymbol{u}_{eq} + \boldsymbol{u}_r \tag{7}$$

Where $\boldsymbol{u}_{eq}$ is chosen as follow as:

$$\boldsymbol{u}_{eq} = \boldsymbol{M}(\boldsymbol{x}_1)(-\boldsymbol{\lambda}\dot{\boldsymbol{e}} + \ddot{\boldsymbol{x}}_{1\boldsymbol{d}}) + \boldsymbol{C}(\boldsymbol{x}_1, \boldsymbol{x}_2)(\dot{\boldsymbol{x}}_{1d} - \boldsymbol{\lambda e}) + \boldsymbol{G}(\boldsymbol{x}_1) - \boldsymbol{Ks} \tag{8}$$

is an equivalent control with $\boldsymbol{K} = diag([k_1, k_2, \ldots, k_6]) \in \Re^{6\times 6}$, and robust control:

$$\boldsymbol{u}_r = -\boldsymbol{\eta} signn(\boldsymbol{s}) \tag{9}$$

is employed to ensure the stability of systems where lumped uncertainties are present. $\boldsymbol{\eta} = diag([\eta_1, \eta_2, \ldots, \eta_6]) \in \Re^{6\times 6}$expresses a diagonal positive matrix which is satisfied by the condition $\eta_i \geq |\delta_i|.signn(\boldsymbol{s}) = [sign(s_1), sign(s_2)\ldots, sign(s_6)]^T \in \Re^{6\times 1}$
From equation (2.18) the control law of the system can be define as:

$$\boldsymbol{u} = \boldsymbol{M}(\boldsymbol{x}_1)(-\boldsymbol{\lambda}\dot{\boldsymbol{e}} + \ddot{\boldsymbol{x}}_{1\boldsymbol{d}}) + \boldsymbol{C}(\boldsymbol{x}_1, \boldsymbol{x}_2)(\dot{\boldsymbol{x}}_{1d} - \boldsymbol{\lambda e}) + \boldsymbol{G}(\boldsymbol{x}_1) - \boldsymbol{Ks} - \boldsymbol{\eta} signn(\boldsymbol{s}) \tag{10}$$

Remark 1: In the conventional SMC, the boundaries of the uncertainties need to be predetermined to guarantee the effectiveness of the control design. This is the limitation of the conventional SMC. It can cause chattering effect in the control signal.

3.2 Adaptive Sliding mode control

In practice, the external disturbance usually changes with respect to time, and the robust gains can cause the chattering effect when the robust gains are significantly larger than the external disturbance. To overcome this issue, an adaptive law is utilized to adjust the switching gain online. The robust term is modified as follows:

$$\boldsymbol{u}_r = \hat{\boldsymbol{\eta}} signn(\boldsymbol{s}) \tag{11}$$

The adaptive law is selective as follows:

$$\dot{\hat{\eta}}_i = \begin{cases} k_{1i} s_i |s_i| > \varepsilon_i \\ -k_{2i} s_i |s_i| \leq \varepsilon_i \end{cases} (i = 1,2,3,\ldots,6) \tag{12}$$

where k_{1i}, k_{2i} are positive constant and $\varepsilon_i \in \Re^{6\times 1}$is the predefined error boundaries of each joint. It is not necessary for these adaptive robust gain laws to understand the upper boundary of the external disturbance. The adaptive robust gains will rapidly grow and drive the sliding surface s_i to the desired areas that are smaller than ε_i. Otherwise, when the sliding variable stays in these regions, the gains will gradually reduce. From the property of the adaptive robust gain laws, under the influence of the external time-varying disturbance, the chattering effect is greatly diminished.

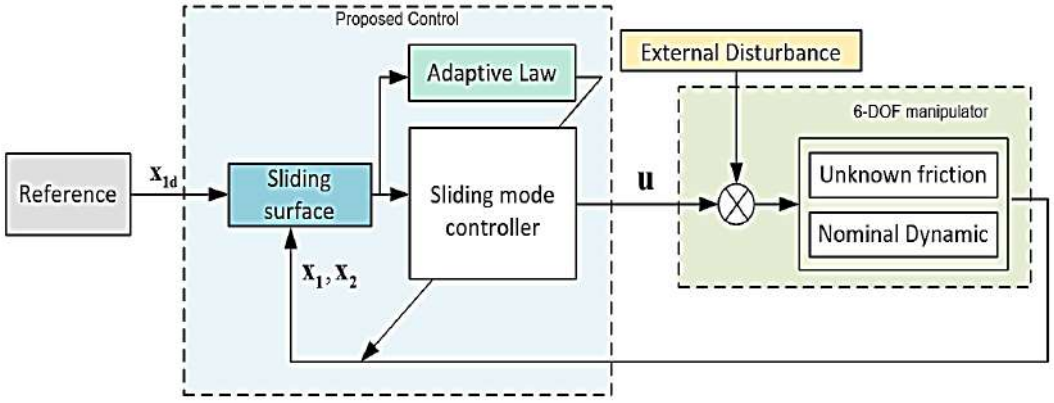


Figure 1. Schematic of the proposed control

The new control law is modified as:

$$\boldsymbol{u} = \boldsymbol{M}(\boldsymbol{x}_1)(-\boldsymbol{\lambda}\dot{\boldsymbol{e}} + \ddot{\boldsymbol{x}}_{1\boldsymbol{d}}) + \boldsymbol{C}(\boldsymbol{x}_1, \boldsymbol{x}_2)(\dot{\boldsymbol{x}}_{1d} - \boldsymbol{\lambda}\boldsymbol{e}) + \boldsymbol{G}(\boldsymbol{x}_1) - \boldsymbol{K}\boldsymbol{s} - \hat{\boldsymbol{\eta}}\,signn(\boldsymbol{s}) \tag{13}$$

Then the time derivative of the sliding surface is computed as:

$$\begin{aligned} \dot{\mathbf{s}} &= \boldsymbol{\lambda}\dot{\mathbf{e}} - \ddot{\mathbf{x}}_{1\mathbf{d}} + \mathbf{M}^{-1}(\mathbf{x}_1)\left(\mathbf{u} - \mathbf{C}(\mathbf{x}_1, \mathbf{x}_2)\mathbf{x}_2 - \mathbf{G}(\mathbf{x}_1) - \boldsymbol{\Delta}\right) \\ &= -\mathbf{K}\mathbf{s} - \hat{\boldsymbol{\eta}}\,signn(\mathbf{s}) - \boldsymbol{\Delta} - \mathbf{C}(\mathbf{x}_1, \mathbf{x}_2)\mathbf{s} \end{aligned} \tag{14}$$

Theorem 1: *When the outside disturbance is took into account in the robot manipulator with the sliding mode surface shown in* (5)*, the controller defined in* (10)*, and the adaptive law then the sliding variables, s_i , are bounded. Consequently, the closed-loop system is proven to be uniformly ultimately bounded (UUB). Additionally, the chattering effect is reduced significantly.*

Proof of theorem 1: To demonstrate the effectiveness of the ASMC with the selected adaptive law, the Lyapunov function of the whole system is chosen as follows:

$$V = \frac{1}{2}\boldsymbol{s}^{\boldsymbol{T}}\boldsymbol{M}(\boldsymbol{x}_1)\boldsymbol{s} + \frac{1}{2}\tilde{\boldsymbol{\eta}}^{\boldsymbol{T}}\boldsymbol{P}^{-1}\tilde{\boldsymbol{\eta}} \tag{15}$$

where $\boldsymbol{P} \in \Re^{6\times 6}$ is a positive diagonal matrix, and $\tilde{\boldsymbol{\eta}} = \boldsymbol{\eta}^* - \hat{\boldsymbol{\eta}} \in \Re^{6\times 1}$ Where $\boldsymbol{\eta}^* \in \Re^{6\times 1}$ is the upper boundary of the uncertainties and external disturbances. The Lyapunov function's time derivative is:

$$\begin{aligned} \dot{V} &= \mathbf{s}^{\mathbf{T}}\mathbf{M}(\mathbf{x}_1)\dot{\mathbf{s}} + \frac{1}{2}\mathbf{s}^{\mathbf{T}}\dot{\mathbf{M}}(\mathbf{x}_1)\mathbf{s} + \tilde{\boldsymbol{\eta}}^{\mathbf{T}}\mathbf{P}^{-1}\dot{\tilde{\boldsymbol{\eta}}} \\ &= \mathbf{s}^{\mathbf{T}}\left(-\mathbf{K}\mathbf{s} - \hat{\boldsymbol{\eta}}\,signn(\mathbf{s}) - \boldsymbol{\Delta}\right) + \left(\frac{1}{2}\mathbf{s}^{\mathbf{T}}\dot{\mathbf{M}}(\mathbf{x}_1)\mathbf{s} - \mathbf{C}(\mathbf{x}_1, \mathbf{x}_2)\mathbf{s}\right) + \tilde{\boldsymbol{\eta}}^{\mathbf{T}}\mathbf{P}^{-1}\dot{\tilde{\boldsymbol{\eta}}} \end{aligned} \tag{16}$$

By applies the Skew-matrix in property 1, then the derivative of Lyapunov function is calculated as:

$$\begin{aligned} \dot{\mathbf{V}} &= -\mathbf{s}^{\mathbf{T}}\mathbf{K}\mathbf{s} - \mathbf{s}^{\mathbf{T}}\left(\boldsymbol{\Delta} + \hat{\boldsymbol{\eta}}\,signn(\mathbf{s})\right) + \tilde{\boldsymbol{\eta}}^{\mathbf{T}}\mathbf{P}^{-1}\dot{\tilde{\boldsymbol{\eta}}} \\ &\le -\mathbf{s}^{\mathbf{T}}\mathbf{K}\mathbf{s} - \mathbf{s}^{\mathbf{T}}signn(\mathbf{s})\left(-\boldsymbol{\eta}^* + \hat{\boldsymbol{\eta}}\right) + \tilde{\boldsymbol{\eta}}^{\mathbf{T}}\mathbf{P}^{-1}\left(\dot{\boldsymbol{\eta}}^* - \dot{\hat{\boldsymbol{\eta}}}\right) \\ &\le -\sum_{i=1}^{6} s_i^2 K_i + \sum_{i=1}^{6}\tilde{\eta}_i\left(|s_i| - \kappa_i^{-1}\dot{\hat{\eta}}_i\right) \end{aligned} \tag{17}$$

Now, taking into account two cases of the adaptive law: $|s_i| \ge \varepsilon_i$ and $|s_i| \le \varepsilon_i$, when the sliding variable remains outside of the region in the first instance, $|s_i| \ge \varepsilon_i$, from the adaptive law (12) then the inequality (17) becomes:

$$\dot{V} \le -\sum_{i=1}^{6} K_i s_i^2 \le -\sum_{i=1}^{6} K_i \varepsilon_i^2 \tag{18}$$

The time derivative of the Lyapunov function is a negative function, so that the sliding variables converge to the sliding surface, and they stay in the region, $|s_i| < \varepsilon_i$. When the sliding variable enter this region, then the time derivative of the Lyapunov function becomes:

$$\begin{aligned} \dot{V} &\le -\sum_{i=1}^{6} K_i s_i^2 + \sum_{i=1}^{6} \tilde{\eta}_i \left(|s_i| - \kappa_i^{-1} \dot{\hat{\eta}}_i \right) \\ &\le \sum_{i=1}^{6} |s_i| \left(2\eta_i^* - k_i |s_i| \right) \end{aligned} \tag{19}$$

In this case, $\dot{V}$ is not sure to be non-positive because of uncertain equation, $2\eta_i^* - k_i|s_i|$, so that the sliding variable can still approach the sliding manifold or leaves this region belong the above equation. To make sure the time derivative of the Lyapunov function is always negative, the system is designed with the condition:

$$\varepsilon_i > |s_i| > \frac{2\eta_i^*}{k_i} \tag{20}$$

Based on the above analysis, the controlled system is stable. The radius of ultimate steady state error is depended on the upper boundary of the uncertainties, $\boldsymbol{\eta}^*$, and the control gain, $\boldsymbol{K}$. Then, the accuracy of the controlled system can be enhanced by increasing the gain $\boldsymbol{K}$.

4 NUMERICAL SIMULATIONS

4.1 Simulation description

To ensure the accuracy and appropriateness of the suggested control, simulations are carried out based on 6-DOF manipulator model, as shown in figure 2 and the parameters in detail are shown in Table 1.

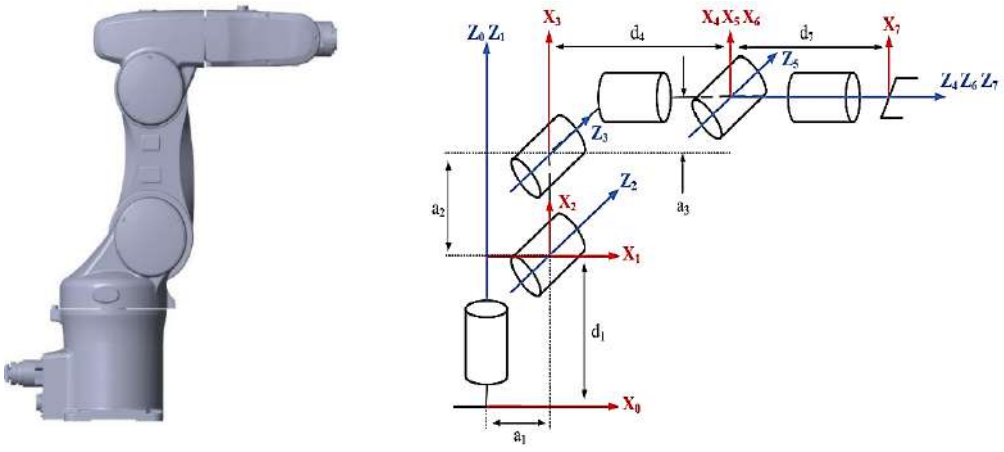


Figure 2. The model of the 6-DOF manipulator

Table 1: The parameters of 6-DOF manipulator

Variables	Description	Variables	Description
$a_1 = 0.0075(m)$	Distance between z_0-axis and z_2- axis	$m_2 = 4.6(Kg)$	Weight of link 2
$a_2 = 0.27(m)$	Distance between z_2-axis and z_3- axis	$m_3 = 4.3(Kg)$	Weight of link 3
$a_3 = 0.09(m)$	Distance between z_3-axis and z_5- axis	$m_4 = 1.9151(Kg)$	Weight of link 4
$d_1 = 0.335(m)$	Distance between x_0-axis and x_1- axis	$m_5 = 0.4(Kg)$	Weight of link 5

$d_4 = 0.295(m)$	Distance between x_3-axis and $x_4x_5x_6$- axis	$m_6 = 0.0714(Kg)$	Weight of link 6
$d_7 = 0.08(m)$	Distance between x_4-axis and x_7- axis	$\boldsymbol{F}_c = diag([5,1,1,0.001, 0.0001,0.0001])(\boldsymbol{N}/m)$	Coefficients of Coulomb friction
$m_1 = 5.65(Kg)$	Weight of link 1	$\boldsymbol{F}_v = diag([5\ 1\ 1\ 0.001\ 0.0001\ 0.0001])(\boldsymbol{N}/m)$	Coefficients of Viscous friction

The robot manipulator's initial position was set to $\boldsymbol{x_1}(0) = [0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0]^T(deg())$, and $\boldsymbol{x_2}(0) = [0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0]^T(deg.s^{-1})$. The desired trajectory for each joint is set the same as:

$$\boldsymbol{x_{1d}} = 60\,sin(2\pi ft) \tag{21}$$

where the amplitude of each joint is equal by 60 with f is the frequency and t is the operating time.
The manipulator operation is separated into two stages based on the operating conditions described in the simulation description with sampling time of 0.01 seconds, simulation time of 20 seconds, and $f = 0.1Hz$.
Remark 2: To confirm the productivity of the proposed control, compare the conventional sliding mode control with the suggested control, ASMC.
The SMC with control laws:

$$\boldsymbol{u} = \boldsymbol{M}(\boldsymbol{x}_1)(-\boldsymbol{\lambda}\dot{\boldsymbol{e}} + \ddot{\boldsymbol{x}}_{1d}) + \boldsymbol{C}(\boldsymbol{x}_1, \boldsymbol{x}_2)(\dot{\boldsymbol{x}}_{1d} - \boldsymbol{\lambda e}) + \boldsymbol{G}(\boldsymbol{x}_1) - \boldsymbol{Ks} - \boldsymbol{\eta} signn(\boldsymbol{s}) \tag{22}$$

The proposed control presents as:

$$\boldsymbol{u} = \boldsymbol{M}(\boldsymbol{x}_1)(-\boldsymbol{\lambda}\dot{\boldsymbol{e}} + \ddot{\boldsymbol{x}}_{1d}) + \boldsymbol{C}(\boldsymbol{x}_1, \boldsymbol{x}_2)(\dot{\boldsymbol{x}}_{1d} - \boldsymbol{\lambda e}) + \boldsymbol{G}(\boldsymbol{x}_1) - \boldsymbol{Ks} - \hat{\boldsymbol{\eta}} signn(\boldsymbol{s}) \tag{23}$$

With the adaptive law:

$$\dot{\hat{\eta}}_i = \begin{cases} k_{1i}s_i & |s_i| > \varepsilon_i \\ -k_{2i}s_i & |s_i| \le \varepsilon_i \end{cases} \tag{24}$$

In the first stage, with two controllers, this study analyzes the impact of friction torques and specified performances on output responses. It happens in the first ten seconds. Next, from 10th to 20th second, when an external force, $\boldsymbol{F}_{ext} = [0 \quad 0 \quad 8 \quad 0 \quad 0 \quad 0]^T(N)$, is applied to the end-effector (EE) in the z-axis direction.
Remark 3: The external force applied to the EE in the z-axis direction is simulated for the "time-varying payload" in the system.
Remark 4: The parameter of the SMC and ASMC are chosen to make the derivative Lyapunov function (15) is a negative function. With the proof of theorem 2 and trial-error method, the parameter can be selected as the table below.

Table 2: The parameters of the two controls discussed before.

Controllers	Control Parameters
SMC	$\boldsymbol{\lambda} = diag([15,20,15,50,120,20]), \boldsymbol{K} = diag([10,10,10,0.4,0.1,0.06]),$ $\boldsymbol{\eta} = diag([7,4,3,0.002,0.3,0.005])$
ASMC	$\boldsymbol{\lambda} = diag([15,20,15,50,120,20]), \boldsymbol{K} = diag([10,10,10,0.4,0.1,0.06]),$ $\boldsymbol{\kappa}_1 = diag([50,80,10,0.001,1,0.01]), \boldsymbol{\kappa}_2 = diag([100,100,50,0.05,1,1]),$ $\boldsymbol{\varepsilon} = [0.05,0.05,0.05,0.15,2,0.05]^T$

4.2 Simulation results

Figure 3 presents the angular position response at each joint of 6-DOF manipulator with black lines of reference, blue lines of SMC and red lines of ASMC. The figure shown that the SMC and ASMC guaranteed the tracking problem for the 6-DOF manipulator in both case of simulation.

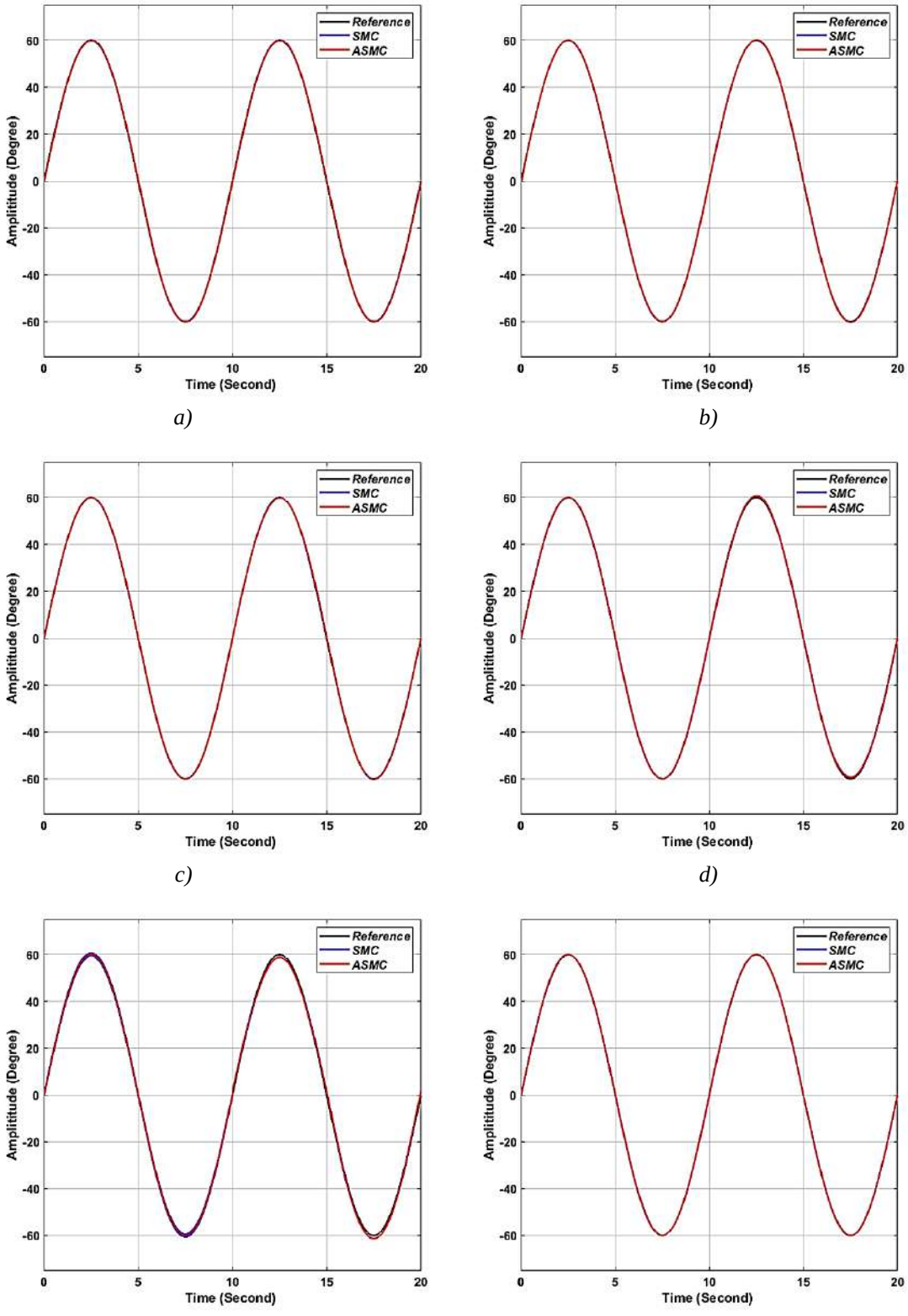


Figure 3. Output response of the manipulator with SMC and ASMC in a) joint 1, b) joint 2, c) joint 3, d) joint 4, e) joint 5, and f) joint 6.

To declare the performance of each controller, Figure 4 present the error efforts at each joint. In Figure 4, error performances of SMC and ASMC are displayed, in turn, by black lines and red lines.

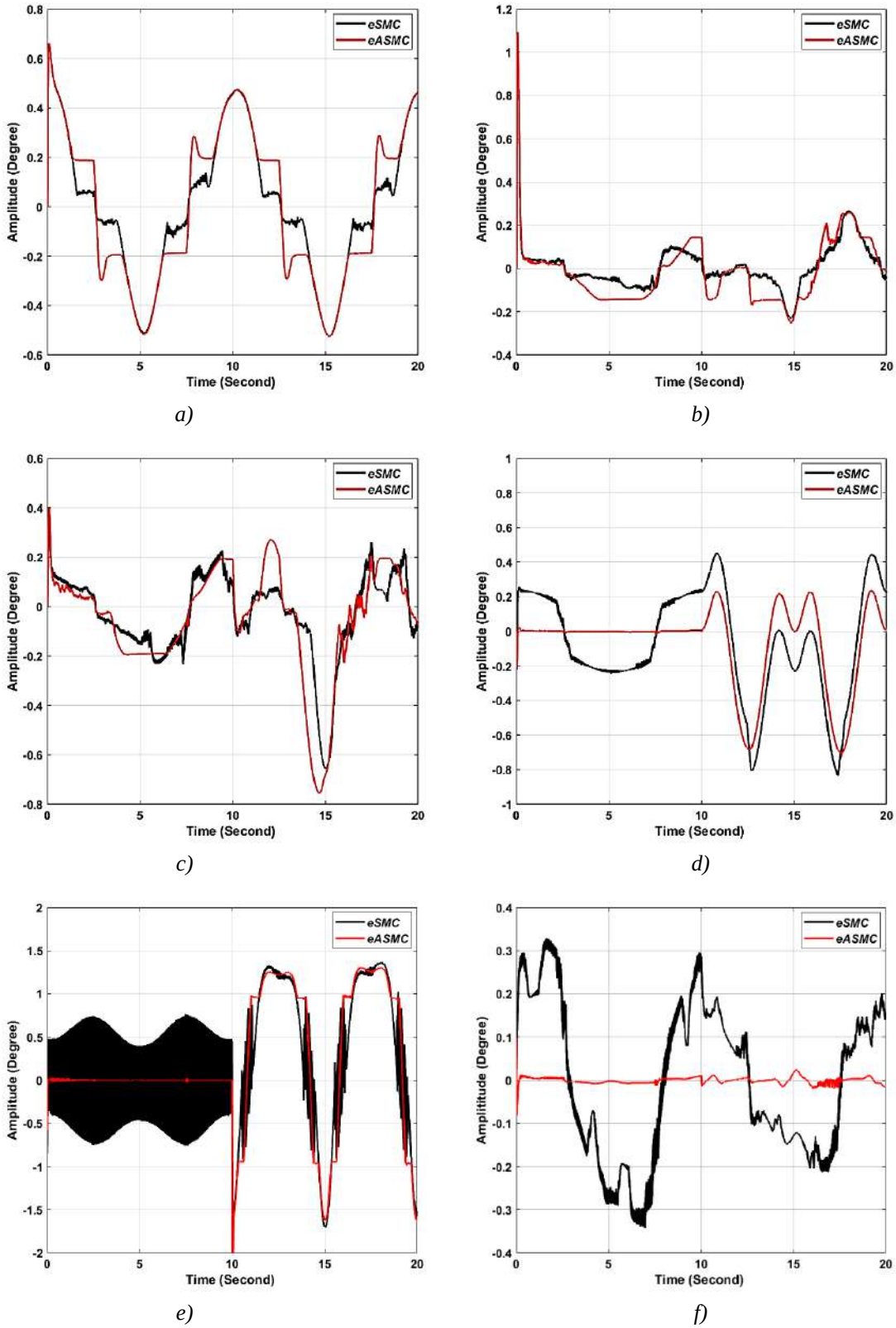


Figure 4. Errors performance of each joint with two different controls SMC and ASMC in a) joint 1, b) joint 2, c) joint 3, d) joint 4, e) joint 5, and f) joint 6.

Figure 5 shows the control signals at each joint of six-DOF manipulator in 2 lumped controllers. The control signal of SMC is black line and red line for ASMC.

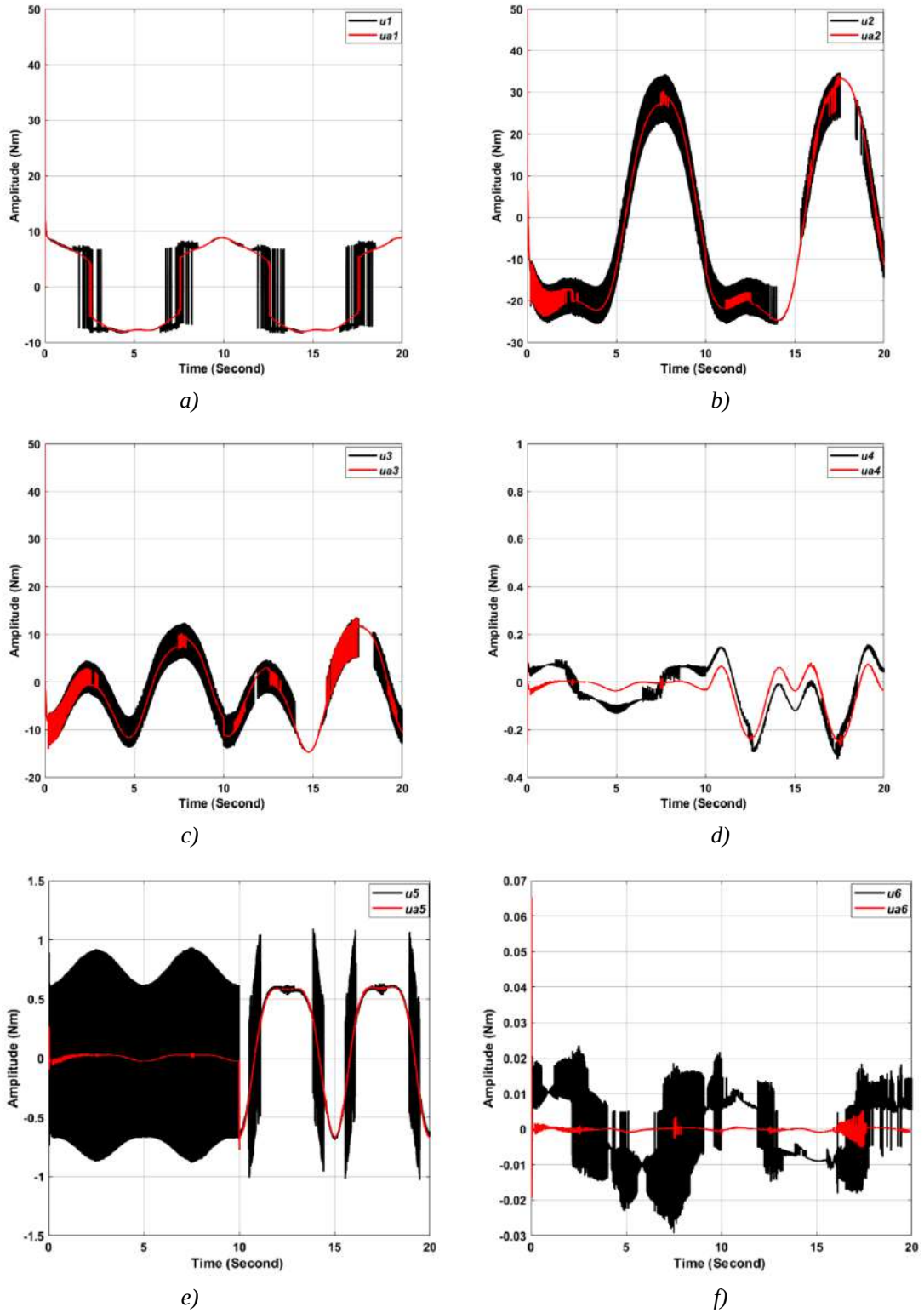


Figure 5. Control signals responses of the manipulator with SMC and ASMC in a) joint 1, b) joint 2, c) joint 3, d) joint 4, e) joint 5, and f) joint 6.

Compare with the conventional SMC, to obtain the accuracy as present in figure 4, the control signal of SMC arises high-frequency vibration, named chattering effect, when using the ASMC, which helps the system not only reduced the chattering effect but also improved the accuracy of tracking problem.

5 CONCLUSIONS

In this study, an adaptive sliding mode control for 6-DOF manipulator with time-varying payload under the present of frictions and external disturbances was aimed. The tracking issue and quick response for the manipulator are guaranteed by the SMC control in the recommended control. Additionally, a strong gain adaptive law in the SMC is used to deal with unknown friction and monitor the system's chattering effect. As a result, it helps the proposed control not only ameliorated the performance of the output response, but also reduced the high-frequency vibration in the system. Finally, several simulations are implemented on MATLAB Simulink with the SMC and ASMC for 6-DOF manipulator to prove the superiority of the suggested control.

ACKNOWLEDGEMENT

This research was implemented at Robotics and Intelligent Control Laboratory (RIC Lab), Faculty of Electrical and Electronics Engineering, Ho Chi Minh City University of Technology and Education, Vietnam.

REFERENCES

[1] G. Feng, A new adaptive control algorithm for robot manipulators in task space," *IEEE Transactions on Robotics and Automation*, vol. 11, no. 3, pp. 457-462, 1995, doi: 10.1109/70.388790.

[2] T. Yang, N. Sun, Y. Fang, X. Xin, and H. Chen, "New Adaptive Control Methods for n-Link Robot Manipulators With Online Gravity Compensation: Design and Experiments," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 1, pp. 539-548, 2022, doi: 10.1109/tie.2021.3050371.

[3] Q. Guo, Y. Zhang, B. G. Celler, and S. W. Su, "Neural Adaptive Backstepping Control of a Robotic Manipulator With Prescribed Performance Constraint," *IEEE Trans Neural Netw Learn Syst*, vol. 30, no. 12, pp. 3572-3583, Dec 2019, doi: 10.1109/TNNLS.2018.2854699.

[4] N. Nikdel, M. A. Badamchizadeh, V. Azimirad, and M. A. Nazari, "Adaptive backstepping control for an n-degree of freedom robotic manipulator based on combined state augmentation," *Robotics and Computer-Integrated Manufacturing*, vol. 44, pp. 129-143, 2017, doi: 10.1016/j.rcim.2016.08.007.

[5] D.-T. Tran, H.-V.-A. Truong, and K. K. Ahn, "Adaptive Backstepping Sliding Mode Control Based RBFNN for a Hydraulic Manipulator Including Actuator Dynamics," *Applied Sciences*, vol. 9, no. 6, 2019, doi: 10.3390/app9061265.

[6] K. Lochan, S. Suklabaidya, and B. K. Roy, "Sliding mode and adaptive sliding mode control approaches of two link flexible manipulator," presented at the Proceedings of the 2015 Conference on Advances In Robotics - AIR '15, 2015.

[7] A. T. Vo and H. Kang, "Adaptive Neural Integral Full-Order Terminal Sliding Mode Control for an Uncertain Nonlinear System," *IEEE Access*, vol. 7, pp. 42238-42246, 2019, doi: 10.1109/ACCESS.2019.2907565.

[8] G. Dereje, "Sliding mode control of a two dergee of freedom robot arm using permanent magnet synchronous motor," 2018.

[9] N. Yagiz, Y. Hacioglu, and Y. Taskin, "Fuzzy Sliding-Mode Control of Active Suspensions," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 11, pp. 3883-3890, 2008, doi: 10.1109/tie.2008.924912.

[10] M. Van, "An Enhanced Robust Fault Tolerant Control Based on an Adaptive Fuzzy PID-Nonsingular Fast Terminal Sliding Mode Control for Uncertain Nonlinear Systems," *IEEE/ASME Transactions on Mechatronics*, vol. 23, no. 3, pp. 1362-1371, 2018, doi: 10.1109/TMECH.2018.2812244.

[11] T. Đ. Thiện, D. M. T. , and T. T. L. , "Adaptive sliding mode control design with genetic algorithm for a 2-DOF " *VCCA*, vol. 6, 2021.

[12] J. J. Craig, *Introduction to robotics: mechanics and control*. Pearson Educacion, 2005.

[13] L. N. Truc, N. V. Quyen, and N. P. Quang, "Dynamic Model with a New Formulation of Coriolis/Centrifugal Matrix for Robot Manipulators," *Journal of Computer Science and Cybernetics*, vol. 36, no. 1, pp. 89-104, 2020, doi: 10.15625/1813-9663/36/1/14557.

ICATSD2F.109

FUSE DETECTION ON INDUSTRIAL PRODUCTION LINE WITH MACHINE LEARNING VISION

THANH HAI DIEP, THANH SON NGUYEN, NGOC HAY NGUYEN, VAN LUAN TRAN

School of Engineering, Eastern International University, Vietnam

hai.diep@eiu.edu.vn, son.nguyen@eiu.edu.vn, hay.nguyen@eiu.edu.vn, luan.tran@eiu.edu.vn

Abstract. Recently, machine learning vision has explicitly developed in the industrial application of object detection, classification, and pose estimation of objects. In this paper, we propose a method for fuse detection to check the assembly completed fuse task for each production order on the industrial production line. The fuse detection has been performed based on the YOLOv4-tiny detector. It is a real-time object recognition algorithm created for low-power and edge devices and offers the best speed and accuracy of object detection. We labeled and trained this model on our dataset. The proposed detection method improves the ability of fuse detection on the industrial production line and significantly reduces the procedure to check the completed fuses for each production order. Our experimental results show that the proposed method improves the ability of fuse detection significantly.

Keywords. Fuse Detection, Machine Learning Vision, Object Detection

1 INTRODUCTION

In recent years, smart manufacturing had a high impact on product assembly and quality assurance procedures. In industry 4.0, an expected standard capability of assembly systems is applying machine learning vision for real-time tracking of defected products to minimize failure rate [1]. Therefore, the Automatic optical inspection (AOI) technique based on machine vision has been widely used in PCBA manufacturing because it has more advantages such as accuracy, high-speed detection, and cost reduction than conventional manual inspection. The printed circuit board (PCB) is one of the key components in electronic devices. In the printed circuit board assembly (PCBA) manufacturing process, the quality control of PCBA becomes more challenging and demanding because of the small quantity of batch order production in compliance with the high-tech production system. Visual inspection is responsible for detecting defects in PCB and it mostly causes a large amount of cost [2]. Due to rapidly growing technologies in both camera production and computing systems, precision automatic machine vision gradually replaces manual visual inspection [3].

During the PCBA process, different defects [4] can be generated in various steps of the production line. The most common fault could be missing components like a capacitor, a resistor, or a fuse. The PCB automatic inspection process includes two main phases: defect detection and defect classification. In the past few decades, several different algorithms have been developed for PCB defects automatic detection. The existing techniques can be categorized into two groups: Contact methods and noncontact methods [5]. Contact methods use several fixtures with intensive setup to perform In-Circuit Testing (ICT) that requires electrical access to check the functionality of the electronic components and connectivity of the circuit [6], [7]. Numerous noncontact methods are described in [4], [5], [8], and [9], such as X-ray imaging, laminographic, ultrasonic imaging, thermal imaging, resonance-based sensing, computer vision, machine learning, and neural networks.

Detecting a fuse on a circuit board usually faces many challenges due to the object's transparency. To improve machine vision on the industrial production lines, we propose an efficient fuse detection method based on YOLOv4-tiny [19] which is a higher version of YOLO algorithms [10]. In this paper, we collected our dataset to train and test with an industrial production line based on YOLOv4-tiny. It is a real-time object detection method with high speed and accuracy for embedded devices. Experimental results of our fuse detection show that cases of missing fuse in the circuit board on the plastic back cover can be detected with

high efficiency. The fuse detection method study to apply on an industrial production line as shown in Figure 1.

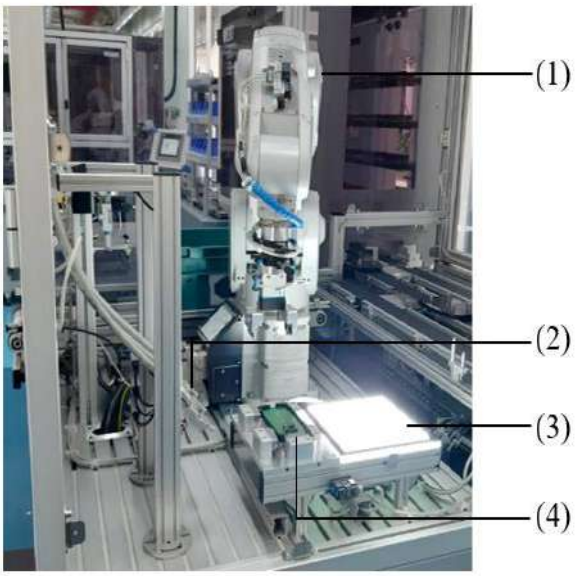


Figure 1. The automatic fuse assembly phase in the production line. (1) Robotic manipulator. (3) Fuse supplier. (3) Backlight illumination. (4) Assembly point.

2 RELATED WORKS

The Automatic optical inspection algorithms for PCB defects detection are mainly divided into three groups: Referential comparison methods, non-referential comparison methods, and hybrid methods [5]. The referential comparison method is based on the image comparison approach pixel-by-pixel by the logic operator. It requires an image without defects as a template image. Any difference between the template and the inspected image is considered a defect. In [11] Wu et al. first converted gray image to binary image for subtracting the inspected image from the template image, then detected and classified defects in PCB by using positive and the negative difference. Ibrahim et al. proposed a referential method for detecting six types of defects on PCB based on the image subtraction technique [12].

The non-referential comparison method is also called the design-rule checking technique. This technique does not need any template image to compare with. This method determines whether inspected image meets the predefined circuit design rules or not, including features of components, solder joints, and conductive pathways. In [13] Shankar and Zhong set up a series of five logic rules based on the energy features of the inspected images for detecting and classifying the defected pattern on semiconductor wafer surfaces. Such that, if the energy features of inspected image exceed 1µm, this means the inspected image has a defect.

The hybrid method combines the referential comparison approach and the non-referential comparison approach. Although it has the advantages of these two methods, this method is complex and high computational cost. Wu and Zhang [14] proposed an approach for inspecting seven types of the solder joint and component defects in PCB. A particular image acquisition system composed of a 3-CCD color digital camera and a 3-color (red, green, and blue) hemispherical LED array illumination was used for features extraction of the special solder joint image. The region, evaluation, and color grads features were extracted from the acquired images. In the next step, the Boolean rules method is used to classify solder joint types into eight different solder joint types, including one acceptable solder joint and seven defected solder joint types. The proposed method achieved a detection rate of 97.7% with a detection speed of 11 seconds per PCB. In [15] Crispin and Rankov proposed a method for detecting missing, misaligned, or incorrectly rotated components in PCB. This method uses a normalized cross-correlation template-matching that could reduce the computational cost by constraining the search space and optimizes the search strategy of template positions using a genetic algorithm.

3 FUSE DETECTION

Recently, object detection is one of the important subjects in machine learning vision. Most machine vision problems relate to detecting visual objects for classification, and localization of the object. Machine vision is an important problem for the optimization and automatic monitoring of the industrial production line of smart factories. In Industry 4.0, machine learning vision with industrial cameras is an essential component of automated production. To increase the accuracy and speed of the Convolutional Neural Network (CNN), several tests and investigations were conducted to create YOLOv4 [16]. YOLOv4 is a significant upgrade over YOLOv3 [17]. The addition of new architecture in the backbone and changes to the neck have increased the mean average precision (mAP) and the number of frames per second (FPS) by 10% and 16%, respectively. The ability to train this network on a single GPU has also improved.

In this study, fuse detection method is carried out based on YOLOv4-tiny [19]. It is the best speed and object detection accuracy are provided by the real-time object identification algorithm, which was developed for low-power and edge devices on embedded systems [20]. It can recognize and provide bounding boxes for several different objects in a single image. By segmenting a picture into areas and forecasting bounding boxes and probabilities for each region. The YOLOv4-tiny network structure as simulated in Figure 2.

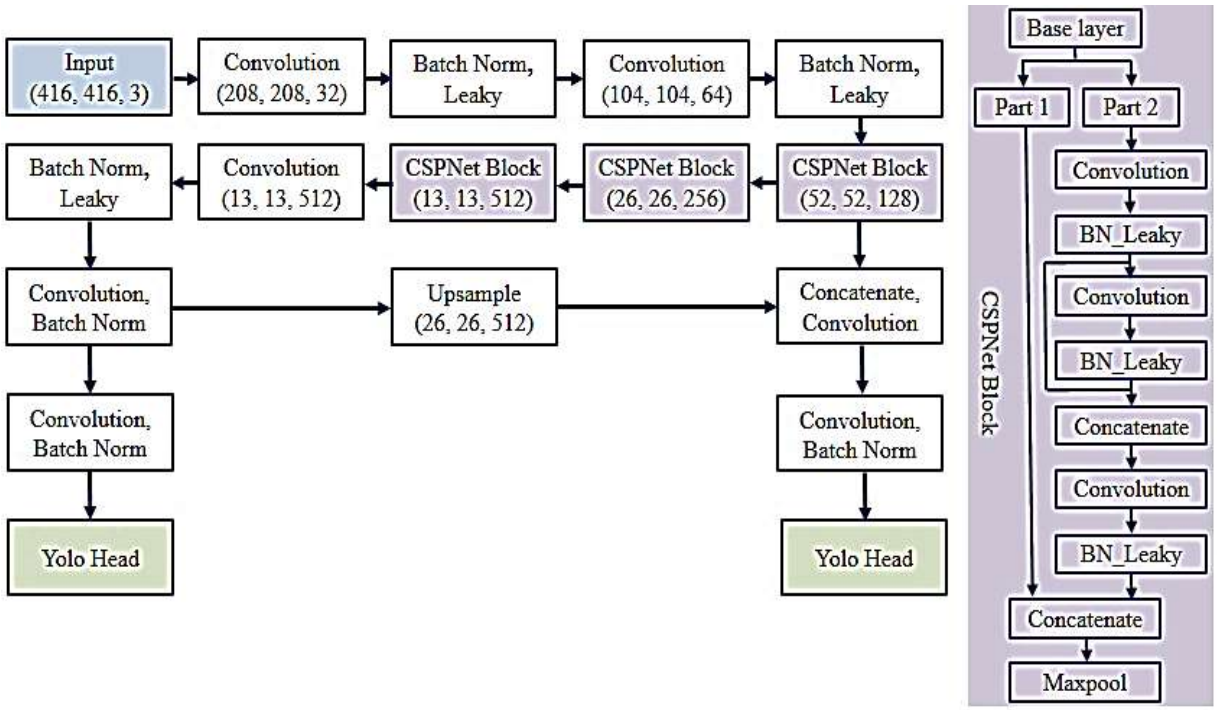


Figure 2. YOLOv4-tiny network structure and cross stage partial network (CSPNet) block.

To achieve these quick speeds, YOLOv4-tiny uses a few distinct modifications from the original YOLOv4 network. First and foremost, there are only a total of 29 pretrained convolutional layers in the Cross Stage Partial Network (CSPNet) backbone, which reduces the number of convolutional layers. The purpose of CSPNet is to ascribe the issue to redundant gradient information inside network optimization, allowing for a significant reduction in complexity while preserving accuracy. Additionally, there are fewer anchor boxes for prediction and there are only two YOLO layers now rather than three. Cross Stage Partial Network is derived from the DenseNet [21] architecture that takes the previous input and concatenates it with the current input before passing it to the Dense layer. It is responsible for dividing the input of the block into 2 parts, one part will go through the convolutional blocks, and the other part will go straight to the end of the block. Then the two parts will be concatenated together and put into the max-pooling layer. The idea here is to remove computational bottlenecks in DenseNet and improve learning by passing an unedited version of the feature map. The last module in the architecture for feature output results is called YOLO Head. The prediction is performed by YOLO Head in the one-stage detector. The outcome consists of a vector holding

the label, the confidence score, and the coordinates of the predicted bounding boxes.

4 EXPERIMENTAL RESULTS

We collected our own dataset, a novel fuse identification dataset, then trained and tested our technique on this dataset. They were gathered from various locations in the room using roughly 1400 images taken at various viewing angles and lighting levels. The images were captured in JPG format with a resolution of 850x480 pixels. Two different types of classes are included in the dataset: Fuse and Without-fuse on circuit boards. The circuit boards are shown in Figure 3. This dataset includes 1200 images for training, 100 images for verification, and 100 images for testing. We trained our network on a computer with 32GB RAM and an Intel i7-7700 CPU running at 3.6 GHz, The GPU is an NVIDIA GeForce RTX 3060 with 12GB RAM.

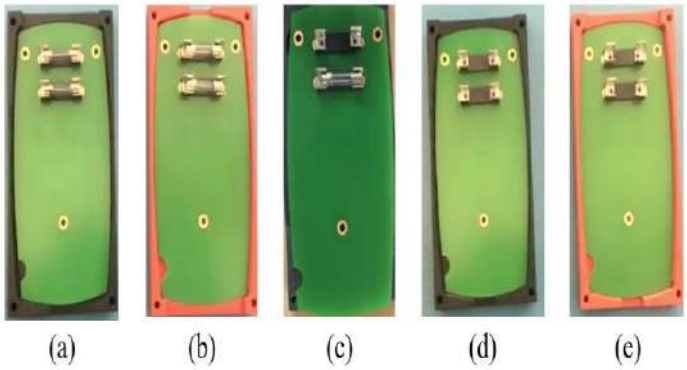


Figure 3. Some types of circuit board were tested on the production line. (a) two fuses circuit board on the black plastic cover. (b) two fuses circuit board on the red plastic cover. (c) one missing fuse circuit board on a black plastic cover. (d) two missing fuses circuit board on the black plastic cover. (e) two missing fuses circuit board on the red plastic cover.

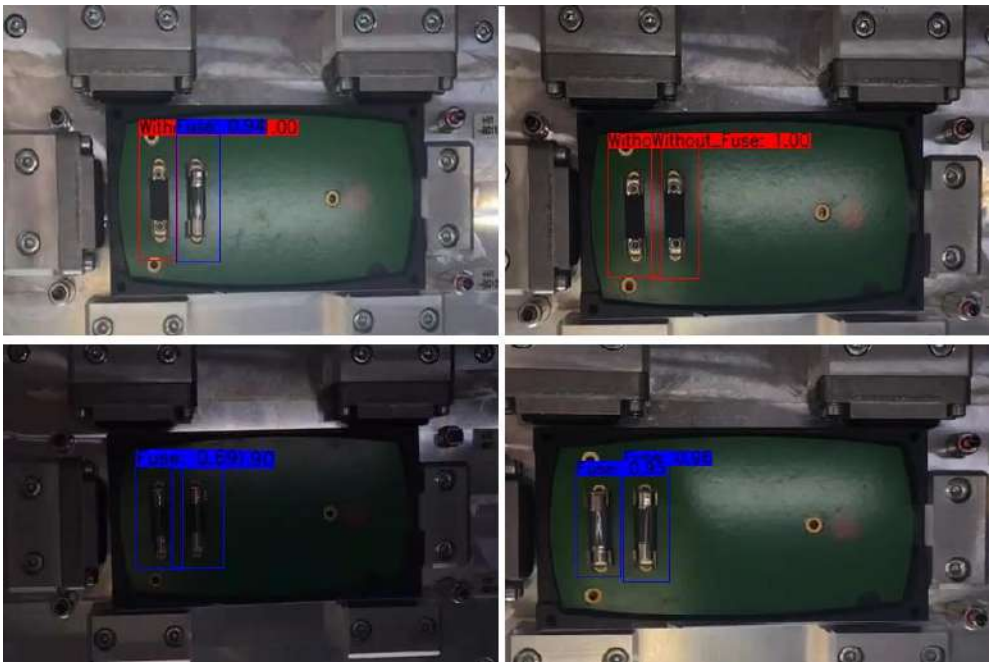


Figure 4. Our experimental results from our dataset, the red color is without a fuse on a circuit board and the blue color is without a fuse on a circuit board based on YOLOv4-tiny.

In Figure 4, we presented some of our experimental results for fuse detection on a circuit board based on YOLOv4-tiny. We can obtain a high accuracy of the fuse detection. The red color is without a fuse, the blue color is with a fuse on a circuit board. We can obtain a high accuracy of the fuse detection on a circuit

board. Figure 5 shows our experimental result, which was tested on a stored box with many circuit boards. We evaluated and compared our fuse detection on this dataset with other methods as presented in Table 1. This table tabulates the evaluation results of our fuse detection, YOLOv4 [16], YOLOv3 [17], and Faster R-CNN [18] on 100 images for testing. We can achieve high accuracy by mean Average Precision (mAP) of Fuse, Without-fuse with 90.9%, and 91.8%, respectively. The average mAPs of Ours, YOLOv4, YOLOv3, and Faster R-CNN are 91.4%, 92.1%, 88.6%, and 86.5%, respectively. The time processing of our fuse detection based on YOLOv4-tiny is 245 FPS more efficient than others as reported in Table 1. So, it is better suited for development on embedded devices and for real-time object detection.

Figure 5. Our experimental results were tested on a stored box with many circuit boards. Red and blue colors are without a fuse and with a fuse on the circuit board, respectively.

Table 1. Our evaluation and comparison of our fuse detection based on YOLOv4-tiny and the evaluation results of YOLOv4, YOLOv3, and Faster R-CNN by mean Average Precision (mAP).

Methods	Fuse	Without-fuse	mAP	FPS
Faster R-CNN	87.8%	85.1%	86.5%	7
YOLOv3	89.1%	88.6%	88.6%	45
YOLOv4	91.4%	92.7%	92.1%	50
Ours	90.9%	91.8%	91.4%	245

5 CONCLUSIONS

This paper proposed a fuse detection method to recognize fuse components in the electronics circuits. Due to the small size of the electrical components, the need for high accuracy and real-time performance in machine vision, and the difficulty in object characteristics such as transparency, affected by the light, this fuse detection work is difficult. We built a dataset of fuse components on the YOLOv4-tiny detector and conducted experiments to confirm the method's efficacy. The proposed method has a good performance with on the testing time and the accuracy rate of fuse detection on the production line are reliable enough to reduce the procedure for each production order.

ACKNOWLEDGMENT

This research was supported by Eastern International University, Binh Duong, Vietnam.

REFERENCES

[1] Y. Cohen, H. Naseraldin, A. Chaudhuri and F. Pilati, "Assembly systems in Industry 4.0 era: a road map to understand Assembly 4.0", The International Journal of Advanced Manufacturing Technology, vol. 105, no. 9, pp. 4037-4054, 2019.

[2] R. T. Chin and C. A. Harlow, "Automated Visual Inspection: A Survey," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-4, no. 6, pp. 557-573, 1982.

[3] S. Shirmohammadi and A. Ferrero, "Camera as the instrument: the rising trend of vision-based measurement," in IEEE Instrumentation & Measurement Magazine, vol. 17, no. 3, pp. 41-47, 2014

[4] T. Qiu, C. Tek and S. Huang, "A Compact High-Resolution Resonance-Based Capacitive Sensor for Defects Detection on PCBAs", IEEE Access, vol. 8, pp. 203758-203768, 2020.

[5] M. Moganti, F. Ercal, C. Dagli and S. Tsunekawa, Automatic PCB Inspection Algorithms: A Survey, *Computer Vision and Image Understanding*, vol. 63, no. 2, pp. 287-313, 1996

[6] J. E. Monagas Martin, A. S. Nevado, and A. V. Martinez, "Low-cost programmable modular system to perform in-circuit test (ICT) full development of the hardware, software and mechanics of an ICT machine,'' in Proc. Technol. Appl. Electron. Teaching (TAEE), pp. 1–7, 2016.

[7] A. J. Albee, ''The evolution of ICT: PCB technologies, test philosophies, and manufacturing business models are driving in-circuit test evolution and innovations,'' in Proc. Conf. Exhib. (IPC APEX EXPO), vol. 1, 2013, pp. 381–401.

[8] D. B. Anitha and M. Rao, "A survey on defect detection in bare PCB and assembled PCB using image processing techniques," 2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), pp. 39-43, 2017.

[9] A. Liu et al., "X-ray inspection methods for controlling PCBA potting process — 2DX and partial angle computer tomography," 2016 Pan Pacific Microelectronics Symposium (Pan Pacific), 2016, pp. 1-5,

[10] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779-788, 2016.

[11] W. Wu, M. Wang and C. Liu, ''Automated inspection of printed circuit boards through machine vision", Computers in Industry, vol. 28, no. 2, pp. 103-111, 1996.

[12] Ibrahim, I., Rahman, S.A., Bakar, S.A., Mokji, M.M., Mukred, J.A.A., Yusof, Z., Ibrahim, Z., Khalil, K., Mohamad, M.S.: A printed circuit board inspection system with defect classification capability. Int. J. Innov. Manag. Inf. Prod. ISME Int. 3(1), pp. 82-87, 2012

[13] N. Shankar and Z. Zhong, "A rule-based computing approach for the segmentation of semiconductor defects", Microelectronics Journal, vol. 37, no. 6, pp. 500-509, 2006.

[14] F. Wu and X. Zhang, "An inspection and classification method for chip solder joints using color grads and Boolean rules", Robotics and Computer-Integrated Manufacturing, vol. 30, no. 5, pp. 517-526, 2014.

[15] A. Crispin and V. Rankov, "Automated inspection of PCB components using a genetic algorithm template-matching approach", The International Journal of Advanced Manufacturing Technology, vol. 35, no. 3-4, pp. 293-300, 2006.

[16] A. Bochkovskiy, C. Wang, and H. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," CoRR, vol. abs/2004.10934, 2020.

[17] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," CoRR, vol. abs/1804.02767, 2018.

[18] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137-1149, 2017.

[19] S. Saponara, A. Elhanashi, and Z. Qinghe, "Developing a real-time social distancing detection system based on yolov4-tiny and bird-eye view for covid-19", Journal of Real-Time Image Processing, 19:1–13, 2022.

[20] Z. Jiang, L. Zhao, S. Li, and Y. Jia. Real-time object detection method based on improved yolov4-tiny. CoRR, abs/2011.04244, 2020.

[21] G. Huang, Z. Liu, L. Van Der Maaten and K. Weinberger, "Densely Connected Convolutional Networks", 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 20

ICATSD2F.110

FRÉCHET MEANS IN SUB-RIEMANNIAN MANIFOLD

THANH-SON TRINH

Faculty of Information Technology, Industrial University of Ho Chi Minh City

trinhthanhson@iuh.edu.vn

Abstract. Riemannian center of mass or the Fréchet means of manifold has many applications in machine learning, computer vision and medical image analysis ([6], [14], [17]). In this paper, we study Fréchet means in sub-Riemannian manifold (M, Δ, g). Let d_{SR} be the sub-Riemannian distance, for an integer number $N \geq 2$ and $x_1, x_2, \dots, x_N \in M$, we consider the following problem

$$\inf_{z \in M} \frac{1}{N} \sum_{i=1}^{N} d_{SR}^2(z, x_i).$$

First, we show that the existence of Fréchet means $y \in arg\min_{z \in M} \frac{1}{N} \sum_{i=1}^{N} d_{SR}^2(z, x_i)$. After that, as Kim and Pass did for manifolds [12], under some mild conditions we prove that Fréchet means y are not belongs to sub-Rieamnnian cut locus set of x_i for $i = 1,2, \dots, N$.

Keywords. Fréchet mean, Sub-Rieamannian manifold, Cut locus.

1 INTRODUCTION

The Fréchet mean is the set of global minima of the variance, which was introduced by Fréchet in general metric spaces, in 1948 [9]. In the case of Riemannian manifolds, the Fréchet mean problem is presented as follows. Let N ($N \geq 2$) points $x_1, \dots, x_N$ in a manifold M, we consider the following problem:

$$\inf_{z \in M} \frac{1}{N} \sum_{i=1}^{N} d_R^2(x_i, z),$$

where d_R is the Riemannian metric on M.

Riemannian center of mass has applications in the data processing which lie in Riemannian manifolds. Since forms of data appear in robotics, computer graphics, medical imaging, statistical analysis of shapes…, finding the mean in Riemannian manifolds has applications in many fields ([6], [14], [17]).

After works of Fréchet, the theory and application of Fréchet mean in Riemannian manifolds have been studied extensively by many authors such as Karcher, Afsari, Yang, Arnaudon, Miclo ([11], [1], [8], [4]). Recently, this problem had many applications in the theory of optimal transport, see [12] and [13].

The aim of this paper is to give some properties of Fréchet means in sub-Riemannian manifolds. The theory of sub-Riemannian manifold (M, Δ, g) will be presented in Section 2. We now introduce our problem as follows. Let $N \geq 2$ be an integer number and $x_1, \dots, x_N \in M$, we consider the following problem:

$$\inf_{z \in M} \frac{1}{N} \sum_{i=1}^{N} d_{SR}^2(z, x_i),$$

where d_{SR} is the sub-Riemannian metric.

Our first main result is the existence of Fréchet means in sub-Riemannian manifolds.

Theorem 1.1. *Assume that (M, d_{SR}) is complete metric space. Let $(x_1, \dots, x_k) \in M^k := M \times \cdots \times M$ then there exists a Borel map $T: M^k \to M$ that associate $(x_1, \dots, x_k)$ to a minimum of $y \mapsto \frac{1}{N} \sum_{i=1}^{k} d_{SR}^2(x_i, y)$.*

In our second main result, we show that Fréchet means $y \in arg\min_{z \in M} \frac{1}{N} \sum_{i=1}^{N} d_{SR}^2(z, x_i)$ do not belongs to sub-Rieamannian cut locus set of x_i for $i = 1,2, \dots, N$.

Theorem 1.2. *Let $(x_1, \dots, x_k) \in M^k$ and suppose that y is a minimizer of $z \mapsto \frac{1}{N}\sum_{i=1}^{k} d_{SR}^2(x_i, z)$. If (Δ, g) is ideal and $y \neq x_i$, for every $i \in \{1, \dots, k\}$ then $y \notin Cut_{SR}(x_i)$ for every $i \in \{1, \dots, k\}$.*

This theorem is a version of [12, Lemma 3.1] in sub-Riemannian manifolds.

Our paper is organized as follows. In Section 2, we review on sub-Riemannian manifolds, it includes sub-Riemannian geometry, sub-Riemannian distance, sub-Riemannian minimizing geodesics and sub-Riemannian cut locus. Then, we recall basic facts on semi concavity in Section 3. In Section 4, we give proofs of Theorem 1.1 and Theorem 1.2. Finally, we give conclusion in Section 5.

2 PRELIMINARIES

2.1 Sub-Riemannian geometry

In this sub section, we review on sub-Riemannian structure. For details, readers can see in [2], [3], [5] and [15].

In this paper, we denote M by a smooth connected manifold of dimension $n \geq 2$. A sub-Riemannian manifold is given by a triple (M, Δ, g), where Δ is a totally nonholonomic on M and g is a Riemannian metric on M. We recall that a smooth distribution of rank $m \leq n$ is a rank m sub bundle of the tangent bundle TM, this means that for every $x \in M$there exists a neighborhood $V_x$of x in $M$and m smooth vector fields $(f_x^1, \dots, f_x^m)$ linearly independent on V_x such that:

$$\Delta(z) = span\{f_x^1(z), \dots, f_x^m(z)\}, \text{ for every} z \in V_x.$$

Then we say that the distribution Δ is represented locally by $(f_x^1, \dots, f_x^m)$. We call that a distribution Δ on M admits a global frame if there exists m smooth vector fields $f_1, \dots, f_m$ on M such that:

$$\Delta(x) = span\{f_1(x), \dots, f_m(x)\}, \text{ for every } x \in M.$$

A distribution Δ on M is nonholonomic (or totally nonholonomic) if for every $x \in M$ there exists $(f_x^1, \dots, f_x^m)$ of smooth vector fields on V_x which represents locally Δ and such that:

$$Lie\{f_x^1, \dots, f_x^m\}(z) = T_z M, \text{ for every } z \in V_x.$$

2.2 Sub-Riemannian distance

A *horizontal* curve $\theta: [0,1] \to M$ is a continuous curve such that there exists $u \in L^2([0,1], \mathbb{R}^m)$ satisfying that

$$\dot{\theta}(t) = \sum_{i=1}^{m} u_i(t) f_i(\theta(t)), \text{ a.e. } t \in [0,1].$$

The element u is unique in the sense that if there is another such $v \in L^2([0,1], \mathbb{R}^m)$ then $u(t) = v(t)$ a.e. $t \in [0,1]$. We call that u is the control of θ.

If θ is horizontal then by [2, Lemma 3.11] one gets that the map $t \mapsto \sqrt{g\big(\dot{\theta}(t), \dot{\theta}(t)\big)}$ is measurable on $[0,1]$ and thus integrable. Hence, we can define its length by

$$l(\theta) := \int_0^1 \sqrt{g\big(\dot{\theta}(t), \dot{\theta}(t)\big)}\, dt.$$

For every $x, y \in M$ the *sub-Riemannian distance* is defined by

$$d_{SR}(x, y) := inf\{l(\theta) | \theta \text{ is a horizontal, } \theta(0) = x, \theta(1) = y\}.$$

We note that since θ is continuous, its length $l(\theta)$ is finite. From M is connected and Δ is totally nonholonomic distribution on M, by Chow-Rashevsky Theorem [15, Theorem 1.14], for every $x, y \in M$ there exists a horizontal path $\theta: [0,1] \to M$ such that $\theta(0) = x$ and $\theta(1) = y$. This means that the sub-Riemannian distance is finite and furthermore, it is continuous on $M \times M$ (see [15, Proposition 1.16]).

2.3 Sub-Riemannian minimizing geodesics

For each $x, y \in M$, a horizontal path $\theta: [0,1] \to M$ is called a *sub-Riemannian minimizing geodesic* between x and y if $d_{SR}(x, y) = l(\gamma)$.

Definition 2.1. *A metric space (M, d_{SR}) is called a geodesic space if for every $x, y \in M$ there exists a sub-Riemannian minimizing geodesic between x and y.*

Let $x, y \in M$ and $\theta_u: [0,1] \to M$ be a horizontal path joining x and y, where $u \in L^2([0,1], \mathbb{R}^m)$ is its control. By [15, Proposition 1.7], there exists the maximal open subset $U_x \subset L^2([0,1], \mathbb{R}^m)$ such that for every $v \in U_x$, the Cauchy problem

$$\dot{\theta}_v(t) = \sum_{i=1}^{m} v_i(t) f_i(\theta_u(t)),\ \theta_v(0) = x,$$

has a well-defined unique solution for a.e. $t \in [0,1]$. The *end-point map* at x is the map $E_x: U_x \to M$ defined by $E_x(v) := \theta_v(1)$ for every $v \in U_x$. Since [15, Proposition 1.8], the end-point mapping E_x is of class C^1 on U_x. For $u \in U_x$, we define:

$$Im_x(u) := D_u E_x\left(L^2([0,1], \mathbb{R}^m)\right).$$

Then $Im_x(u)$ is a vector subspace of $T_y M$ with $y = E_x(u)$. The dimension of $Im_x(u)$ is called *rank* of u and is denoted by $rank(u)$. We say that a horizontal path $\theta_u: [0,1] \to M$ with control $u \in U_x$ is *singular* if u is a critical point of E_x, i.e., $rank(u) < n$. In this case, θ_u is called an *abnormal* path.

Definition 2.2. *A sub-Riemannian structure (Δ, g) on M is ideal if it has no non-trivial abnormal minimizing geodesic and the metric space (M, d_{SR}) is complete.*

2.4 Sub-Riemannian cut locus

We now review sub-Riemannian cut locus. For every $x \in M$, the cut locus of x is defined by

$$Cut_{SR}(x) := \overline{\{y \in M | d_{SR}(x,y) is not continuously differentiable at y\}},$$

and we define the global cut locus of M as follows

$$Cut_{SR}(x) := \{(x,y) \in M | y \in Cut_{SR}(x)\}.$$

The following result is well-known (see [5, Theorem 19] and [8, Proposition 5.10]).

Theorem 2.1. *The squared sub-Riemannian distance is smooth on $M \times M / Cut_{SR}(M)$.*

3 LOCALLY SEMICONCAVE FUNCTIONS

In this section, we recall basic properties on semi concavity. This is a main tool to prove our second main result Theorem 1.2.

Let U be an open subset of $\mathbb{R}^n$. A function $f: U \to \mathbb{R}$ is locally semiconcave if for every $z \in U$, there exists $C, r > 0$ such that

$$tf(y) + (1-t)f(x) - f(tx + (1-t)y) \le t(1-t)C|x-y|^2,$$

for every $t \in [0,1]$, $x, y \in B(z,r)$.

Let M be a smooth connected manifold of dimension n. Let U be an open subset of M, a function $f: U \to M$ is locally semiconcave if for every $x \in U$, there exists a neighborhood U_x of x and smooth diffeomorphism $\varphi_x: U_x \to \varphi_x(U_x) \subset \mathbb{R}^n$ such that $f \circ \varphi_x^{-1}$ is locally semiconcave on the open subset $\varphi_x(U_x)$ of $\mathbb{R}^n$.

Lemma 3.1. ([16, Proposition 9]) *For every $x \in M$ the squared sub-Riemannian distance $d_{SR}^2(x,\cdot)$ is locally semiconcave on $M/\{x\}$.*

Lemma 3.2. ([8, Lemma A.1]) *Let U be an open subset of $\mathbb{R}^n$ and a function $f: U \to \mathbb{R}$. Suppose that for every $z \in U$ there exist a neighborhood $V_z \subset U$ and $\sigma > 0$ such that, for every $x \in V_z$, there is $\rho_x \in \mathbb{R}^n$ such that*

$$f(y) \le f(x) + \langle \rho_x, y - x \rangle + \sigma |y-x|^2, \text{ for every } y \in V_z.$$

Then the function f is locally semiconcave on U.

4 PROOFS OF MAIN RESULTS

We will now prove that the existence of Fréchet means $y \in arg\min_{z \in M} \frac{1}{N} \sum_{i=1}^{N} d_{SR}^2(z, x_i)$ by giving proof of Theorem 1.1.

Proof of Theorem 1.1. As (M, d_{SR}) is complete, applying [15, Theorem 2.4] we get that for every $x, y \in M$ there exists a horizontal path $\theta: [0,1] \to M$ such that $\theta(0) = x$, $\theta(1) = y$ and

$$d_{SR}(x,y) = l(\theta).$$

This implies that (M, d_{SR}) is a geodesic space. Moreover, the topology defined by the distance d_{SR} coincides with the original topology of M (see [3, Theorem 3.31] or [15, Theorem 2.4]). Hence, (M, d_{SR}) is a separable space.

Thanks to [2, Corollary 3.37], the metric space (M, d_{SR}) is a locally compact space. Hence, (M, d_{SR}) is a separable locally compact geodesic space. Therefore, by [10, Lemma 7] we get the results. □

Next, we will prove Theorem 1.2.

Proof of Theorem 1.2. Argue by contradiction, we assume that there exists $y \in \{1, \dots, k\}$ such that $y \in Cut_{SR}(x_j)$.

For each $i \in \{1, \dots, k\}$, setting $\Omega_i := M \backslash \{x_i\}$. By using Lemma 3.1 we obtain that $d_{SR}^2(x_i, \cdot)$ is locally semiconcave in Ω_i. Hence, up to working in chart, we get that there exists $p \in \mathbb{R}^n$ and $C > 0$ such that:

$$d_{SR}^2(x_i, y + v) - d_{SR}^2(x_i, y) \leq p \cdot v + \frac{N}{2} C_i |v|^2, \text{ for every } |v| < 1.$$

By replacing v by $-v$ one has that:

$$d_{SR}^2(x_i, y - v) - d_{SR}^2(x_i, y) \leq -p \cdot v + \frac{N}{2} C_i |v|^2, \text{ for every } |v| < 1.$$

This yields that:

$$d_{SR}^2(x_i, y + v) + d_{SR}^2(x_i, y - v) - 2d_{SR}^2(x_i, y) \leq N C_i |v|^2, \text{ for every } |v| < 1.$$

Putting $C := max\{C_1, \dots, C_k\}$, then since $y \in Cut_{SR}(x_j)$ and $y \neq x_i$, applying [5, Theorem 6] we get that there exists $v \in \mathbb{R}^n$, $|v| < 1$ such that:

$$\frac{d_{SR}^2(x_j, y+v) + d_{SR}^2(x_j, y-v) - 2d_{SR}^2(x_j, y)}{|v|^2} \leq -k \cdot C \cdot N.$$

Therefore,

$$\begin{aligned}
\frac{2}{N} \sum_{i=1}^{N} d_{SR}^2(x_i, y) &= \frac{2}{N} \sum_{i \neq j}^{N} d_{SR}^2(x_i, y) + \frac{2}{N} d_{SR}^2(x_j, y) \\
&\geq |v|^2 \left(N \cdot C - \sum_{i \neq j}^{N} C_i \right) + \frac{1}{N} \sum_{i=1}^{N} d_{SR}^2(x_i, y + v) + \frac{1}{N} \sum_{i=1}^{N} d_{SR}^2(x_i, y - v) \\
&> \frac{1}{N} \sum_{i=1}^{N} d_{SR}^2(x_i, y + v) + \frac{1}{N} \sum_{i=1}^{N} d_{SR}^2(x_i, y - v).
\end{aligned}$$

This is a contradiction with y minimizes $z \mapsto \frac{1}{N} \sum_{i=1}^{k} d_{SR}^2(x_i, z)$. Hence, we get the result. □

5 CONCLUSIONS

In this work, we study the Fréchet means on sub-Riemannian manifolds. The contributions of our paper are two-fold.

1) We show that the existence of Fréchet means over sub-Riemannian manifolds structure.

2) By considering the semi concavity properties of the square sub-Riemannian distance d_{SR}^2, we show an important relationship between sub-Riemannian means and sub-Rieamannian cut locus. This result is an extension of [12, Lemma 3.1]. Moreover, our second main result plays an important role in applying the theory of sub-Riemannian manifolds to study multi-marginal transport on sub-Riemannian manifolds and Wasserstein barycenter over sub-Riemannian structure (see [12] and [13]).

However, in our second main result, we need a condition that Fréchet means y must be different from N points $x_1, \dots, x_N$. Without this condition, we do not know that the result in Theorem 1.2 is correct or not? This question remains open.

As we mention in the introduction part, Fréchet means has many applications in optimal transport, we can apply main results in this paper to study optimal transport over sub-Riemannian structure. These works can be considered in the future.

REFERENCES

[1] B. Afsari, *Riemannian L^p center of mass: existence, uniqueness, and convexity*, Proceedings of the American Mathematical Society, 139 (2011), no. 2, 655-673.

[2] Andrei Agrachev, Davide Barilari, and Ugo Boscain, *Introduction to geodesics in sub-Riemannian geometry*, Geometry, analysis and dynamics on sub-Riemannian manifolds. Vol. II, EMS Ser. Lect. Math., Eur. Math. Soc., Zurich, 2016, 1 – 83.

[3] Andrei Agrachev, Davide Barilari, and Ugo Boscain, *A Comprehensive Introduction to Sub-Riemannian Geometry*, Cambridge Studies Advanced Mathematics, vol. 181, Cambridge University Press, 2019.

[4] M. Arnaudon and L. Miclo, *Means in complete manifolds: Uniqueness and Approximation*, ESAIM: Probability and Statistics, PS 18 (2014) 185-206.

[5] Davide Barilari and Luca Rizzi, *Sub-Riemannian interpolation inequalities*, Invent. Math. 215 (2019), no. 3, 977-1038.

[6] Rudrasis Chakraborty and Baha C. Vemuri, *Recursive Fréchet mean computation on the grassmannian and its applications to computer vision*, Proceedings of the IEEE International Conference on Computer Vision (2015), 4229–4237.

[7] Nhan-Phu Chung and Thanh-Son Trinh, *Unbalanced optimal total variation transport problems and generalized Wasserstein barycenters*, Proceedings of the Royal Society of Edinburgh. Section A: Mathematics. 152.3 (2022), 674–700.

[8] Alessio Figalli and Ludovic Rifford, *Mass transportation on sub-Riemannian manifolds*, Geom. Funct. Anal. 20 (2010), no. 1, 124–159.

[9] M. Fréchet, *Les éléments aléatoires de nature quelconque dán un espace distancié*, Annales de l'Institut Henri Poincaré. 10 (1948), 215–310.

[10] Thibaut Le Gouic and Jean-Michel Loubes, *Existence, and consistency of Wasserstein barycenters*, Probab. Theory Related Fields 168 (2017), no. 3 – 4, 901–917.

[11] H. Karcher, *Riemannian center of mass and mollifier smoothing*, Com. Pure. Appl. Math. 30 (1977), 509-541

[12] Young-Heon Kim and Brendan Pass, *Multi-marginal optimal transport on Riemannian manifolds*, Amer. J. Math. 137 (2015), no. 4, 1045–1060.

[13] Young-Heon Kim and Brendan Pass, *Wasserstein barycenters over Riemannian manifolds*, Adv. Math. 307 (2017), 640–683.

[14] Aaron Lou, Isay Katsman, Qingxuan Jiang, Serge Belongie, Ser-Nam Lim, and Christoppher De Sa, *differentiating through the Fréchet Mean*, Proceedings of the 37th International Conference on Machine Learning (2020).

[15] Ludovic Rifford, *Sub-Riemannian geometry and optimal transport*, Springer Briefs in Mathematics, Springer, Cham, 2014.

[16] L. Rifford and E. Trélat, *On the stabilization problem for nonholonomic distributions*, J. Eur. Math. Soc. (JEMS) 11 (2009), no. 2, 223-255.

[17] Hesamoddin Salehian, Rudrasis Chakraborty, Edward Ofori, David Vaillancourt, and Baha C. Vemuri, *An efficient recursive estimator of the Fréchet mean on a hypersphere with applications to Medical Image Analysis*, Mathematical Foundations of Computational Anatomy 3 (2015).

[18] Le Yang, *Some properties of Fréchet medians in Riemannian manifolds*, arXiv: 1110.3899

ICATSD2F.111

PERFORMANCE ANALYSIS OF TWO-WAY NETWORK WITH NONLINEAR ENERGY HARVESTING RELAY AND DIGITAL NETWORK CODING

THU- THUY THI DAO[1,2], PHAM NGOC SON[2]

[1]*Industrial University of Ho Chi Minh City, Ho Chi Minh City, Vietnam*

[2]*Ho Chi Minh City University of Technology and Education, Ho Chi Minh City, Vietnam*

daothithuthuy@iuh.edu.vn, thuydtt.ncs@hcmute.edu.vn, sonpndtvt@hcmute.edu.vn

Abstract. This article studies a two-way cooperative scheme including two sources and the assist of a decode-and-forward relay which equipped a non-linear energy harvester. This scheme operates under a half-duplex mode and has three-time slots to communicate information. In two first slots, the relay uses power splitting method to harvest energy and decode signal from sources, then the digital network coding technique is used to make a new signal by the XOR operation. The relay uses harvested energy in two first time slots to send back the encoded signal to two sources. The exact expressions of the outage probabilities are analyzed to evaluate the effect of a non-linear energy harvesting on the system performance over Rayleigh fading channels. Besides, the outage probabilities of two sources are compared with cases in linear energy harvesting. Finally, the Monte Carlo simulation results verify the exactly of the analysis expressions.

Keywords. two-way network, nonlinear energy harvesting, decode-and-forward relay, digital network coding.

1 INTRODUCTION

In recent years, the fast growth number of users and applications of wireless communication have led to more challenges for wireless networks as improving system performance, spectrum efficiency and energy efficiency. Specially, energy efficiency are critical problems for low-cost and resource-constrained wireless networks like wireless sensor networks and Internet of Things systems. Energy harvesting (EH) via radio-frequency signals in wireless cooperative networks has been appealing many attentions, the relay can harvest energy from the sources signals and then spend the harvested energy to aid its transmission. There are two main methods to energy harvesting at relay base on time switching (TS) and power splitting (PS) techniques. Relay node divides the received signal into two-time phases for EH and signal decoding in TS method, whereas it splits the received signal power for EH and signal decoding in PS method [1]. The cooperative network has been well-known solution for these challenges because it brings an improvement in the performance, coverage, and reliability of wireless networks. There are two methods for relay nodes to process signals as decode-and-forward (DF) and amplify-and-forward (AF) [2]. Besides, the two-way relay network (TWR) combines with the digital network coding (DNC) technique bring to rise bandwidth use efficiency by reducing transmission time among devices [3].

In [4, 5], the performance of EH schemes for a two-way half-duplex relay model with two sources and a relay had been analyzed over block Rayleigh fading channels. The relay harvests the energy from two sources by PS protocol then using that energy to send the information back sources. In [4], the DF relay collects energy from the received signals of two sources in the first two time-slots and then uses the acquired energy and DNC technique for signal transmission in the third one. In [5], the authors use the direct channel between two sources to improve the overall performance and consider three relay styles as AF, DF and hybrid-DF. However, the authors in two papers studied the linear EH model, this may be unrealistic as the electronic devices executed in the energy harvester at the relay are nonlinear elements. The nonlinear energy-harvesting at relays was studied in [6-8]. In [6], the DF relay collects energy from the source and interference signals with TS technique in one-way network over Nakagami-m fading channels. In [7], the authors studied a hybrid TS-PS harvest energy scheme in case of transceiver hardware impairments of two-way AF relay network. In [8], the authors analyzed the outage probability (OP) of a three-step two-

way DF relay network with the nonlinear EH using PS method, where the PS ratios are animatedly changed according to instantaneous channel gains to maximize the capacity of the system.

Motivated by prior works for the two-way network to enhance spectrum utilization efficiency, system performance and energy efficiency, we propose a two-way DF relay system with two sources and a relay which equipped a non-linear energy harvester. In the proposed model, there are three-time slots to communicate information, in two first slots the relay uses PS method to harvest energy and decode signal from to sources, then it uses the DNC technique to create a new signal and uses harvested-energy to send back this signal to two sources, called as NEH-TW-DNC protocol. To evaluate the system performance, the exact expressions of the source OPs are considered. Additionally, we compare the OP of the proposed protocol with the ideal linear EH ones. Our proposed system can be applicated in low-cost and resource-constrained wireless networks like wireless sensor networks, vehicular networks and Internet of Things systems.

The rest of our article includes sections as follows. A two-way cooperative model with nonlinear energy harvesting and DNC technique is described in Section 2. Then Section 3 analyzes the OPs of the two sources. Section 4 displays the results and discussions. Lastly, Section 5 is conclusions.

The notations: $f_{\square}(\cdot)$and $F_{\square}(\cdot)$are the probability density function (PDF) and the cumulative distribution function (CDF) of a random variable (RV) $\square$, respectively; $Pr[\ \ ]$ is the probability operation; $\Gamma[.,.]$ is the upper incomplete Gamma function [9] (eq. 8.350.2) ; $X \sim CN(0,\sigma^2)$is a complex Gaussian random variable (RV) X with zero mean and variance σ^2.

2 SYSTEM MODEL

The system model of the NEH-TW-DNC scheme gives in Figure 1. In this model, there are two sources S_1,S_2 communicate to each other through the DF relay R. The sources and the relay have a single antenna. The direct link among the sources S_1 and S_2 is absent owing to severe fading and path-loss, so communication among S_1 and S_2 just be established via relay. We denote h_{RS_i}, h_{S_iR} and d_i as the fading channel coefficients and the normalized distances between S_i and R respectively, where $i \in \{1,2\}$. The relay has a limited power and so it must harvest the energy from the radio frequency signals of two sources to send the information of these sources by working the three-time slots protocol as Table 1.

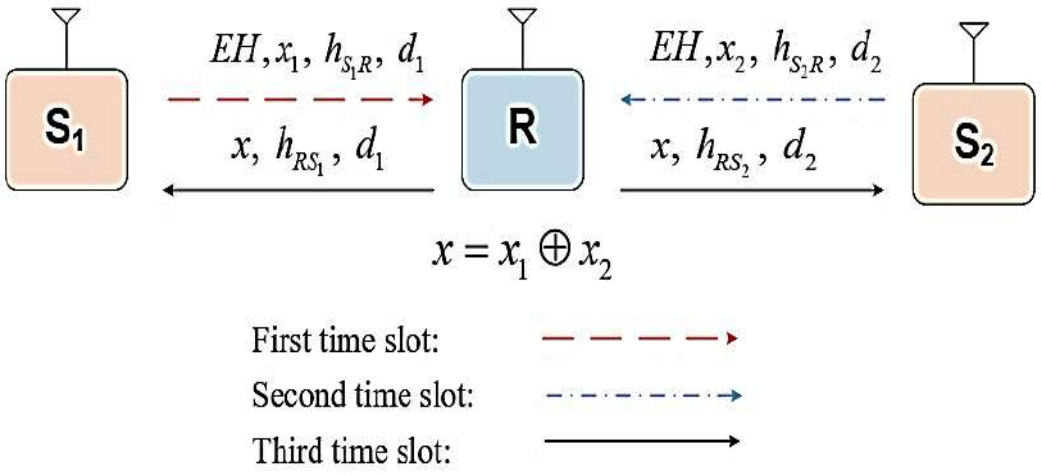


Figure 1. System model of the NEH-TW-DNC scheme.

Table 1: The three-time slots protocol of the NEH-TW-DNC scheme.

$S_1 \rightarrow R$	$S_2 \rightarrow R$	$R \rightarrow S_1, S_2$
NEH at Relay (ρ) Decoding signal x_1 from S_1	NEH at Relay (ρ) Decoding signal x_2 from S_2	$x = x_1 \oplus x_2$ Relay forwards signal x to sources
Time slot 1 αT	Time slot 2 αT	Time slot 3 $(1-2\alpha T)$
T		

The first time slot

The receiver signals at R

$$y_{S_1R} = \sqrt{P_1} h_{S_1R} x_1 + n_R, \tag{1}$$

where n_R is the additive Gaussian noise $CN(0, N_0)$ at the relay.

The signal-noise ratio (SNR) at R

$$\gamma_{S_1R} = \frac{(1-\rho)|h_{S_1R}|^2 P_1}{N_0}. \tag{2}$$

The EH at R in the first time slot:

$$E_{h1} = \eta\rho\alpha P_1 \left|h_{S_1R}\right|^2 T, \tag{3}$$

where $0 < \eta \leq 1$and $0 < \rho < 1$ denote the energy conversion efficiency and the power allocation efficiency, respectively.

The second time slot

The receiver signals at R:

$$y_{S_2R} = h_{S_2R} x_2 + n_R. \tag{4}$$

The SNR at R:

$$\gamma_{S_2R} = \frac{(1-\rho)\left|h_{S_2R}\right|^2 P_2}{N_0}. \tag{5}$$

The EH at R in the second time slot:

$$E_{h2} = \eta\rho\alpha P_2 \left|h_{S_2R}\right|^2 T. \tag{6}$$

From **Error! Reference source not found.** and **Error! Reference source not found.** we have sum of EH at the R following:

$$E_h = E_{h1} + E_{h2} = \eta\rho\alpha\left(P_1\left|h_{S_1R}\right|^2 + P_2\left|h_{S_2R}\right|^2\right)T. \tag{7}$$

Reality, the electronic devices used in the energy harvester at the relay are nonlinear elements. If the values of the input power are under the saturation threshold power P_{th}, the harvested energy is linearly relative to the input power. When the saturation effect happens, the harvested energy does not rise though the input power still rises. Hence, the transmit power of the relay is presented as [6]:

$$P_R = \frac{E_h}{(1-2\alpha)T} = \begin{cases} \dfrac{\eta\rho\alpha}{1-2\alpha}\left(P_1\left|h_{S_1R}\right|^2 + P_2\left|h_{S_2R}\right|^2\right) & if \quad \left(P_1\left|h_{S_1R}\right|^2 + P_2\left|h_{S_2R}\right|^2\right) \leq P_{th} \\ \dfrac{\eta\rho\alpha}{1-2\alpha} P_{th} & if \quad \left(P_1\left|h_{S_1R}\right|^2 + P_2\left|h_{S_2R}\right|^2\right) > P_{th} \end{cases}. \tag{8}$$

Assuming that the power of two source equal $P = P_1 = P_2$

$$P_R = \begin{cases} \varsigma\ P\left(\left|h_{S_1R}\right|^2 + \left|h_{S_2R}\right|^2\right) & if \quad P\left(\left|h_{S_1R}\right|^2 + \left|h_{S_2R}\right|^2\right) \le P_{th} \\ \varsigma P_{th} & if \quad P\left(\left|h_{S_1R}\right|^2 + \left|h_{S_2R}\right|^2\right) > P_{th} \end{cases}, \tag{9}$$

where $\varsigma = \frac{\eta\rho\alpha}{1-2\alpha}$.

The third time slot, the relay creates an encoded signal $x = x_1 \oplus x_2$ by the DNC techniques [10] and uses P_R to broadcasts this signal to two sources.

$$y_{RS_i} = \sqrt{P_R} h_{RS_i} x + n_{RS_i}, \tag{10}$$

$$\gamma_{RS_i} = \frac{\left|h_{RS_i}\right|^2 P_R}{N_0} = \begin{cases} \frac{\left|h_{RS_i}\right|^2}{N_0} \varsigma\ P\left(\left|h_{S_1R}\right|^2 + \left|h_{S_2R}\right|^2\right) & if \quad P\left(\left|h_{S_1R}\right|^2 + \left|h_{S_2R}\right|^2\right) \le P_{th} \\ \frac{\left|h_{RS_i}\right|^2}{N_0} \varsigma P_{th} & if \quad P\left(\left|h_{S_1R}\right|^2 + \left|h_{S_2R}\right|^2\right) > P_{th} \end{cases}. \tag{11}$$

The SNR at the source S_1, S_2 can be obtained as:

$$\gamma_{S_1} = min\{\gamma_{S_2R}, \gamma_{RS_1}\}, \gamma_{S_2} = min\{\gamma_{S_1R}, \gamma_{RS_2}\}. \tag{12}$$

The estimated channel gains $g_{S_iR} = \left|\hat{h}_{S_iR}\right|^2$ and $g_{RS_i} = \left|\hat{h}_{RS_i}\right|^2$ are exponentially distributed RVs with probability density function (PDF) and cumulative distribution function (CDF):

$$f_{g_{S_iR}}(z) = \frac{1}{\lambda_i} e^{-z/\lambda_i}, f_{g_{RS_i}}(x) = \frac{1}{\lambda_{i+2}} e^{-z/\lambda_{i+2}}, \tag{13}$$

$$F_{g_{S_iR}}(z) = 1 - e^{-z/\lambda_i}, F_{g_{RS_i}}(z) = 1 - e^{-z/\lambda_{i+2}}, \tag{14}$$

where $\lambda_i = d_i^{-\beta_i}, \lambda_{i+2} = d_{i+2}^{-\beta_{i+2}}$ and β_i, β_{i+2} are path-loss exponents.

3 OUTAGE PROBABILITY ANALYSIS

In this section, we consider the outage probabilities of the sources. The outage at S_i happens when the equivalent SNR exceeds the threshold$\gamma_{S_i} < \gamma_{th}$, where$\gamma_{th} = 2^R - 1$, R(bits/s/Hz) is the threshold data rate [11]. Specifically, the outage probability of the S_1 in model present as:

$$OP_{S_1} = Pr[\gamma_1 \le \gamma_{th}] = Pr\left[\gamma_{S_2R} \le \gamma_{th}\right] + Pr\left[\gamma_{R_{S_1}} \le \gamma_{th}, \gamma_{S_2R} > \gamma_{th}\right]. \tag{15}$$

Substituting (5) and (11) into (15), we obtain [6]:

$$OP_{S_1} = \Pr\left[\frac{(1-\rho)\left|h_{S_2R}\right|^2 P}{N_0} \le \gamma_{th}\right] + \Pr\left[\frac{\left|h_{RS_1}\right|^2}{N_0}\varsigma \le \frac{\gamma_{th}}{P_{th}}, \frac{(1-\rho)\left|h_{S_2R}\right|^2 P}{N_0} > \gamma_{th}\right] + \tag{16}$$

$$+ \Pr\left[P\left(\left|h_{S_2R}\right|^2 + \left|h_{S_1R}\right|^2\right)\frac{\left|h_{RS_1}\right|^2}{N_0}\varsigma \le \gamma_{th}, \frac{\left|h_{RS_1}\right|^2}{N_0}\varsigma > \frac{\gamma_{th}}{P_{th}}, \frac{(1-\rho)\left|h_{S_2R}\right|^2 P}{N_0} > \gamma_{th}\right]$$

The formula (16), can rewrite as:

$$OP_{S_1} = \Pr\left[g_{S_2R} \le k_1\right] + \Pr\left[g_{RS_1} \le k_3, g_{S_2R} > k_1\right] + \\ + \Pr\left[\left(g_{S_2R} + g_{S_1R}\right) g_{RS_1} \le k_2, g_{RS_1} > k_3, g_{S_2R} > k_1\right] \quad (17)$$

$$= \int_0^{k_1} f_{g_{S_2R}}(y)dy + \int_{k1}^{\infty} f_{g_{S_2R}}(y)dy \int_0^{k_3} f_{g_{RS_1}}(x)dx + \int_{k3}^{\frac{k2}{k1}} f_{g_{RS_1}}(x) \int_{k1}^{\frac{k2}{x}} f_{g_{S_2R}}(y) \int_0^{\frac{k2}{x}-y} f_{g_{S_1R}}(z)dzdydx$$

where $\gamma_0 = \frac{P}{N_0}, k_1 = \frac{\gamma_{th}}{(1-\rho)\gamma_0}, k_2 = \frac{\gamma_{th}}{\varsigma\gamma_0}, k_3 = \frac{\gamma_{th}N_0}{\varsigma P_{th}}$.

After applicating Taylor's series [5]: $exp\left(-\frac{x}{\lambda_{p+2}}\right) = \sum_{k=0}^{\infty} \frac{(-x)^k}{k!\lambda_{p+2}^k}$ and calculating the complex integrations, we have the OP as the two case follows:

- In case of $\lambda_1 = \lambda_2 = \lambda$:

$$OP_{S_1}\big|_{\lambda_1=\lambda_2=\lambda} = 1 - e^{-(k_1/\lambda + k_3/\lambda_3)} + e^{-k_1/\lambda}\left(e^{-k_3/\lambda_3} - e^{-k2/(k_1\lambda_3)}\right) + \\ + (\frac{k_1}{\lambda} - 1)\sum_{k=0}^{\infty} \frac{(-1)^k}{k!}\left(\frac{k_2}{\lambda\lambda_3}\right)^{k+1}\left[\Gamma\left(-k-1, \frac{k_1}{\lambda}\right) - \Gamma\left(-k-1, \frac{k_2}{k_3\lambda}\right)\right] \\ + \sum_{k=0}^{\infty} \frac{(-1)^{k-1}}{k!}\left(\frac{k_2}{\lambda\lambda_3}\right)^{k+1}\left[\Gamma\left(-k, \frac{k_1}{\lambda}\right) - \Gamma\left(-k, \frac{k_2}{k_3\lambda}\right)\right]. \quad (18)$$

- In case of $\lambda_1 \neq \lambda_2$:

$$OP_{S_1}\big|_{\lambda_1\neq\lambda_2} = 1 - e^{-(k_1/\lambda_2 + k_3/\lambda_3)} + e^{-k_1/\lambda_2}\left(e^{-k_3/\lambda_3} - e^{-k_2/(k_1\lambda_3)}\right) + \\ + \frac{\lambda_2}{(\lambda_1 - \lambda_2)}\sum_{k=0}^{\infty} \frac{(-1)^k}{k!}\left(\frac{k_2}{\lambda_2\lambda_3}\right)^{k+1}\left[\Gamma\left(-k-1, \frac{k_1}{\lambda_2}\right) - \Gamma\left(-k-1, \frac{k_2}{k_3\lambda_2}\right)\right] \\ - \frac{\lambda_1 e^{-(1/\lambda_2 - 1/\lambda_1)k_1}}{(\lambda_1 - \lambda_2)}\sum_{k=0}^{\infty} \frac{(-1)^k}{k!}\left(\frac{k_2}{\lambda_1\lambda_3}\right)^{k+1}\left[\Gamma\left(-k-1, \frac{k_1}{\lambda_1}\right) - \Gamma\left(-k-1, \frac{k_2}{k_3\lambda_1}\right)\right]. \quad (19)$$

Likewise, the OP of the S_2 in the model is presented as (20) and (21):

$$OP_{S_2}\big|_{\lambda_1=\lambda_2=\lambda} = 1 - e^{-(k_1/\lambda + k_3/\lambda_4)} + e^{-k_1/\lambda}\left(e^{-k_3/\lambda_4} - e^{-k2/(k_1\lambda_4)}\right) + \\ + (\frac{k_1}{\lambda} - 1)\sum_{k=0}^{\infty} \frac{(-1)^k}{k!}\left(\frac{k_2}{\lambda\lambda_4}\right)^{k+1}\left[\Gamma\left(-k-1, \frac{k_1}{\lambda}\right) - \Gamma\left(-k-1, \frac{k_2}{k_3\lambda}\right)\right] \\ + \sum_{k=0}^{\infty} \frac{(-1)^{k-1}}{k!}\left(\frac{k_2}{\lambda\lambda_4}\right)^{k+1}\left[\Gamma\left(-k, \frac{k_1}{\lambda}\right) - \Gamma\left(-k, \frac{k_2}{k_3\lambda}\right)\right]. \quad (20)$$

$$OP_{S_2}\big|_{\lambda_1\neq\lambda_2} = 1 - e^{-(k_1/\lambda_1 + k_3/\lambda_4)} + e^{-k_1/\lambda_1}\left(e^{-k_3/\lambda_4} - e^{-k_2/(k_1\lambda_4)}\right) + \\ + \frac{\lambda_1}{(\lambda_2 - \lambda_1)}\sum_{k=0}^{\infty} \frac{(-1)^k}{k!}\left(\frac{k_2}{\lambda_1\lambda_4}\right)^{k+1}\left[\Gamma\left(-k-1, \frac{k_1}{\lambda_1}\right) - \Gamma\left(-k-1, \frac{k_2}{k_3\lambda_1}\right)\right] \\ - \frac{\lambda_2 e^{-(1/\lambda_1 - 1/\lambda_2)k_1}}{(\lambda_2 - \lambda_1)}\sum_{k=0}^{\infty} \frac{(-1)^k}{k!}\left(\frac{k_2}{\lambda_2\lambda_4}\right)^{k+1}\left[\Gamma\left(-k-1, \frac{k_1}{\lambda_2}\right) - \Gamma\left(-k-1, \frac{k_2}{k_3\lambda_2}\right)\right] \quad (21)$$

When $P_{th} \to \infty$ we have linear energy harvesting at relay nearly like the paper [5]

4 RESULTS AND DISCUSSIONS

In the section, we analyze and evaluate the outage probabilities of NEH-TW-DNC scheme. To validate the accuracy of the exact theory expressions, we use Monte Carlo simulations (markers point to simulated results in all figures). At all our analyzations and assessments, we set the target SNR as $\gamma_t = 1$, the path-loss exponent as$\beta_k = 3$ and the variance of noise at the nodes as $N_0 = 1$. Besides, we use the first 20 terms for all the infinite sum formulas.

Figure 2 illustrations the OPs of the S_1 aid S_2 of NEH-TW-DNC system versus $P_s/N_0(dB)$ when$P_{th} \in \{20,40,\infty\}dB$, $\rho = 0.5$, $\eta = 0.8$ $\alpha = 1/3$, $d_1 = 0.3$, $d_2 = 1 - d_1$. Due to the symmetry of the model, when the relay is equidistant from the two sources, the OP of the two sources are equal. With these values of d_1, d_2, the relay is nearer S_1 than S_2. Observing Figure 2, when P_s is less than P_{th}, the OP_{S_2} has better

than OP_{S_1}, conversely when P_s is more than P_{th}, the OP_{S_1} has better than OP_{S_2}. It is worth noting that when $P_{th} = P_s$, the OP of the two sources are equal. When $P_s/N_0(dB)$ is in the small value region, the OP of the two sources decrease as P_s/N_0 increases, and the OP of each source is roughly equal for all values of P_{th}. Besides, when P_s/N_0reaches a large enough value, the OP will reach a threshold value (OP_{th}), that means P_s/N_0increases but OP no longer increase. This OP_{th} depends on the value of P_{th}, when the P_{th} is larger the OP_{th} is smaller. Moreover, the OP curves of the nonlinear EH model will approach the OP of the linear EH model as P_{th} approaches infinity.

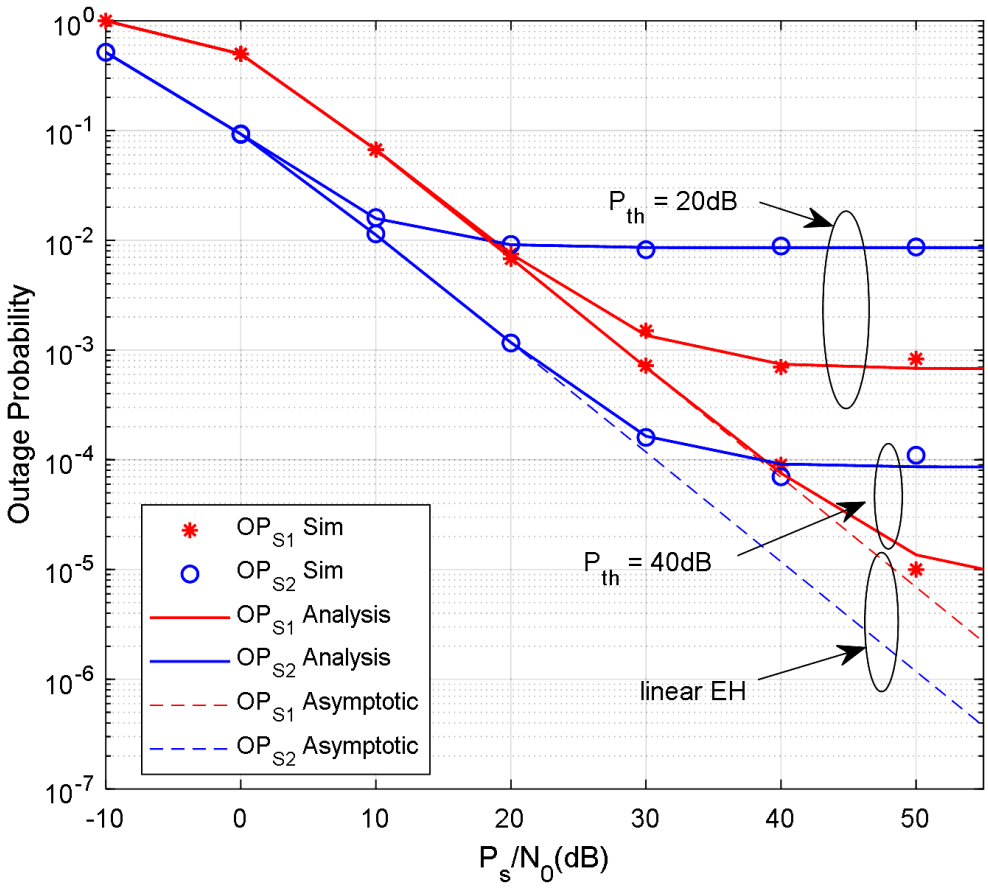


Figure 2. The outage probabilities of the S_1 and S_2 versus $P_s/N_0(dB)$ with three values$P_{th} \in \{20,40,\infty\}(dB)$

Fig.3 shows the outage probabilities of the source S_1 of NEH-TW-DNC system as a function of the saturation threshold power $P_{th}(dB)$ when $P_s/N_0 \in \{20,40\}dB$, $\rho = 0.5$, $\eta = 0.8$ $\alpha = 1/3$, $d_1 \in \{0.3,0.5,0.7\}$, $d_2 = 1 - d_1$. Due to symmetry of the system model, the OP_{S_1} is equal the OP_{S_2}if the equivalent parameters of the two sources are the same, so in fig.3 we just draw the curves of OP_{S_1}. We can observe that the OP_{S_1} is better when P_{th} and P_s/N_0increases and the OP_{S_1} achieves the saturation values when P_{th} increases to a high value, and this value is change according to the values of P_s/N_0, and d_1. Besides, when $P_{th}$is smaller than P_s/N_0 the smaller d_1 is the better OP_{S_1} and when P_{th} is bigger than P_s/N_0 the higher $d_1$is the better OP_{S_1}.

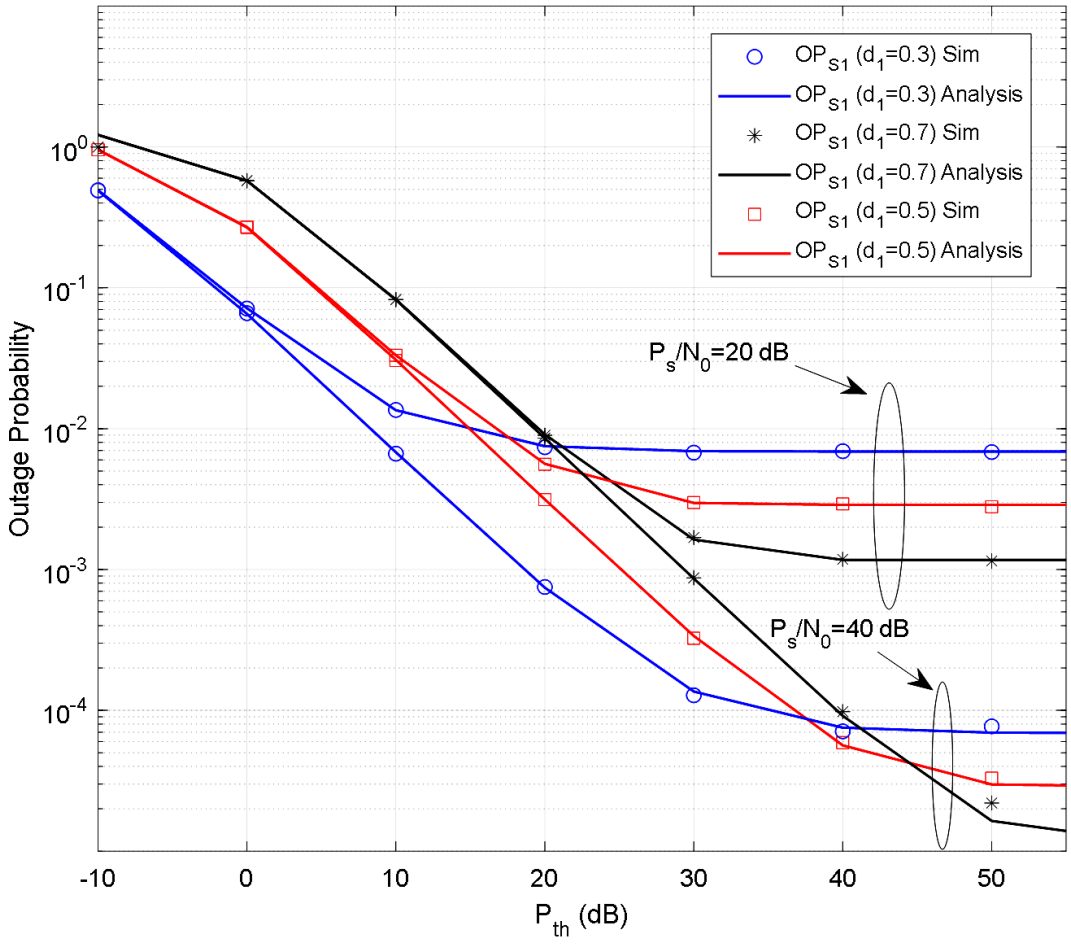


Figure 3. The outage probabilities of the S_1 versus $P_{th}(dB)$ with two values of P_s/N_0 and three values of d_1

Fig. 4 shows the outage probabilities of the source S_1 of NEH-TW-DNC system as a function of the power allocation coefficient ρ when $P_s/N_0 = 30(dB)$, $P_{th} \in \{20,30,40\}dB$, $\eta = 0.8$ $\alpha = 1/3$, $d_1 = 0.5$, $d_2 = 1 - d_1$. It is observed that the OP_{S_1} decreases as the P_{th} increases. Moreover, for each set of parameters, there is always a value of ρ so that the source has the smallest outage probability.
Finally, the theory analyses and the Monte Carlo simulations coincide for all figures.

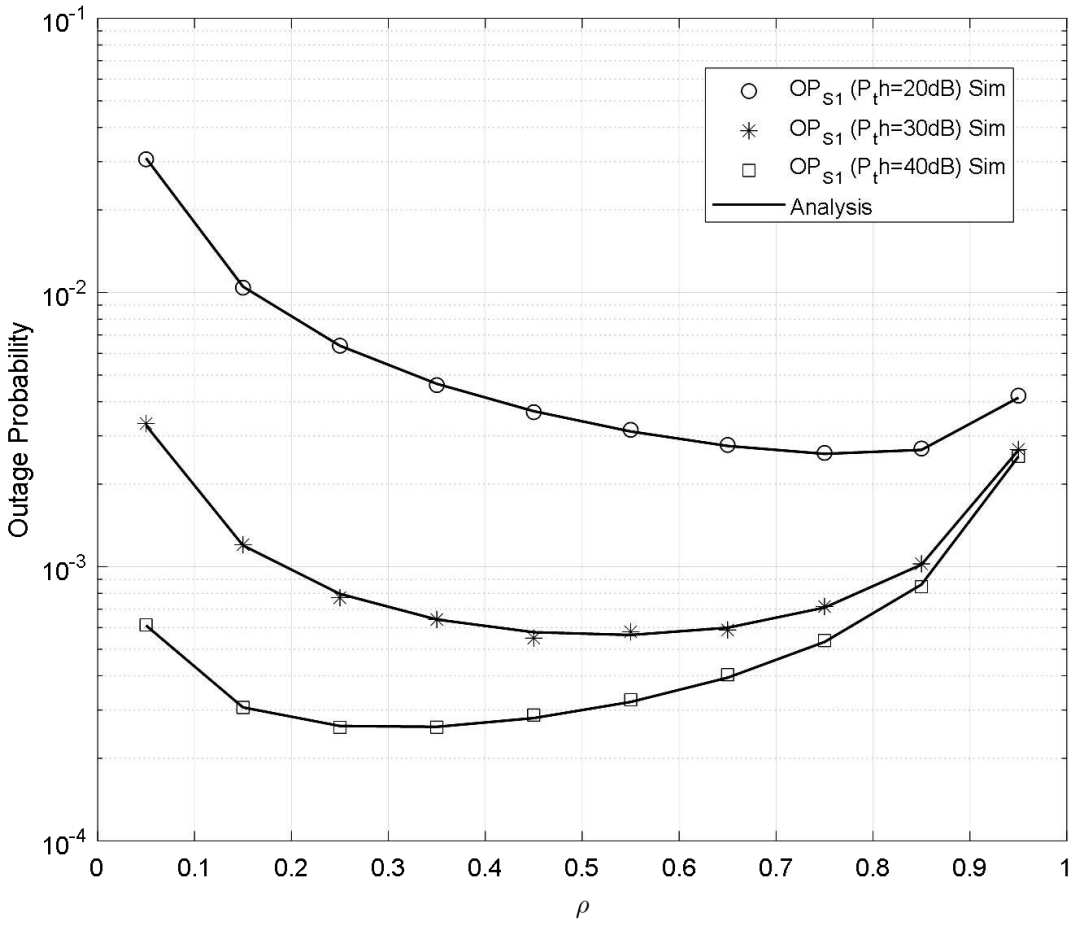


Figure 4. The outage probabilities of the S_1 versus ρ with three values of P_{th} and d_1

5 CONCLUSIONS

In this article, we have derived the exact sources outage probabilities of a two-way cooperation system. And the relay has a non-linear equip to collect the energy from the received radio frequency signals in two first time slots. Then it converts this energy into the broadcast power in the third time slot. The outage probabilities are investigated according to the saturation threshold power, the power allocation coefficient, and the relative position of the relay and two sources. Numerical results show that when the saturation threshold power of a non-linear equip to harvest the energy at relay increases, the outage probabilities decrease. And the outage probabilities achieve the saturation value when P_s/N_0 and P_{th} are enough large. Furthermore, we can define the optimal the power allocation coefficient where the outage probability is minutest. Finally, the theory analyses and the Monte Carlo simulations coincide for all figures. Our future work may consider more techniques and reality conditions in the NEH two-way system to enhance the system performance.

REFERENCES

[1] A. A. Nasir, X. Zhou, S. Durrani, and R. A. Kennedy, "Relaying protocols for wireless energy harvesting and information processing," *IEEE Transactions on Wireless Communications,* vol. 12, no. 7, pp. 3622-3636, 2013.

[2]T.-T. T. Dao and P. N. Son, "Uplink non-orthogonal multiple access protocol in two-way relaying networks: realistic operation and performance analysis," in *2020 7th NAFOSTED Conference on Information and Computer Science (NICS)*, 2020, pp. 399-404: IEEE.

[3] T.-T. T. Dao and P. N. Son, "Cancel-Decode-Encode Processing on Two-Way Cooperative NOMA Schemes in Realistic Conditions," *Wireless Communications and Mobile Computing*, vol. 2021, 2021.

[4] H. V. Toan and T. M. Hoang, "Outage Probability Analysis of Decode-and-Forward Two-Way Relaying System with Energy Harvesting Relay," *Wireless Communications and Mobile Computing*, vol. 2020, 2020.

[5] T. N. Nguyen, P. T. Tran, and M. Voznak, "Wireless energy harvesting meets receiver diversity: A successful approach for two-way half-duplex relay networks over block Rayleigh fading channel," *Computer Networks*, p. 107176, 2020.

[6] A. Cvetkovic, V. Blagojevic, and P. Ivaniš, "Performance analysis of nonlinear energy-Harvesting DF relay system in interference-limited Nakagami-m fading environment," *ETRI journal*, vol. 39, no. 6, pp. 803-812, 2017.

[7] D. Kumar, P. K. Singya, and V. Bhatia, "Performance Analysis of Hybrid Two-Way Relay Network with NLPA and Hardware Impairments," in *2021 IEEE Wireless Communications and Networking Conference (WCNC)*, 2021, pp. 1-6: IEEE.

[8] L. Shi, W. Cheng, Y. Ye, H. Zhang, and R. Q. Hu, "Heterogeneous power-splitting based two-way DF relaying with non-linear energy harvesting," in *2018 IEEE Global Communications Conference (GLOBECOM)*, 2018, pp. 1-7: IEEE.

[9] I. S. G. a. I. M. Ryzhik, *Table of integrals, series, and products*, Seventh ed. Elsevier, 2007.

[10] P. N. Son and T. T. Duy, "A new approach for two-way relaying networks: improving performance by successive interference cancellation, digital network coding and opportunistic relay selection," *Wireless Networks*, pp. 1-15, 2019.

[11] T.-T. T. Dao and P. N. Son, "Multi-constraint two-way underlay cognitive network using reconfigurable intelligent surface," *Wireless Networks*, p. 14, 2022/04/06 2022.

ICATSD2F.112

OPTIMAL TRACKING CONTROL FOR ROBOT MANIPULATORS WITH INPUT CONSTRAINT BASED REINFORCEMENT LEARNING

NGUYEN DUC DIEN[1], NGUYEN TAN LUY[2], LAI KHAC LAI[3], TRAN THANH HAI[2]

[1]*University of Economics – Technology for Industries*

[2]*Industrial University of Ho Chi Minh City*

[3]*Thai Nguyen University of Technology*

nddien@uneti.edu.vn, nguyentanluy@iuh.edu.vn, laikhaclai@tnu.edu.vn, tranthanhhai@iuh.edu.vn

Abstract. An important problem in motion control of robot manipulators is tracking the reference joint angle position trajectory. The controller is designed to balance both tracking performance and the control energy from the torque at the joints of the robot manipulators, which means it is necessary to minimize a predefined multi-objective cost function. In addition, due to the torque limitation of actuators in robot manipulators, control inputs are saturated. This paper introduces an optimal tracking controller for robot manipulators with saturation control torque based on a reinforcement learning (RL) method based on a neural network (NN). A cost function is designed, and then a constrained optimal tracking control is proposed, which ensures that the NN weights and the tracking error can achieve convergence in the optimal sense. The constrained Hamilton-Jacobi-Bellman (HJB) equation is built, which is solved by the online RL algorithm using only a single NN instead of two. Then the optimal control law with the input constraint is determined. The technique of experience replay is used to relax a persistent citation condition when updating the critic NN weights to global optimal values. The control performance is given in the comparative simulation results, showing the effectiveness of the proposed method.

Keywords. Robot manipulators, reinforcement learning, optimal control, input constraints.

1 INTRODUCTION

Robot manipulators have brought excellent efficiency in manufacturing, healthcare, and services. Controller designs to control the robot manipulators for proximity tracking according to the reference signal have continuously received the attention of researchers [1, 2]. The traditional method which builds a controller that satisfies the above requirements is the PD gravity controller [1]. During the working process, the robot manipulators are affected by external noise, load weight change, nonlinear friction, unexpected changes in model parameters (due to equipment wear, parameter deviation, etc.) technology of the system after a long working time. Thus, the PD gravity controller does not guarantee the control performance for the robot manipulators during real-time working. Advanced control options have been studied and applied successfully, such as adaptive control, sliding control, and backstepping control [3-11]. In [3], they used a sliding controller combined with a system identifier and disturbance observer. In [4], a widely and stably adaptive sliding mode has been used. In [5], a terminal sliding mode control has been used. Furthermore, the fast terminal slide mode controller has been based on third-order sliding introduced in [6] and the fixed time slide controller introduced in [7]. Chi et al. [9] used a backstepping controller, and a terminal slide controller combined with the backstepping controller was used in [10] [11]. Intelligent controllers based on fuzzy control, neural networks (NN) control combined with sliding controllers, or backstepping controllers have also been applied in tracking controllers for robot manipulators [12] [13] [14]. Zhou et al. [12] proposed an adaptive fuzzy backstepping controller, Van and Ge [13] used an adaptive fuzzy sliding controller, and Jie et al. [14] proposed a terminal slide mode controller combined with RBF NN controller. In practice, with the actuators of joints in robot manipulators, the input that drives the motors of the joints is always torque-bounded, which reduces control performance or instability. The adaptive fuzzy controller [15] and the NN controller [16] [17] were designed and applied control to the robot manipulators with input constraints and dead zones.

The fuzzy or neural controllers combined with modern controllers have effectively ensured the tracking quality for robot manipulators with input constraints. However, the above articles did not consider the optimal control problem. The controller is designed to balance both the quality of tracking performance and the control energy from the torque at the joints of the robot manipulators.

Mathematically, the optimal control problem is solved when solving the solution of the HJB equation. The nonlinear HJB equation is difficult to solve. RL is a helpful method to approximate online HJB equation solutions [18, 19]. RL was used to control the tracking for robot manipulators [20-22]. In [20], Modares et al. used an offline policy algorithm, which leads to a slow learning rate. Kamalapurkar et al. [21] used an online learning algorithm and analyzed convergence and stability, but the algorithm used 2 NN.

For the saturated optimal tracking control problem, Luy et al. [23-25] have designed an optimal tracking controller using RL for a nonlinear system and successfully applied it to a mobile robot [24], a PMSM system [25] with input constraints. Recently, Zhao et al. [26] have used RL for the problem of saturated optimal tracking control for robot manipulators in discrete time. To our best knowledge, there are currently no studies for the optimal tracking problem for robot manipulators with saturated inputs in continuous time.

The main contributions of this article are summarized as follows:

- The robot manipulator dynamics are transformed into a strict-feedback nonlinear system with input constraints. The feedforward control inputs are designed to transform a position tracking control problem into an optimal control problem.
- The constrained HJB equation is built, which is solved by the online RL algorithm using only a single NN instead of two NNs, and the optimal control law is also determined. Furthermore, the experience replay is used to relax a persistent citation condition.

Next, the paper is organized as follows. In Section 2, the robot manipulator dynamics and the design problem of feedforward control inputs are presented. Section 3 offers a design of the optimal control law with the input constraint. The simulation results are shown in Section 4. The conclusion of the article is given in Section 5.

2 ROBOT MANIPULATORS AND FEEDFORWARD CONTROL

2.1 Robot manipulators nonlinear

Consider a model of a robot manipulator in the form of Euler – Lagrange system with full actuators [1]:

$$M(q)\ddot{q} + C(q,\dot{q})\dot{q} + G(q) = \tau \tag{1}$$

where $q \in \mathbb{R}^{n\times 1}$ is the angular position vector, $\dot{q} \in \mathbb{R}^{n\times 1}$ is the angular velocity vector, $\ddot{q} \in \mathbb{R}^{n\times 1}$ is the angular acceleration vector, $M(q) \in \mathbb{R}^{n\times n}$ is the inertia matrices, $C(q,\dot{q}) \in \mathbb{R}^{n\times n}$is the Coriolis-centripetal matrices, $G(q) \in \mathbb{R}^{n}$ is the dynamic friction, $\tau \in \mathbb{R}^{n\times 1}$is the control input vector of joints torque.

Dynamics (1) has some properties [1]: $M(q)$ is a symmetric matrix, positive definite, and is bounded $m_1 I \leq M(q) \leq m_2 I$ ($m_1, m_2$are positive constants). The matrix $C(q,\dot{q})$ and $G(q)$ are bounded. $\dot{M}(q) - 2C(q,\dot{q})$ is the skew-symmetric matrix.

We transform dynamics (1) into a strict-feedback nonlinear system with saturation torques as follows:

$$\begin{cases} \dot{q} = f_q(q) + g_q(q)v \\ \dot{v} = f_v(q,v) + g_v(q,v)\tau \end{cases} \tag{2}$$

where $f_q(q) = 0_{nx1}, g_q(q) = I_n$, $f_v(q,v) = -M^{-1}(Cv + G) \in \mathbb{R}^{n\times 1}$, $g_v(q,v) = M^{-1} \in \mathbb{R}^{n\times n}$, $\tau = [\tau_1, \tau_2, \ldots, \tau_n]^T$ are the control inputs limited by $|\tau_i| \leq \lambda$ $(i = 1 \div n)$.

2.2 Feedforward Control

In this section, we design the feedforward control inputs by applying the backstepping method. Then, we transform (2) into an optimal control problem [19].

Step 1: Define $e_q = q - q_d$, the position tracking error dynamics are calculated as:

$$\dot{e}_q = -\dot{q}_d + g_q(q)v_d + g_q(q)(v - v_d) = f_q(e_q) + g_g(q)v_d + g_q(q)e_v \tag{3}$$

wherev_d is virtual control input vector and designed as $v_d = v_d^* + v_{da}$, with v_d^* is the virtual optimal control input vector, and $v_{da}$is designed as follows:

$$v_{da} = g_q^{-1}(q)\big(\Lambda_1 e_v + \dot{q}_d + f_q(e_q)\big) \tag{4}$$

Substituting (4) into (3), one obtains

$$\begin{aligned}\dot{e}_q &= -\dot{q}_d + g_q(q)v_d + g_q(q)(v - v_d)\\ &= f_q(e_q) + g_g(q)v_d^* + \Lambda_1 e_v + g_q(q)e_v\\ &= \bar{f}_q(e_q, e_v) + g_g(q)v_d^* + g_q(q)e_v\end{aligned} \tag{5}$$

where $\bar{f}_q(e_q, e_v) = f_q(e_q) + \Lambda_1 e_v$.

Step 2: Defining $e_v = \dot{q} - v_d$, we have

$$\dot{e}_v = -\dot{v}_d + f_v(q, v) + g_v(q, v)\tau \tag{6}$$

where τ is designed as $\tau = \tau^* + \tau_a$ with τ^* is the optimal control input vector, and τ_a is designed as

$$\tau_a = g_v^{-1}(q,v)\left[\, f_v(e_q, e_v) + \dot{v}_d - f_v(q,v) - g_q^T(q)e_q - \Lambda_2 e_v \,\right] \tag{7}$$

Substituting (7) into (6), one obtains

$$\dot{e}_v = \bar{f}_v(e_q, e_v) + g_v(q,v)\tau^* - g_q^T(q)e_q \tag{8}$$

where $\bar{f}_v(e_q, e_v) = f_v(e_q, e_v) - \Lambda_2 e_v$.

Lemma 1: Consider dynamic tracking in (3) and (5), if u^* is designed to stabilize the following system

$$\dot{e} = \bar{f} + g u^* \tag{9}$$

where $e = [e_q^T, e_v^T]^T \in \mathbb{R}^{2n\times 1}$, $\bar{f} = [\bar{f}_q^T(e_q, e_v), \bar{f}_v^T(e_q, e_v)]^T \in \mathbb{R}^{2n\times 2n}$, $u^* = [v_d^{*T}, \tau^{*T}] \in U \in \mathbb{R}^{2n\times 1}$, $u_a = [v_{da}^T, \tau_a^T]^T \in \mathbb{R}^{2n\times 1}$, $g = diag\big[g_q(q), g_v(q,v)\big] \in \mathbb{R}^{2n\times 2n}$.

Then, the position tracking control problem for (2) is equivalent to an optimal tracking control problem for (9).

Proof. A Lyapunov function is chosen for system (2) as

$$J_1 = \frac{1}{2}e_q^T e_q + \frac{1}{2}e_v^T e_v \tag{10}$$

and taking derivative (10) along trajectories of (5) and (8), we get the following:

$$\begin{aligned}\dot{J}_1 &= e_q^T \bar{f}_q(e_q) + e_q^T g_q(q)v_d^* + e_q^T g_q(q)e_v + e_v^T \bar{f}_v(e_q, e_v) + e_v^T g_v(q,v)\tau^* - e_v^T g_q^T(q)e_q\\ &= e_q^T\left(\bar{f}_q(e_q) + g_q(q)v_d^*\right) + e_v^T\left(\bar{f}_v(e_q, e_v) + g_v(q,v)\tau^*\right)\\ &= e^T\left(\bar{f} + g u^*\right)\end{aligned} \tag{11}$$

For dynamics (9), choose a Lyapunov function as

$$J_2 = \frac{1}{2}e^T e \tag{12}$$

Taking derivative (12), one obtains

$$\dot{J}_2 = \frac{1}{2}e^T\left(\bar{f} + g u^*\right) \tag{13}$$

Comparing (11) with (13), we can be seen that if $\dot{J}_1 < 0$ implies $\dot{J}_2 < 0$. In other words, the stabilities of closed-loop dynamics (2) and closed-loop dynamics (9) are equivalents.

The Proof is completed.

If the feedforward control inputs are designed with the following constraint [25]:

$$\|\tau_a\| \le \lambda - \rho\, tanh(1) \tag{14}$$

and the saturated optimal feedback control inputs $\|\tau^*\| \le \rho$, then the control inputs will be constrained by $\|\tau\| \le \lambda$.

3 OPTIMAL TRACKING CONTROL

After converting the tracking control problem of the system (2) into the optimal tracking control problem of the system (9). In this section, we design the optimal tracking controller with input constraints of the system (9). First, a cost function is designed, and constrained optimal tracking control is proposed. Then the constrained HJB equation is built and solved by the online RL algorithm using only one NN.

We denote the control input of the system (9) as u and the optimal control input of the system (9) as u^*.

The cost function for system (9) is defined as

$$V(e,u) = \int_t^{\infty} r(e,u)\,dt \tag{15}$$

with $r(e,u) = e^T Qe + U(u)$, where Q is a positive definite matrix, $U(u)$ is non-negative, and u approximates u^*at each given time.

We follow the approach in [27] to define the energy cost function, i.e. $U(u) = 2\int_0^u \Omega^{-T}(s)Rds$, where $\Omega(.)$is a continuous one-to-one, bounded, the real-analytic integrable function of class C^{π}, $\pi > 1$, used to map $\mathbb{R}$ onto the interval $(-\rho, \rho)$. We choose $U(u)$ of the following form:

$$U(u) = 2\rho\int_0^u tanh^{-T}(s/\rho)Rds \tag{16}$$

whereρ is the saturating bound of u, R is a positive definite diagonal matrix.

The optimal cost function is derived from the following law:

$$V^*(e,u) = \min_u\big(V(e,u)\big) \tag{17}$$

Define $V_e = \frac{\partial V}{\partial e}$ and the Hamilton function for system (9) with a cost function (15) as follows:

$$H(e,u,V_e) = r(e(t),u(t)) + V_e^T\big(\bar{f} + gu\big) \tag{18}$$

The saturated optimal control law $u^*$is determined using the stationary condition of (17). We have

$$u^* = arg\min_u H(e,u,V_e) \tag{19}$$

$$\Rightarrow u^* = -\rho\, tanh\left(\frac{1}{2\rho}R^{-1}g^T V_e^*\right) \tag{20}$$

Substituting (20) into (18), we get the following HJB equation:

$$H^*(e,u,V_e^*) = e^T Qe + U(u^*) + V_e^{*T}\big(\bar{f} + gu^*\big) = 0 \tag{21}$$

Using integral by parts (16), we get

$$U(u^*) = 2\rho\int_0^{u^*} tanh^{-T}(s/\rho)Rds = 2\rho\, tanh^{-T}\left(\frac{s}{\rho}\right)Rs\Big|_0^{u^*} - 2\rho\bar{R}\int_0^{u^*} sd(tanh^{-T}(s/\rho)) \tag{22}$$

with $\bar{R}$is a column vector of the diagonal elements of R.

$$\rightarrow U(u^*) = 2\rho\, tanh^{-T}\left(\frac{u^*}{\rho}\right)Ru^* + U(u^*)_2 \tag{23}$$

where

$$U(u^*)_2 = -2\rho\bar{R}\int_0^{u^*} sd(tanh^{-T}(s/\rho)) \tag{24}$$

Define $v = tanh^{-T}(s/\rho) \rightarrow s = \rho\, tanh(v)$

$$\rightarrow U(u^*)_2 = -2\rho^2\bar{R}\int_0^{tanh^{-T}(u^*/s)} tanh(v)dv \tag{25}$$

Define $z = tanh(v) \rightarrow dz = (\bar{1} - tanh^2(v))dv \rightarrow dv = \frac{dz}{\bar{1}-tanh^2(v)} = \frac{dz}{\bar{1}-z^2}$ (26)

where $\bar{1} = [1,1,\cdots,1]^T$.

Substituting (26) into (25), we have

$$U(u^*)_2 = -2\rho^2\bar{R}\int_0^{u^*/\rho}\frac{z}{\bar{1}-z^2}dz = -2\rho^2\bar{R}\int_0^{u^*/\rho}\frac{-1/2}{\bar{1}-z^2}d(1-z^2) \tag{27}$$

$$\rightarrow U(u^*)_2 = \rho^2\bar{R}\, ln(\bar{1} - tanh^2(v))\Big|_0^{tanh^{-T}\left(\frac{u^*}{\rho}\right)} \tag{28}$$

$$\rightarrow U(u^*)_2 = \rho^2\bar{R}\left(ln\left(\bar{1} - \left(tanh\left(tanh^{-1}\left(\frac{u^*}{\rho}\right)\right)^2\right)\right) - ln(1)\right) = \rho^2\bar{R}\, ln\left(\bar{1} - \left(\frac{u^*}{\rho}\right)^2\right) \tag{29}$$

From (29) and (23), we have

$$U(u^*) = 2\rho\, tanh^{-T}\left(\frac{u^*}{\rho}\right)Ru^* + \rho^2\bar{R}\, ln\left(\bar{1} - \frac{u^{*2}}{\rho^2}\right) \tag{30}$$

To solve the HJB equation, V_e^* is approximated by a NN. As follows:

$$V^*(e) = W^T\phi(e) + \varepsilon(e) \tag{31}$$

where $W \in \mathbb{R}^N$is the NN weight vector, $\phi(e)$: $\mathbb{R}^n \rightarrow \mathbb{R}^N$is the NN activation function vector with $\phi(0) = 0$, N is the number of neurons in the hidden layer, $\varepsilon(e)$ is the NN approximation error.

Assumptions 1 [25]: One can select an entirely independent basis set $\phi(e)$such that $\|\phi(e)\| \le s_{\phi}$, $\|\phi_e\| = \|\partial\phi(e)/\partial e\| \le s_{\nabla\phi}$,$\|\varepsilon(e)\| \le s_{\varepsilon}$,$\|\varepsilon_e\| = \|\partial\varepsilon(e)/\partial e\| \le s_{\nabla\varepsilon}$, where$s_{\phi}$, $s_{\nabla\phi}$, s_{ε}, $s_{\nabla\varepsilon}$are positive constants.

Taking derivative (31), we have

$$V_e^* = W^T \frac{\partial \phi(e)}{\partial e} + \frac{\partial \varepsilon(e)}{\partial e} = W^T \phi_e + \varepsilon_e \tag{32}$$

Substituting (32) into (21), the Hamiltonian equation (21) becomes

$$H^*(e, u, W^T \phi_e) = e^T Q e + U(u^*) + W^T \phi_e \left(\bar{f} + g u^*\right) + \varepsilon_e \left(\bar{f} + g u^*\right) = 0 \tag{33}$$

Define $\varepsilon_H = -\varepsilon_e \left(\bar{f} + g u^*\right)$, (33) becomes

$$H^*(e, u, W^T \phi_e) = e^T Q e + U(u^*) + W^T \phi_e \left(\bar{f} + g u^*\right) - \varepsilon_H = 0 \tag{34}$$

The weight vector (31) is unknown, so $V^*(e)$ is approximated by

$$\hat{V}(e) = \widehat{W}^T \phi(e) \tag{35}$$

where $\widehat{W} \in R^N$ is the NN weight vector. The control law u^* (20) becomes

$$\hat{u} = -\rho \, tanh\left(\frac{1}{2\rho} R^{-1} g^T \phi_e^T \widehat{W}\right) \tag{36}$$

Substituting (35) into (21), the Hamiltonian equation (21) becomes

$$\hat{H}\left(e, \hat{u}, \widehat{W}^T \phi_e\right) = e^T Q e + U(\hat{u}) + \widehat{W}^T \phi_e \left(\bar{f} + g \hat{u}\right) = e_1 \tag{37}$$

where

$$e_1 = -\widetilde{W}^T \phi_e \left(\bar{f} + g \hat{u}\right) + \varepsilon_H \tag{38}$$

with $\widetilde{W} = W - \widehat{W}$.

The online parameter tuning law for the approximator is designed to minimize the squared error of $E = \frac{1}{2} e_1^T e_1$ to $\widehat{W} \to W$. Using the algorithm Normalized gradient descent and the experience replay (namely concurrent learning (CL)) [27, 28], the update law $\widehat{W}$ is defined as follows:

If $e^T \left(\bar{f} + g \hat{u}\right) < 0$, then

$$\dot{\widehat{W}} = -\alpha_1 \frac{\sigma}{(\sigma^T \sigma + 1)^2} \left(\sigma^T \widehat{W} + \mathrm{e}^T Q e + U(\hat{u})\right) - \alpha_1 \sum_{i=1}^{P} \frac{\sigma(t_i)}{(\sigma(t_i)^T \sigma(t_i) + 1)^2} \left(\sigma(t_i)^T \widehat{W} + \mathrm{e}(t_i)^T Q e(t_i) + U(\hat{u}(t_i))\right) \tag{39}$$

and if $e^T \left(\bar{f} + g \hat{u}\right) \geq 0$, then

$$\dot{\widehat{W}} = \left(-\alpha_1 \frac{\sigma}{(\sigma^T \sigma + 1)^2} \left(\sigma^T \widehat{W} + \mathrm{e}^T Q e + U(\hat{u})\right) - \alpha_1 \sum_{i=1}^{P} \frac{\sigma(t_i)}{(\sigma(t_i)^T \sigma(t_i) + 1)^2} \left(\sigma(t_i)^T \widehat{W} + \mathrm{e}(t_i)^T Q e(t_i) + U(\hat{u}(t_i))\right)\right) + \alpha_2 \frac{1}{2} \phi_e G e \tag{40}$$

where $\sigma = \phi_e \left(\bar{f} + g \hat{u}\right)$, $G = g R^{-1} g^T$, $\alpha_1 > 0$, $\alpha_2 > 0$.

In (40), the experience replay is used to relax a persistent citation condition PE when updating the critic NN weights. $\sigma(t_i)$, $r(t_i) = e^T Q e + U(\hat{u})$ are recorded and stored in $\{\sigma(t_i)\}_{i=1}^P$, $\{r(t_i)\}_{i=1}^P$, where $\{\sigma(t_i)\}_{i=1}^P$ must be linearly independent. That is $rank(\sigma(t_1), \sigma(t_2), \ldots, \sigma(t_P)) = N$. The number of Pselected is $P \geq N$ [27].

Theorem 1: Consider the dynamic (9), with the equation HJB (21), the critic NN (35), the saturated optimal control law (36), and the tuning law for the NN (39), (40). Then, the closed-loop system states and the approximation error are UUB, and the control law (36) converges to the near-optimal value.

Proof: The Proof is presented in the appendix.

4 SIMULATIONS

To verify the proposed controller's performance, we perform simulations with a two-link robot manipulator [20]. The convergence of NN weights with PE and CL is compared to show the effectiveness of the relaxation of PE condition.

The inertia and Coriolis-centripetal matrices are

$$M = \begin{bmatrix} h_1 + 2h_3c_2 & h_2 + h_3c_2 \\ h_2 + h_3c_2 & h_2 \end{bmatrix}$$

$$C = \begin{bmatrix} -h_3s_2\dot{q}_2 & -h_3s_2(\dot{q}_1 + \dot{q}_2) \\ h_3s_2\dot{q}_1 & 0 \end{bmatrix}$$

with $c_2 = cos(q_2)$, $s_2 = sin(q_2)$, $h_1 = 3.473$ kgm^2, $h_2 = 0.196$ kgm^2, and $h_3 = 0.242$ kgm^2. Moreover, the dynamic friction $G(q) = [8.45\,tanh(\dot{q}_1), 2.35\,tanh(\dot{q}_2)]^T$, τ is limited $\|\tau\| \leq 5\,kgm$.

The desired trajectory q_d are given as $q_d = [q_{d1}, q_{d2}]^T$, where $q_{d1} = sin(0.1t)$, $q_{d2} = sin(0.1t)$. The initial value is chosen $q(0) = [0.5, -0.5]^T$, $\dot{q}(0) = [0,0]^T$. The activation function is chosen $\phi(e) = \left[e_{q1}^2, e_{q1}e_{q2}, e_{q1}e_{v1}, e_{q1}e_{v2}, e_{q2}^2, e_{q2}e_{v1}, e_{q2}e_{v2}, e_{v1}^2, e_{v1}e_{v2}, e_{v2}^2\right]$. Select $Q = I \in R^{4\times4}$, $R = 1$, $\alpha_1 = 50$, $\alpha_2 = 0.01$, $\lambda = 5$, $\rho = 0.6$.

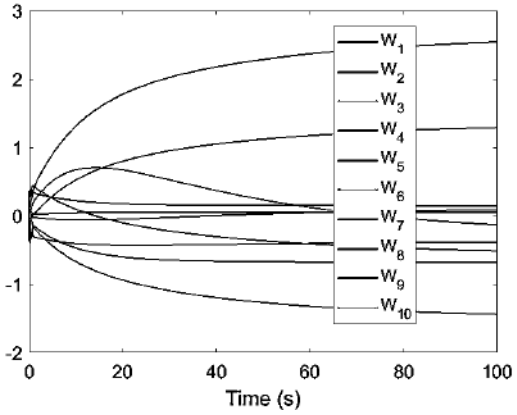


Fig.1. The weight values of the NN in the case of using CL

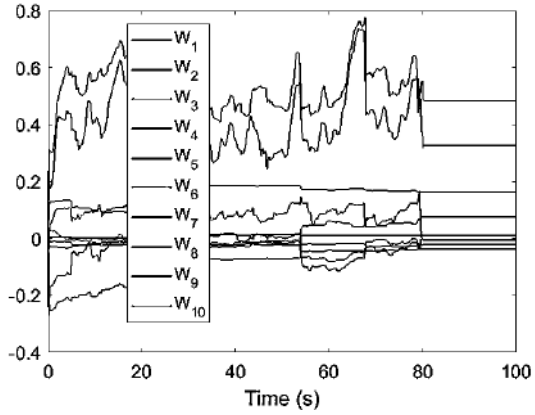


Fig.2. The weight values of the NN in the case of using only PE

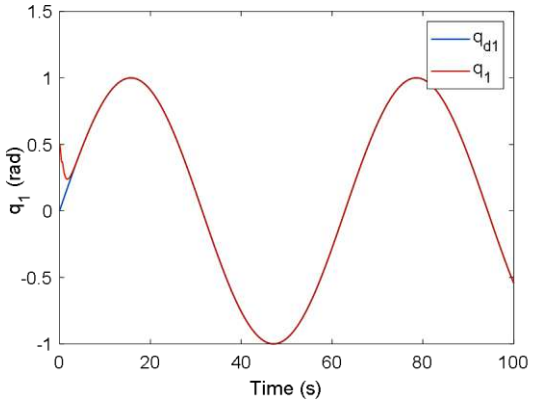


Fig. 3. Compare the position of the first joint with the reference position

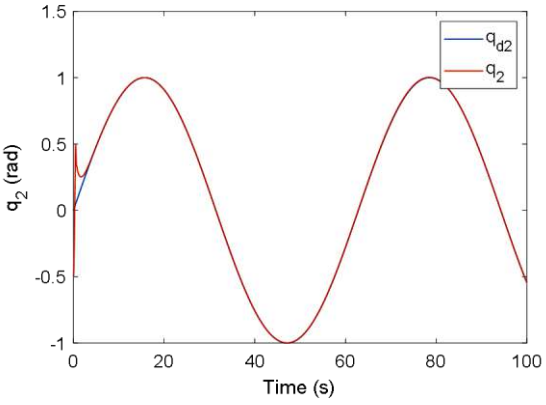


Fig. 4. Compare the position of the second joint with the reference position

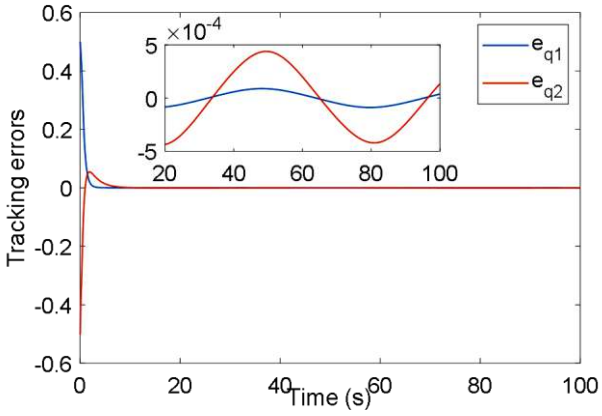


Fig. 5. The error tracking the position after learning

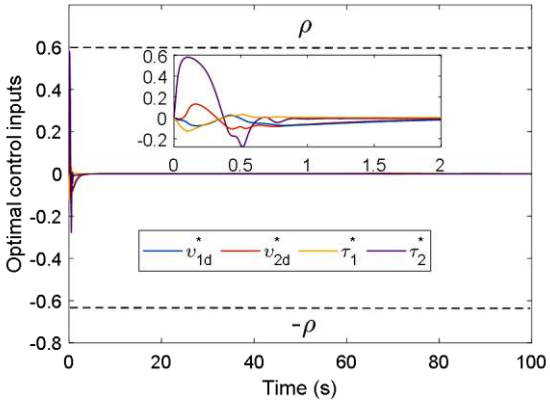


Fig. 6. The saturated optimal feedback control inputs

The feedforward controller parameter is selected as follows: $\Lambda_1 = diag(10,10)$, $\Lambda_2 = diag(10,10)$. Simulation time is 100s with sampling period T=0.01s.

For nonlinear systems, it is very difficult to obtain convergent weights using only PE. Thus, the PE is used during$0 \leq t \leq 0.4s$ and then the CL technique is used to relax a persistent citation condition PE. The convergent weight values of the NN have the results shown in Fig. 1 and have the following value: $W = [2.54, -0.6844, -1.4416, -0.3887, -0.5136, 1.2886, 0.1449, 0.0477, 0.0837, -0.1308]^T$. The case of using only PE applied for $0 \leq t \leq 80s$ has the result shown in Fig.2, and it shows that the weights get stuck in a local minimum.

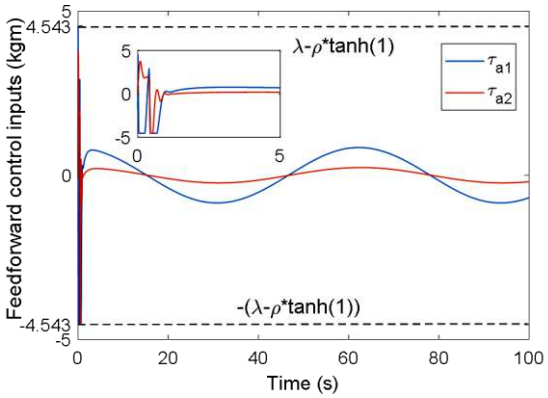


Fig. 7. The feedforward control inputs

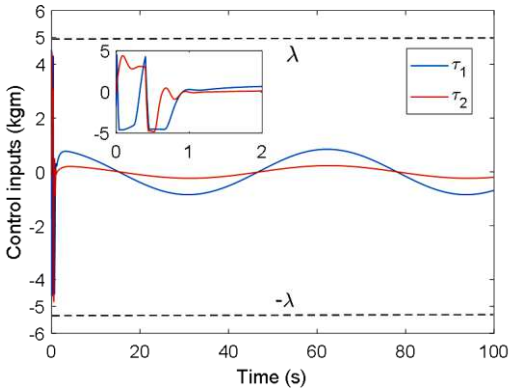


Fig. 8. The control inputs

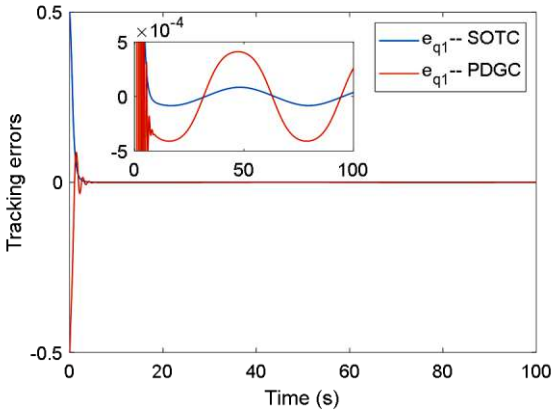


Fig. 9. Compare the error tracking the position of the first joint position

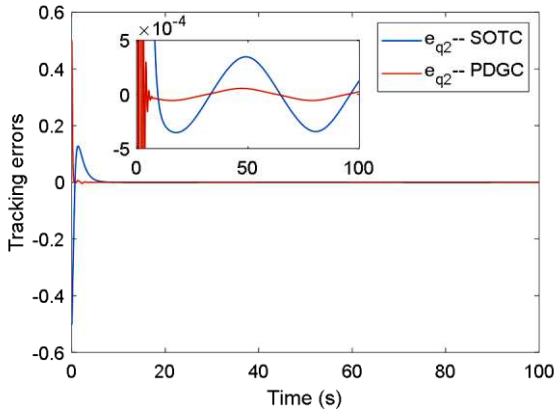


Fig. 10. Compare the error tracking the position of the second joint position

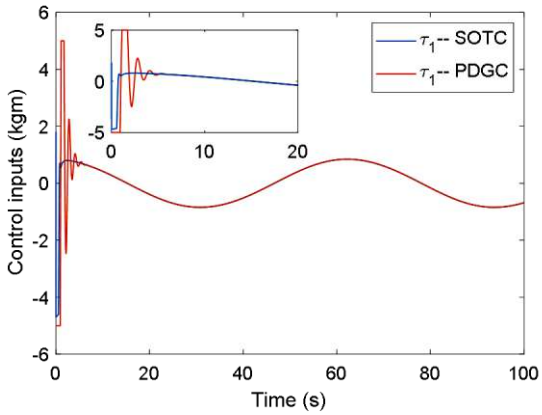


Fig. 11. Compare the inputs torque of the first joint position

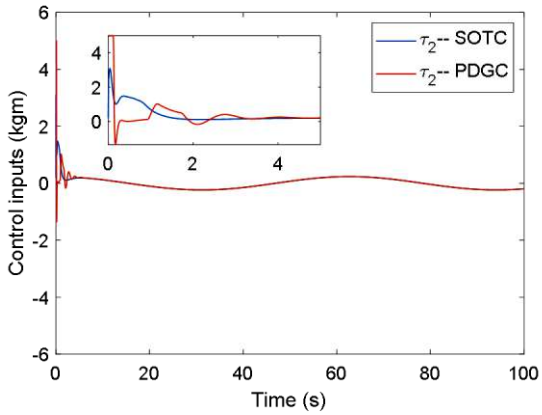


Fig. 12. Compare the inputs torque of the second joint position

The cause of this problem is: Case 1) The states $e$reach 0 too late, i.e., σ becomes 0 after oscillation. Case 2) The states e reach 0 too early, i.e., σ becomes 0 before the weights converge. The result of case 2 is shown in Fig. 2.

Figs. 3 and 4 show the tracking position of the joints during and after learning. The results show that the designed controller has good tracking position performance. Fig. 5 shows that the tracking error is 5.10^{-4}after learning.

With the feedforward control inputs are $\|\tau_a\| \leq 5 - 0.6\,tanh(1)$ and the saturated optimal feedback control inputs are $\|u^*\| \leq 0.6$, then the control inputs will be constrained by $\|\tau\| \leq 5$, shown in Fig.6, 7, and 8.

Fig. 3 to Fig. 8 show that the controller has been designed to balance both the quality of tracking performance and the control energy from the torque at the joints of the robot with input constraints. To clarify the performance of the controller, we compare the saturated optimal tracking control (SOTC) with the PD gravity controller (PDGC).

Fig. 9 to Fig. 12 show that the PDGC oscillates vigorously when the torque is saturated, while the SOTC has no oscillation, is smooth, and has a smaller error tracking the position than the PDGC. Through the above comparison simulation results, the control performance of SOTC is better than PDGC.

4 CONCLUSIONS

This paper proposed an optimal tracking controller for robot manipulators with saturation control torque based on a reinforcement learning method (RL). The feedforward control inputs were designed to transform the position tracking control problem into the optimal tracking control problem. The optimal control law with input constraint is approximated based on the online RL algorithm using only a single NN. The CL

technique was used to relax the PE condition when updating the NN weights. The controller, which has been designed to ensure that the closed-loop system states and the approximation error are UUB, the NN weights, and the control law converge to a near-optimal value. The simulation results show that the SOTC regulator is satisfactory and has good performance. The next development direction will be to build optimal tracking for robot manipulations with saturation torque, disturbances, and unknown dynamics.

APPENDIX

PROOF OF THE THEOREM 1

Choose a Lyapunov function as follows

$$J_3 = \frac{1}{2}\alpha_2 e^T e + \frac{1}{2} trace\left(\widetilde{W}^T \widetilde{W}\right) = J_{31} + J_{32} \quad (41)$$

Consider case 1: $e^T\left(\bar{f} + g\hat{u}\right) < 0$

First, we take the derivative of the functionJ_{31}.

$$\dot{J}_{31} = \alpha_2 e^T \dot{e} = \alpha_2 e^T\left(\bar{f} + g\hat{u}\right) \quad (42)$$

If$\alpha_2 e^T\left(\bar{f} + g\hat{u}\right) < 0$, then it is easy to see that $\exists \lambda_0 > 0$ such that $\alpha_2 e^T\left(\bar{f} + g\hat{u}\right) < -\alpha_2\lambda_0\|e\|$.

So, we have

$$\dot{J}_{31} < -\alpha_2\lambda_0\|e\| \quad (43)$$

Continuing, we calculate$\dot{J}_{32}$. We have

$$\dot{J}_{32} = -\alpha_1 m\left(\sigma^T\widetilde{W} - \varepsilon_H\right) - \alpha_1\sum_{i=1}^{P} m(t_i)\left(\sigma^T(t_i)\widetilde{W} - \varepsilon_H(t_i)\right) \quad (44)$$

With $m = \frac{\sigma\sigma^T}{(\sigma^T\sigma+1)^2}$

$$\rightarrow \dot{J}_{32} = -\alpha_1\widetilde{W}^T\psi\widetilde{W} + \alpha_1\widetilde{W}^T(m\varepsilon_H + \sum_{i=1}^{P} m(t_i)\,\varepsilon_H(t_i)) \quad (45)$$

Where $\psi = m\sigma^T + \sum_{i=1}^{P} m(t_i)\,\sigma^T(t_i)$. We have $m\sigma^T > 0$ and $\{\sigma(t_i)\}_{i=1}^{P}$ are linearly independent, so we have$\psi > 0$.

Using Young's inequality, we have

$$\dot{J}_{32} \le -\lambda_{min}(\psi)\left\|\widetilde{W}\right\|^2 + \left\|\widetilde{W}\right\|^2 m\sigma^T + \frac{\alpha_1^2\varepsilon_H}{4} + \left\|\widetilde{W}\right\|^2\sum_{i=1}^{P} m(t_i)\,\sigma^T(t_i) + \frac{\alpha_1^2 P\varepsilon_H^2}{4}$$

$$\rightarrow \dot{J}_{32} \le -(\alpha_1 - 1)\lambda_{min}(\psi)\left\|\widetilde{W}\right\|^2 + \frac{{\alpha_1}^2}{4}(P+1)\varepsilon_{Hm}^2 \quad (46)$$

where $\varepsilon_{Hm}$is the upper bound of ε_H.

From (43) and (46), we have

$$\dot{J}_3 \le -\alpha_2\lambda_0\|e\| - \beta_1\left\|\widetilde{W}\right\|^2 + \beta_2 \quad (47)$$

where $\beta_1 = (\alpha_1 - 1)\lambda_{min}(\psi)$, $\beta_2 = \frac{{\alpha_1}^2}{4}(P+1)\varepsilon_{Hm}^2$. We choose $\alpha_1 > 1$to $\beta_1 > 0$.

$\dot{J}_3 \le 0$, if and only if

$$\begin{cases} \|e\| \ge \frac{\beta_2}{\alpha_2\lambda_0} = \bar{s}_e \\ \left\|\widetilde{W}\right\| \ge \sqrt{\frac{\beta_2}{\beta_1}} = \bar{s}_{\widetilde{W}} \end{cases} \quad (48)$$

Consider case 2: $e^T\left(\bar{f} + g\hat{u}\right) \ge 0$

Formula (43) becomes the following:

$$\dot{J}_{31} \le -\alpha_2\lambda_0\bar{f} - \frac{1}{2}\alpha_2 e^T G\phi_e^T\widehat{W} \quad (49)$$

According to the properties of (1), the first term (49) $\alpha_2 e^T\bar{f}$ is bounded by

$$\alpha_2 e^T\bar{f} < -\beta_3\|e\|^2 \quad (50)$$

with $\beta_3 = \alpha_2 s_{\bar{f}}$, $s_{\bar{f}}$ is the upper bound of $\bar{f}$.

Substituting (40) into J_{32} of (41) and combining with (49), we have

$$\dot{J}_3 \le -\beta_3\|e\|^2 - \beta_1\left\|\widetilde{W}\right\|^2 + \beta_2 - \frac{1}{2}\alpha_2 e^T G\phi_e^T W \quad (51)$$

Then, (51) becomes

$$\dot{J}_3 \le -\beta_3\left(\|e\| - \frac{\beta_4}{2\beta_3}\right)^2 - \beta_1\left\|\widetilde{W}\right\|^2 + \beta_2 + \frac{\beta_4^2}{4\beta_3^2} \quad (52)$$

where $\beta_4 = \frac{1}{2}\alpha_2 s_G s_{\Delta\phi} s_W$.

$\dot{J}_3 \le 0$, if and only if

$$\begin{cases} \|e\| \geq \sqrt{\frac{4\beta_3^2\beta_2+\beta_4^2}{4\beta_3^3}} + \frac{\beta_4}{2\beta_3} = \bar{\bar{s}}_e \\ \|\widetilde{W}\| \geq \sqrt{\frac{4\beta_3^2\beta_2+\beta_4^2}{4\beta_3^2\beta_1}} = \bar{\bar{s}}_{\widetilde{W}} \end{cases} \quad (53)$$

From (48) and (53), we can be seen that when J_3surpasses a certain bound, then the updated law $\widehat{W}$ (39) and (40) will make$\dot{J}_3 \leq 0$. Thus, according to the extended Lyapunov theorem, the tracking and NN errors are UUB. We notice that $\|e\| \geq s_e$and $\|\widetilde{W}\| \geq s_{\widetilde{W}}$where $s_e = max(\bar{s}_e, \bar{\bar{s}}_e)$and $s_{\widetilde{W}} = max(\bar{s}_{\widetilde{W}}, \bar{\bar{s}}_{\widetilde{W}})$.

From (20), (31), (35), (36), we have$\|V^*(e) - \hat{V}(e)\| \leq s_{\widetilde{W}} s_\phi + s_\varepsilon = s_v$, and $\|u^* - \hat{u}\| \leq \frac{1}{2\rho}\lambda_{min}(R) s_g\left(s_{\widetilde{W}} s_{\nabla\phi} + s_{\nabla\varepsilon}\right) = s_u$, with $s_v \geq 0$, $s_u \geq 0$. We see that the error of the approximation of the cost function and the error of the approximation of the optimal control law are bounded by positive constants. Good convergence quality can be achieved by choosing the appropriate α_1, α_2 factor. The Proof is completed.

REFERENCES

[1] F. L.Lewis, D. M.Dawson, and C. T.Abdallah, *Robot Manipulator Control Theory and Practice*. CRC Press, 2003.

[2] S. Zeghloul, M. A. Laribi, and J.-P. Gazeau, "Robotics and Mechatronics," presented at the Procedings of the 4th IFToMM International Symposium on Robotics and Mechatronics, 2015.

[3] D. Chwa and H. Kwon, "Nonlinear Robust Control of Unknown Robot Manipulator Systems With Actuators and Disturbances Using System Identification and Integral Sliding Mode Disturbance Observer," *IEEE Access*, vol. 10, pp. 35410-35421, 2022.

[4] J. Baek, W. Kwon, and C. Kang, "A new widely and stably adaptive sliding-mode control with nonsingular terminal sliding variable for robot manipulators," *IEEE Access*, vol. 8, pp. 43443-43454, 2020.

[5] W. Jie, Z. Yudong, B. Yulong, H. H. Kim, and M. C. Lee, "Trajectory tracking control using fractional-order terminal sliding mode control with sliding perturbation observer for a 7-DOF robot manipulator," *IEEE/ASME Transactions on Mechatronics*, vol. 25, no. 4, pp. 1886-1893, 2020.

[6] V.-C. Nguyen, A.-T. Vo, and H.-J. Kang, "A non-singular fast terminal sliding mode control based on third-order sliding mode observer for a class of second-order uncertain nonlinear systems and its application to robot manipulators," *IEEE Access*, vol. 8, pp. 78109-78120, 2020.

[7] L. Zhang, Y. Wang, Y. Hou, and H. Li, "Fixed-time sliding mode control for uncertain robot manipulators," *IEEE Access*, vol. 7, pp. 149750-149763, 2019.

[8] R.-D. Xi, X. Xiao, T.-N. Ma, and Z.-X. Yang, "Adaptive Sliding Mode Disturbance Observer Based Robust Control for Robot Manipulators Towards Assembly Assistance," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 6139-6146, 2022.

[9] J. Chi, H. Yu, and J. Yu, "Hybrid tracking control of 2-DOF SCARA robot via port-controlled hamiltonian and backstepping," *IEEE Access*, vol. 6, pp. 17354-17360, 2018.

[10] M. Van, M. Mavrovouniotis, and S. S. Ge, "An adaptive backstepping nonsingular fast terminal sliding mode control for robust fault tolerant control of robot manipulators," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 7, pp. 1448-1458, 2018.

[11] T. N. Truong, A. T. Vo, and H.-J. Kang, "A backstepping global fast terminal sliding mode control for trajectory tracking control of industrial robotic manipulators," *IEEE Access*, vol. 9, pp. 31921-31931, 2021.

[12] J. Zhou, E. Liu, X. Tian, and Z. Li, "Adaptive Fuzzy Backstepping Control Based on Dynamic Surface Control for Uncertain Robotic Manipulator," *IEEE Access*, vol. 10, pp. 23333-23341, 2022.

[13] M. Van and S. S. Ge, "Adaptive fuzzy integral sliding-mode control for robust fault-tolerant control of robot manipulators with disturbance observer," *IEEE Transactions on Fuzzy Systems*, vol. 29, no. 5, pp. 1284-1296, 2020.

[14] W. Jie, L. M. Cheol, K. Jaehyung, and K. H. Hee, "Fast fractional-order terminal sliding mode control with rbfnn based sliding perturbation observer for 7-dof robot manipulator," *IEEE Access*, vol. 9, pp. 67117-67128, 2021.

[15] S. Ling, H. Wang, and P. X. Liu, "Adaptive fuzzy dynamic surface control of flexible-joint robot systems with input saturation," *IEEE/CAA Journal of Automatica Sinica,* vol. 6, no. 1, pp. 97-107, 2019.

[16] C. Yang, D. Huang, W. He, and L. Cheng, "Neural control of robot manipulators with trajectory tracking constraints and input saturation," *IEEE Transactions on Neural Networks and Learning Systems,* vol. 32, no. 9, pp. 4231-4242, 2020.

[17] Q. Zhou, S. Zhao, H. Li, R. Lu, C. Wu, and L. Systems, "Adaptive neural network tracking control for robotic manipulators with dead zone," *IEEE Transactions on Neural Networks,* vol. 30, no. 12, pp. 3611-3620, 2018.

[18] K. G. Vamvoudakis and F. L. Lewis, "Online actor–critic algorithm to solve the continuous-time infinite horizon optimal control problem," *Automatica,* vol. 46, no. 5, pp. 878-888, 2010, doi: 10.1016/j.automatica.2010.02.018.

[19] H. Zargarzadeh, T. Dierks, and S. Jagannathan, "Adaptive neural network-based optimal control of nonlinear continuous-time systems in strict-feedback form," *International Journal of Adaptive Control and Signal Processing,* vol. 28, no. 3-5, pp. 305-324, 2014.

[20] H. Modares, F. L. Lewis, and Z. P. Jiang, "H infinity tracking control of completely unknown continuous-time systems via off-policy reinforcement learning," *IEEE Transactions on Neural Networks and Learning Systems,* vol. 26, no. 10, pp. 2550-62, Oct 2015, doi: 10.1109/TNNLS.2015.2441749.

[21] R. Kamalapurkar, H. Dinh, S. Bhasin, and Warren E Dixon, "Approximate optimal trajectory tracking for continuous-time nonlinear systems," *Automatica,* vol. 51, pp. 40-48, 2015, doi: 10.1016/j.automatica.2014.10.103.

[22] L. Kong, W. He, C. Yang, and C. Sun, "Robust Neurooptimal Control for a Robot via Adaptive Dynamic Programming," *IEEE Trans Neural Networks Learn Systems,* vol. 32, no. 6, pp. 2584-2594, Jun 2021, doi: 10.1109/TNNLS.2020.3006850.

[23] L. N. Tan, "Distributed H∞ optimal tracking control for strict-feedback nonlinear large-scale systems with disturbances and saturating actuators," *IEEE Transactions on Systems, Man, and Cybernetics: Systems,* vol. 50, no. 11, pp. 4719-4731, 2018.

[24] L. Nguyen Tan, "Distributed optimal control for nonholonomic systems with input constraints and uncertain interconnections," *Nonlinear Dynamics,* vol. 93, no. 2, pp. 801-817, 2018.

[25] L. N. Tan and T. C. Pham, "Optimal Tracking Control for PMSM With Partially Unknown Dynamics, Saturation Voltages, Torque, and Voltage Disturbances," *IEEE Transactions on Industrial Electronics,* vol. 69, no. 4, pp. 3481-3491, 2021.

[26] X. Zhao, B. Tao, L. Qian, and H. Ding, "Model-Based Actor-Critic Learning for Optimal Tracking Control of Robots With Input Saturation," *IEEE Transactions on Industrial Electronics,* vol. 68, no. 6, pp. 5046-5056, 2021, doi: 10.1109/tie.2020.2992003.

[27] K. G. Vamvoudakis, M. F. Miranda, and J. P. Hespanha, "Asymptotically Stable Adaptive-Optimal Control Algorithm With Saturating Actuators and Relaxed Persistence of Excitation," *IEEE Trans Neural Netw Learn Syst,* vol. 27, no. 11, pp. 2386-2398, Nov 2016, doi: 10.1109/TNNLS.2015.2487972.

[28] N. T. Luy, "Reinforecement learning-based optimal tracking control for wheeled mobile robot," in *2012 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER)*, 2012: IEEE, pp. 371-376.

ICATSD2F.113

ENRICHING KNOWLEDGE GRAPH OF COMPUTING DOMAIN ONTOLOGY BY HETEROGENEOUS RESOURCES

TA DUY CONG CHIEN

Faculty of Information Technology Department, Industrial University of Ho Chi Minh City,

taduycongchien@iuh.edu.vn

Abstract. Ontologies play an important role in the recent years. Its applications now are more varieties and popularities. Ontologies are used in different areas relevant to Information Technology, Biology, Medicine, especially in Information Retrieval, Information Extraction, and Question Answering. This paper will introduce an overview and explanation of the most successful and efficient algorithms for enriching ontology by heterogeneous resources such as Wikipedia, WordNet. Special focus is on approaches that are fairly generic, i.e., that can be applied for processing heterogeneous data collections rather than text type or web page and that are as much language independent as possible. The paper also mentions necessary knowledge and resources involved in the enriching ontology, which is built by knowledge graph, and how to evaluate ontology. Additionally, a very important focus is on statistical, and natural language processing techniques for information detection and classification.

Keywords. Knowledge graph; Topic-oriented search; Domain Ontology.

1 INTRODUCTION

The main goal of this paper is to give a comprehensive overview of algorithms used to extract concepts from text and XML files to build and enrich domain specific ontology. Whereas traditional methods recognize information only in text or web page in a deterministic way and represents this information in a format with known semantics, in this paper, our ontology can process text documents that have structure such as XML and text file to enrich itself. Almost equal importance is given to early technologies developed in the field primarily with the aim of natural language understanding, as to the most advanced and recent technologies for building Ontologies. The past approaches are an incentive to identify some forgotten avenues that can be researched to advance the state of the art. Machine learning is playing a central role in the development of the novel technologies and contributes to the portability and widespread use of information extraction technology. This paper will especially focus on supervised learning algorithms that are very promising for ontology learning.

A second important thing is to focus on the prospects of domain specific ontology to be used in modern information systems, and more specifically in information retrieval systems. We want to demonstrate that on one hand the statistical and machine learning techniques traditionally used in domain ontology have little attention for the underlying cognitive and linguistic models that shape the patterns that we are attempting to detect. On the other hand, current our algorithms use semantic rules, synonyms, hyponyms, and hypernyms to exhibit the semantic features between concepts in domain ontology. Thereby, this ontology can be used for applications that are semantic-oriented.

Although concept extraction considers certain characteristics of the language and the examples are usually taken from the English language, our discussion is as language independent as possible. Moreover, our ontology can process text documents that have structure such as XML, text file to enrich it automatically.

This paper proposes domain specific ontology that relates to Computing domain, called Computing Domain Ontology (CDO). We want to illustrate the technologies, algorithms to build and enrich the CDO from text documents, ACM categories, Wikipedia, WordNet and evaluate the results as they are currently implemented in this domain. The evaluation should also lead to better insights into the points of attention for future research. It is organized as follows: section 2 describes related works; section 3 details the steps

to build and automatically enrich domain ontology; section 4 focuses on a methodology to evaluate the generated ontology; section 5 discusses the results of the evaluation methodology.

2 RELATED WORKS

With the huge data on internet that they come from many different sources, most of the data access applications, such as Information Retrieval, Information Extraction, Question and Answer trend using Ontology as knowledge base.

First of them is the research of Ana's group [1]. They built the web topic ontologies by classifying Web pages content into groups with different topics. After that, they re-organized them into a hierarchical scheme and build cross-references in kinds of "is-a", "symbolic" and "related" between different topics in a non-hierarchical scheme. As outline from Graham Harman [2], he given a brief history and conceptual overview of object-oriented ontology (OOO) itself and then, he given an assessment of how OOO might fit with some current discussions of trans- and posthumanism and finally, he given some basic examples of how OOO might be applicable to film and television criticism. P. Murphy et al [3] proposed ontology of objects and non-human worlds is central to cocreation work in energy research where there exist assemblages of the Anthropocene. They argue that an ethical, engaged, object-oriented ontology that links with fictive imaginaries is crucial whichever direction Science, Technologies and Social (STS) takes in energy research. In the Oracle database with 21c release [4] it allowed users to build and enrich ontology using RDF graph, to query semantic data and to perform ontology-assisted query of enterprise relational data, and to use supplied or user-defined inference to expand the power of querying on semantic data. G. Yoran [5] said that as computer science works on domain-specific models in order to find solutions to practical problems, employing models of the world, informatics is – like any proper science – applied metaphysics. M. Allahyari and K. Kochut [6] created a topic model that integrates ontological concepts with topic models in a single framework, where each topic and each concept are represented as a multinomial distribution over concepts and over words, respectively, and a topic labeling method based on the ontological meaning of the concepts included in the discovered topics. Niels Wilde [7] provided a method explores the status of relations in OOO. On the one hand, objects often seem to get the upper hand, since what an object is cannot be reduced to the sum of its (external) relations. A.A. Salatino et al [8] provided Computer Science Ontology (CSO). CSO is a large-scale, automatically generated ontology of research areas, which includes about 14K topics and 162K semantic relationships. It was created by applying the Klink-2 algorithm on a very large data set of 16M scientific articles. El Ghosh et al [9] proposed an approach to identify the themes necessary to describe the knowledge structure of an application domain. Meanwhile, their development from scratch is hard and time-consuming task. This paper discusses the development a topic-specific ontology, named Topic-OPA, for modeling topics of old press articles. Topic-OPA is extracted from the open knowledge graph Wikidata by the application of a SPARQL-based fully automatic approach. The development process of Topic-OPA depends mainly on a set of disambiguated named entities representing the articles. R. Jabla et al [10] proposed an automatic ontology-based model evolution approach to account for highly dynamic environments at runtime. The approach can extend initial models expressed as ontologies to cope with rapid changes encountered in surrounding dynamic environments at runtime.

A domain specific ontology usually consists of concepts/entities/types [11] of the domain, which are captured and organized by various types of relationships that hold between them. Relationships, which organize concepts hierarchically, are called vertical relationships (e.g., the 'is-a' relationship) in contrast to horizontal relationships (e.g., the 'synonymy' relationship) which link entities across the hierarchy. The instantiation of the concepts in a particular domain is performed by instances/objects/individuals that are members of concepts.

3 AN APPROACH FOR ENRICHING COMPUTING DOMAIN ONTOLOGY BY HETEROGENEOUS RESOURCES

3.1 Overview Computing Domain Ontology

Some concepts should be defined first:

Definition 1: Domain Ontology includes concepts, attributes and events. Ontology is a tuple O = (C, I, R, T, V, $\leq$, $\perp$, $\in$, =) [6], where:

C is the set of concepts, I is the set of individuals including attributes and events. For instance, in Figure 1, Software is a concept.

R is the set of relations. For instance, in Figure 1, there is a "SUB-OF" relation between Software programming and Software.

T is the set of data types. For instance, in Figure 1, Software has String data type.
V is the set of values (C, I, R, T, V being pair wise disjoint); ≤ is a relation on (C×C) ∪ (R×R) ∪ (T×T) called specialization; ⊥ is a relation on (C×C) ∪ (R×R) ∪ (T×T) called exclusion; ∈ is a relation over (I×C) ∪ (V× T) called instantiation.; = is a relation over I×C× (I∪V) called assignment. For instance, in Figure 1, Software programming is specialization of Software; algorithm can be declared to be exclusive to data structure.

Definition 2: Semantic role regards the assignment semantic to the constituent of sentence. They regard certain actions or states, their participants, and their circumstances. Semantic roles can be very general defined or be more specific.

Definition 3: Entity Relation is relation between two or more entities that is detected, and the relation possibly is typed with a semantic role. Entities can be nouns or compound nouns.

Definition 4: An instance presented a concept. In this CDO, instances are concepts belong to Information Technology area.

To build domain ontology usually requires overcoming many obstacles because the data has been collected from different sources. Therefore, how to consolidate data is a big problem. Our CDO hierarchy has fourth layer:

• First layer is known Topic layer. This layer represents to set of categories (C).
• Second layer is known Ingredient layer. This layer presents to set of instances (I) that are belong to first layer
• Third layer is known Synset layer. This layer presents synonyms, hyponyms, hypernyms of instances in second layer
• Fourth layer is known Sentence layer. This layer presents set of relations (R) between instances of second layer.

The CDO hierarchy is given in Figure 1.

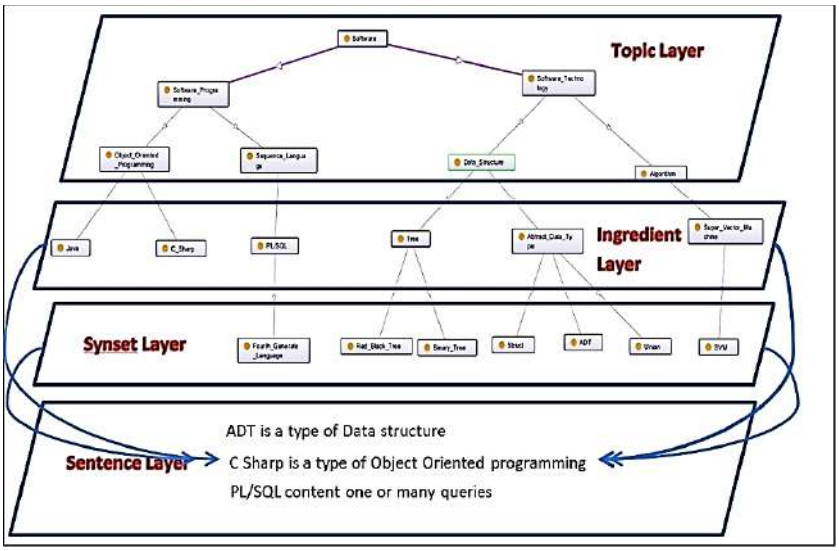


Figure 1. The CDO hireachy.

3.2 Enriching Knowledge Graph of Computing Domain Ontology

Secondly, instead of building the set of attributes (C), we build a set of instances that are belonged to categories in the topic layer. We extract data from Wikipedia ontology and ACM Library Digital [12]. Wikipedia is a rich ontology, which includes many areas and many languages. However, we only focus on

Computer domain and English language. Similarity to text files of ACM Library Digital is also related to Computer domain. The words that are extracted from both sources will be converted to standard form and be put into ingredient and sentence layers of CDO according to their category. We propose an enriching model as shown in Figure 2.

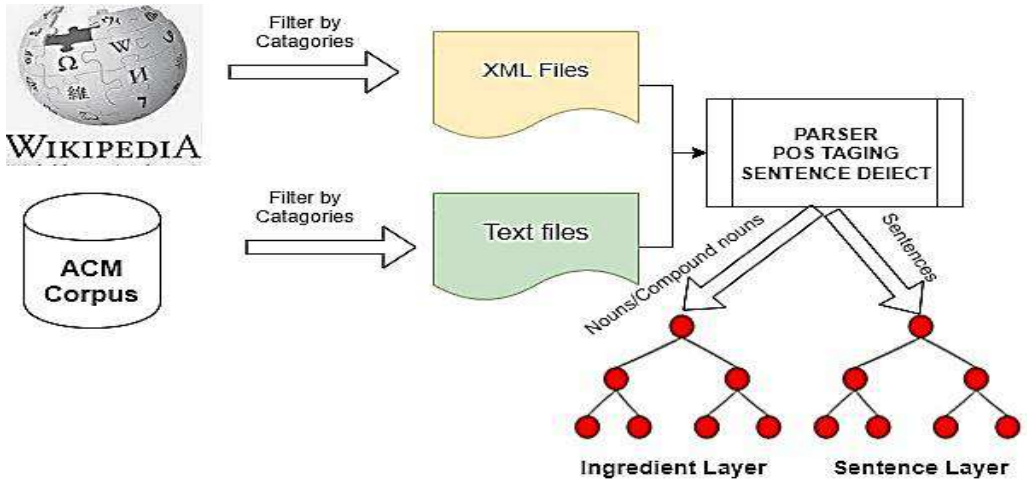


Figure 2. Model of enriching knowledge graph of CDO.

The algorithm for enriching Ingredient and Sentence layers is shown as follows.
Input: a set of XML files from Wikipedia and text file of ACM Digital Library.
Output: instances of Ingredient layer of CDO
Procedure enriching Ingredient and Sentence Layers
Begin
Loading (XML/Text File(categoryID)) into Buffer
While (Buffer is not null)
Begin

```
sentence = Remove (StopWord)
Array linguistic[] = Parser(sentence)
Array Nouns[],  Noun_phrases[], Sentences[]
While (lin in linguistic)
Begin
      If (lin is noun) then
          Add (lin into Nouns[])
      Else
          if (lin is compound_noun) then
              Add (lin into Noun phrases [])
          end if
          else
              if (lin is full_sentence) then
                 Add (lin into Sentences [])
              end if
     End if
     End
  End while
  While (noun in Nouns[])
        Put (noun into Ingredient layer by Category)
  End while
  While (nph in Noun phrases[])
        Put (nph into Ingrdient layer by Category)
```



```
    End while
    While (sen in Sentences[])
        Put (sen into Sentence layer by Category)
    End while
  End Procedure.
```

Enriching ontology also plays an important role in building Ontology. To enrich CDO, we use WordNet and semantic roles for building synset and sentence layer of the CDO.

Synset layer is a set of synonyms, hyponyms, and hypernyms of instances of ingredient layer. They are extracted from WordNet. We propose an algorithm (Enriching_Synset_Layer) for collecting these words as follows.

```
Input: WordNet Lexical Ontology
Output: instances of synset layer of CDO
Procedure Enriching_Synset_Layer ()
  While (instance of Ingredient layer is not null)
  Begin
      Synonym_List = WordNet_query (instance)
      Hyponym_List = WordNet_query (instance)
      Hypernym_List = WordNet_query (instance)
      If (Synonym_List is not null)
         Link (instance to Synonym_list)
      End if;
      If (Hyponym_List is not null)
         Link (instance to Hyponym_list)
      End if
     If (Hypernym_List is not null)
        Link (instance to Hypernym_list)
      End if
  End;
End While
End Procedure
```

The same "WordNet_query" method is invoked three times in this procedure because we want to separate the set of synonyms, hyponyms, and hypernyms of instances when they are stored in this layer.

After building and enriching CDO, the knowledge graph of CDO is shown as Figure 3.

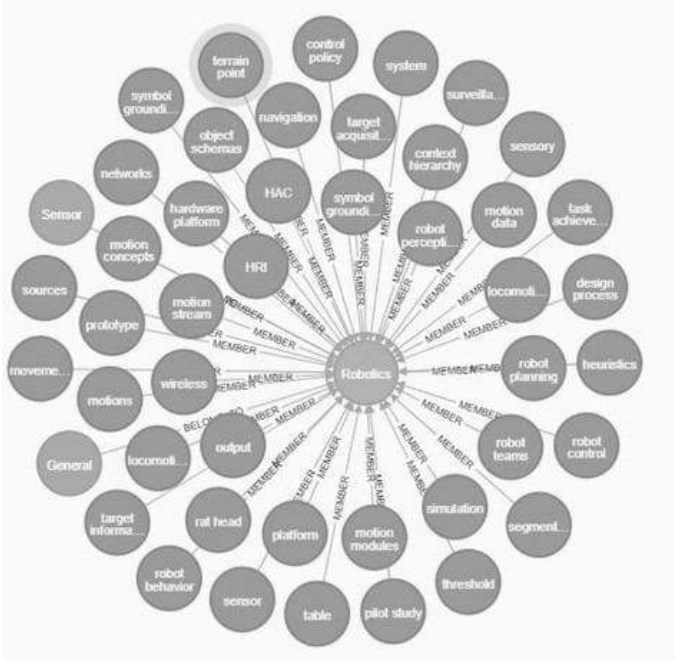


Figure 3. The CDO is represented by graph database (Neo4J).

4 EXPERIMENT AND EVALUATION

Evaluation CDO's performance has been measured using three factors: Precision, Recall and F-Measure. These factors are calculated to each category in CDO as below:

$$P(C_i) = \frac{Correct(Ci)}{Correct(Ci)+Wrong(Ci)} \quad (1)$$

$$RC_i = \frac{Correct(Ci)}{Correct(Ci)+Missing(Ci)} \quad (2)$$

$$\text{F-Measure} = 2\frac{\text{Precision}*\text{Recall}}{\text{Precision}+\text{Recall}} \quad (3)$$

Where C_i represents a category in CDO and correct, wrong, missing represent the number of correct, wrong, missing, respectively.

Table 1 respectively shows the testing results through three above measures. We pick five categories at random at the topic layer for illustration.

Table 1. The result of Wikipedia testing

Categories	Precision	Recall	F-Measure
Artificial Intelligent (AI)	97.03%	88.62%	93.00%
Logic Design (LD)	82.41%	54.72%	71.20%
Operating System (OS)	84.47%	81.37%	83.00%
Process Management (PM)	96.72%	76.02%	86.00%
Software (Soft)	96.52%	92.19%	95.00%

The comparison of Precision, Recall and F-Measure is shown as Figure 4.

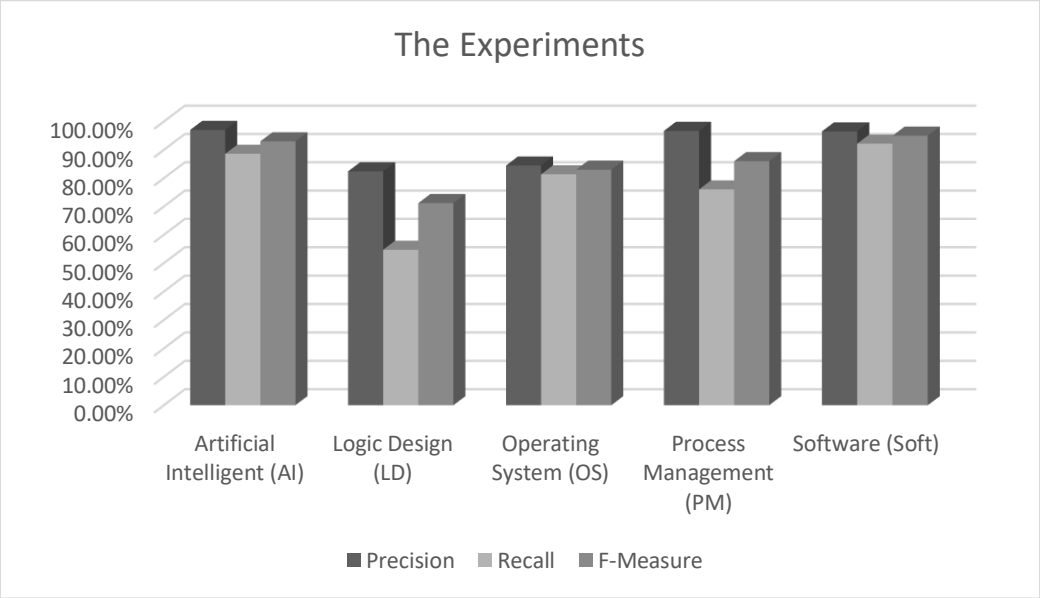


Figure 4. The comparison of Precision, Recall and F-Measure.

The result from Fig. 4 reveals that for different categories, the percentage of precision, recall and F-measure are different. The category "Logic design" has the lowest percentage among that of other categories because the nouns, compound nouns and sentences which are extracted from the different resources are less than the others.

5 CONCLUSIONS AND FUTURE WORKS

In this paper, we dealt with the problem of enriching domain-specific ontology from different sources, such as: Wikipedia, WordNet, and ACM Digital Libraries. However, we only focus on Computing domain and English language. Our goal is to find a way to extract reliable data from those sources, and to convert them into a useable standard form. We propose an approach to enrich CDO by extracting of words from many different sources for enriching the major layers of CDO hierarchy, such as ingredient layer, synset layer and sentence layer. Besides, we use WordNet ontology to enrich the synset layer of CDO. We also used methodologies, algorithms related to Natural Language Processing, and Statistical methods to solve our problems. Results generated by such experiments show that our proposal approaches are valid.

In the future work, we will focus particularly on automated ontology evaluations and how to detect automatically semantic relationship between concepts, the necessary perquisites for the fruitful development of automated ontology processing techniques to solve a number of problems, such as ontology learning and matching.

REFERENCES

[1] A.G. Maguitman, L. Cecchini, C.M. Lorenzetti, F. Menczer, "Using Topic Ontologies and Semantic Similarity Data to Evaluate Topical Search," in *The Proceedings of 36th Latin American Informatics Conference (CLEI)*, Asuncion, Paraguay, 2010.

[2] G. Harman, "Object-Oriented Ontology," in *The Palgrave Handbook of Posthumanism in Film and Television*, 2015, pp. 401-409.

[3] P. Murphy, P. Brereton, F. O'Brolchain, "New materialism, object-oriented ontology and fictive imaginaries: new directions in energy researc," *Energy Research & Social Science Journal,* vol. 79, 2021.

[4] Oracle, [Online]. Available: https://docs.oracle.com/en/database/oracle/oracle-database/21/rdfrm/lof.htm. [Accessed January 2022].

[5] G. Yoran, "Applied Metaphysics – Objects in Object-Oriented Ontology and Object-Oriented," *Interface Critique Journal,* vol. 1, 2018.

[6] M. Allahyari, K. Kochut, "Automatic Topic Labeling Using Ontology-Based Topic Models," in *The Proceedings of 14th International Conference on Machine Learning and Applications (ICMLA)*, 2015.

[7] Wilde, "Burning Bridges: The problem of relations in object-oriented ontology—a topological approach," *Humanities and social sciences communications Journal,* vol. 6, 2020.

[8] A.A. Salatino, T. Thanapalasingam, A. Mannocci, A. Birukou, F. Osborne & E. Motta, "The computer science ontology: A comprehensive automatically-generated taxonomy of research areas," *Data Intelligence Journal,* vol. 2, pp. 379-416, 2020.

[9] M.E. Ghosh, C. Zanni-Merk, N. Delestre, J. Kotowicz, H. Abdulrab, "Topic-OPA: A Topic Ontology for Modeling Topics of Old Press Articles," in *The Proceedings of the 12th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2020)*, 2020.

[10] R. Jabla, M. Khemaja, F. Buendia, S. Faiz, "Automatic Ontology-Based Model Evolution for Learning Changes in Dynamic Environnements," *Applied Sciences Journal,* vol. 11, 2021.

[11] J. Nogueras-Iso, J. Lacasta, J. Teller & G. Falquet, Ontology Theory, Management and Design: Advanced Tools and Model, Ed., IGI Global, 2010.

[12] "ACM," [Online]. Available: http://www.acm.org/about/class/ccs98-html. [Accessed July 2014].

ICATSD2F.114

A NEW APPROACH IN CALCULATING CURRENT-VOLTAGE CHARACTERISTICS OF THE QUANTUM DOT GATE FIELD EFFECT TRANSISTOR

AN NGUYEN VAN

Faculty of Electronic Technology, Industrial University of Ho Chi Minh City

nguyenvanan@iuh.edu.vn

Abstract. This paper presents a new approach in calculating the current - voltage characteristics of the Quantum dot Gate Field effect Transistor (QDGFET), an important component for fabricating microelectronic devices, especially non-volatile memories. In this research, a QDGFET model consisting of two quantum dot layers SiOx cladded Si placed in the gate stack working as an effect element was proposed, analyzed, and evaluated. At first, we study the effects of quantum dots on current - voltage characteristics of QDGFET by the Schrödinger-Poisson equations to determine the transmission probability of electrons that tunnel the oxide layer and the energy levels in quantum dots. Then, using mathematical analysis, the current - voltage is calculated based on the analytical method of Metal-Oxide-Semiconductor Field Effect Transistor (MOSFET). Finally, the calculated results are simulated by Matlab software in the different conditions such as the changing of channel sizes, doping concentration in semiconductor wafer, gate oxide thickness and quantum dot sizes, etc. The obtained results are highly reliable and strongly consistent with recent research.

Keywords. Quantum dot Gate Field effect Transistor, Quantum dot device, Current-voltage characteristics of Quantum dot Gate Field effect Transistor.

1 INTRODUCTION

In recent years, nanoelectronics is one of the most important technologies interested many scientists. Some new materials are being researched and applied for microelectronic fabrication such as Carbon nanotubes (CNT), Graphene, Nanowires, Quantum dots (QD), etc. [1], [2], [3]. Among these, Quantum dots are important materials studied to fabricate Quantum dot Gate-Field effect Transistors (QDG-FETs), an element in making non-volatile memories [4], [5], [6]. Up to now, much research on this device have been published and shown many interesting results. For example, the QD-FET structure with double quantum dot layer, its operation is based on the quantum Spin effect proposed by N. Shaji, the device consist of two quantum dot layers placed in the FET's channel and can operate in two states on-off considered as a switching element [7]; The model of Quantum dot gate field-effect transistor (QDG-FET) operating in 3 states introduced by F. Karmakar applied in manufacturing the high speed ADC/DAC converters [8]; and the typical structure of quantum dot channel field-effect transistor (QDC-FET) operating in 4 states given by J.A. Jain, using II–VI barrier layers, the device is used in fabricating of SRAM memories [9]. Thereby, it shows that QD-FETs have many outstanding advantages such as small size, high speed switching, low energy consumption, etc [10], [11]. Since then, QD-FET became a promising candidate in the application for making microelectronic devices. This paper introduces a QDG-FET model consisting of two quantum dots placed in the gate stack working as a memory element; moreover, it proposes a new analytical method to determine the current-voltage characteristics in the device that gives us an overview of the important properties of this particular device.

2 QUANTUM DOT MATERIAL AND QDG-FET MODEL

2.1 Quantum dot material

Quantum dot is a type of semiconductor crystals, a few nanometers. Its diameter is about 2 nm to 15 nm

and about 100 to 100,000 atoms [12]. Quantum dot can be made from semiconductor materials, metals, or polymers, etc. In quantum dot, charged particles cannot move freely, only existing in discontinuous energy states in three-dimensional space. The typical property of QD is that if its size is changed, the energy band structure is changed. As the result, the band gap in the energy levels is also changed. In other word, when the band gap is changed, the photons' emission wavelength is also affected [13].

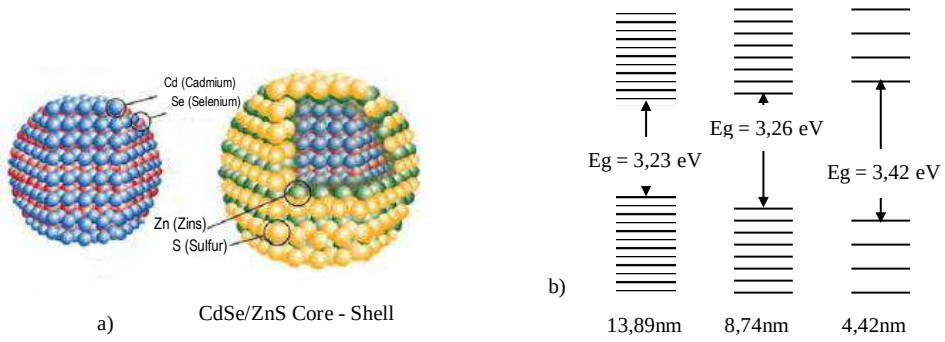


Figure 1. a) Quantum dot structure; b) Energy band diagram of QD ZnO [14]

Regarding the surface effects, the atoms on the surface of the quantum dot have many different properties in comparison to the others inside. Once the size of quantum dot is changed, the surface effect is also affected. The ratio between the number of atoms on the surface and the total number of atoms in the quantum dot is determined by [11]:

$$f = \frac{n_s}{n} = \frac{4n^{2/3}}{n} = \frac{4}{n^{1/3}} = \frac{4r_o}{r} \quad (1)$$

Where, $n_s$is the number of atoms on the surface, n is the total number of atoms, r_o is the radius of the atom and r is the radius of the particle. Thus, as the size of quantum dot decreases, the surface effect will increase, the smaller the particles the larger the surface effect. Figure 2 shows the distribution of atoms on the surface in comparison with the others in the QD.

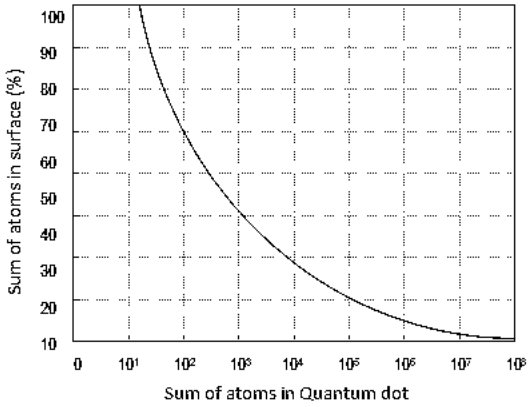


Figure 2. The distribution of atoms on the surface in comparison with the others in the QD [15]

Concerning the quantum confinement effect (size effect), once the size of semiconductor is twice smaller than exciton Bohr radius, the excitons are squeezed, leading to quantum confinement. The energy levels are determined by the "particle in a box'' model [16], energy states depend on the length of the box. Accordingly, there are 3 regimes including strong, weak, and intermediate confinement. The strong confinement regime is defined as the quantum dot radius being smaller than both electron and hole Bohr radius, the weak confinement is given as the quantum dot is larger than both. In intermediate confinement regime, the quantum dot radius is larger than the hole of Bohr radius (in case electron and hole radius are different). The exciton Bohr radius is determined following:

$$r_B = r_o \frac{\varepsilon . m_e^*}{m^*} \quad (2)$$

Where, $r_o = 0{,}529A^o$ is the 1s orbital radius of hydrogen, m_e is the mass of the electron, m_h is the mass of the hole, ε is the dielectric constant, $m^*$is the reduced mass and expressed by $m^* = (m_e^* m_e^*)/(m_e^* + m_e^*)$. The equation (2) shows that the quantum confinement effect in different sizes depends on the properties of the material such as the dielectric constant and the mass of the electron or hole. It makes changes in the physical properties of quantum dots such as the energy band structure and density of state of charged carriers. Thus, quantum dots are considered as an artificial atom with discrete energy states of electron and hole, like an atom. Quantum dots are expressed as intermediate materials between semiconductors and independent molecules.

In the strong confinement regime, in the quantum dot, the electrons and holes are expressed as independent transmission if the Coulomb interaction is ignored. Its absorption spectrum consists of discrete bands with peak positions at the energy level E [16]:

$$E = E_{gap} + \frac{\hbar^2 \chi^2}{2m^* R^2} \quad (3)$$

Where, χ is Bessel function,$R$is size of quantum dot, $m^*$is the reduced mass of electron, $\hbar$ is the reduced Planck's constant ($\hbar = h/2\pi$), $E$is band gap energy.

The electron-hole transition in the quantum dot has the smallest energy level:

$$Ebg \frac{\hbar^2 \pi^2}{2a^2}\left(\frac{1}{m_e^*} + \frac{1}{m_h^*}\right)_{bg} \frac{\hbar^2 \pi^2}{2m^* R^2}_{min} \quad (4)$$

Thus, the energy for transferring electron in the quantum dot is increased by ΔEin comparison to it in semiconductor:

$$\Delta E = \frac{\hbar^2 \pi^2}{2m^* R^2} \quad (5)$$

However, the displacements of charges are not completely independent in a quantum dot. Approximately, the energy corresponding to the two stimulated states of the electron - hole pair is determined by the following expression:

$$Ebg \frac{\hbar^2 \chi^2}{2m^* R^2} \frac{1{,}8e^2}{\varepsilon R}_{min} \quad (6)$$

The third part of equation (6) represents the Coulomb interaction energy (exiton). In the SiOx -cladded Si quantum dots, the energy levels can be controlled by varying the size of the quantum dot and the magnitude of the electric field on material.

2.2 QDG-FET model

The QDG-FET model is built based on the top-gate field-effect transistor (n-MOSFET) structure as shown in Figure 3. The gate stack of the device consists of two layers of SiOx-cladded Si quantum dots [17] placed on top of SiO_2 gate insulator with high dielectric constant and large band gap energy formed an effective gate stack.

Two semiconductor layers $n^+$with high doping concentration (10^{16}cm^{-3}) are diffused onto the p-Silicon semiconductor wafer to form two electrodes D (Drain) and S (Source). Once power is applied, the electric field generated by the gate voltage V_{GS} caused a change in the energy levels of the quantum dot layer leading to the change the threshold voltage V_{TH}. Meanwhile, the current in the conduction channel is also controlled.

As shown in Figure 3, the gate stack of the QDG - FET consists of layers stacked each other. Consequently, the band gap diagram between them is also different and divided into different levels depending on the arrangement of the material layers

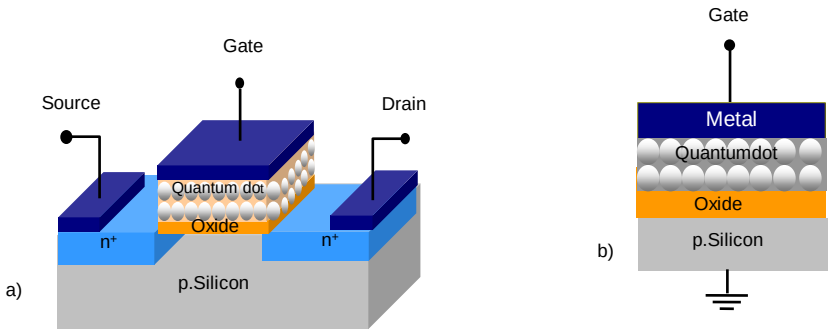


Figure 3. a) QDC-FET structure b) Gate stack of QDG-FET

The operation principle of QDG-FET can be explained by its band diagram across the gate region as shown in Figure 4. The two quantum dot layers are shown with their cores labeled as t_{qd}, the layer t_1 includes SiOx cladding ~1 nm and a tunneling insulator ~2 nm, layer thickness t_2 is 2 nm (comparison to the two SiOx claddings 1 nm each), and t_3 is the outer layer ~1 nm, which represents the cladding SiOx thickness of quantum dots. In addition, the thickness of the quantum well inversion layer is represented by t_{qw}. In the quantum dots, energy levels depend on its parameters as well as the electric field [18].

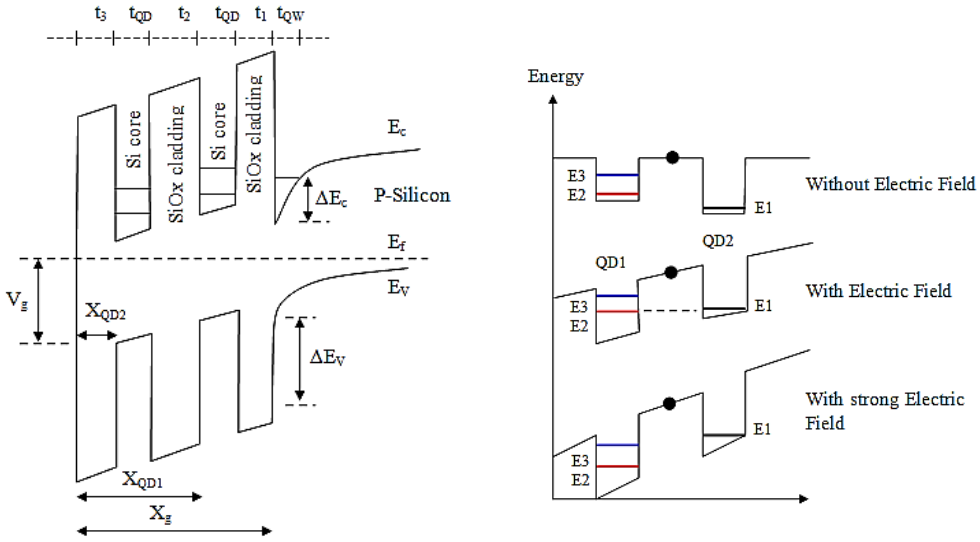


Figure 4. Energy band diagram of QDG-FET

From the energy band diagram in Figure 4, the energy levels in the quantum dot are changed while existing an electric field provided by the gate voltage V_{GS}. In the SiOx-cladded Si quantum dot layer, the energy levels are divided into many narrow sub bands (mini band), about 0.2 eV [18]. The energy levels are controlled by the electric field.

Due to the tunneling effect, the charges transfer from the channel to the quantum dot layers causing change of energy levels in the quantum dots [15], [18]. Therefore, these energy levels can be determined by the Schrödinger-Poisson equations in an inversion channel [7], [17]. This allows it to determine the threshold voltage V_{TH}, considered as a function representing the polarization.

When the gate stack is supplied by positive voltage, creating an electric field will decrease the Fermi energy level in the gate metal layer, a reversion charge layer in the channel is formed below the gate region.

Once the gate voltage increases, the energy level inside the QD moves further down and the electrons tunnel from the reversion channel to the first QD layer near the channel. The tunneling of electrons causes the increase in flat band voltage V_{FB} making the increase of threshold voltage V_T [8]. Therefore, the flat band voltage V_{FB} is expressed by:

$$V_{FB} = -\frac{q}{C_{OX}} \int_0^{X_g} \frac{x\rho(x)}{X_g} dx \qquad (7)$$

$$= -\frac{q}{C_{OX}}\left[\sum\frac{X_{QD1}n_1N_{QD1}}{X_g} + \sum\frac{X_{QD2}n_2}{X_g}\right]$$

Where, X_{QD1} and $X_{QD2}$are the distances from the gate contact to quantum dot cores, n_1 and n_2 are the number of quantum dots in layer 1 and 2, N_{QD1} and N_{QD2} are the charges on each SiOx-cladded Si quantum dots, $X_g$is the distance from the gate to the quantum dot interface, $\rho(x)$ is the charge density, and $q$is the electron charge, C_{ox} is the oxide capacitance.

When the gate voltage is higher than the threshold voltage, electrons from the inversion channel move to the first quantum dot layer close to the channel. If the gate voltage increases higher, the charges will be transferred from the first to the second quantum dot layer near the gate electrode. The intermediate states in the quantum dot can be explained by the tunneling of electrons in the reversion layer to two layers of the quantum dot in the gate stack. These charges are confined in the channel and in different quantum dot layers depending on the tunneling probability of the charges in the channel (quantum well) and the quantum dot layers [5], [6].

The charge in a QD layer is calculated by determining the tunneling proportion. The tunneling transition rate from the channel to the quantum dot layers is expressed by equation (8) [19], [20], [21].

$$P_{w\to d} = \frac{4\pi}{\hbar}\sum_{w,d}|\langle\psi_d|H - H_w|\psi_w\rangle|^2\,(f_w - f_d)\delta(E_d - E_w) \qquad (8)$$

Where, ψ_d and ψ_w are the wave functions in the quantum dot and inversion channel, f_w and $f_d$are the Fermi distribution functions, E_d and $E_w$are the energy levels in the inversion channel and in the quantum dot layers, $\hbar$ is the reduced Planck's constant, $\hat{H}$is the Hamiltonian operator defined by:

$$\hat{H} = -\frac{\hbar^2}{2}\nabla\left(\frac{1}{m^*}\nabla\right) + V_w(z) + V_d(r) \qquad (9)$$

The electron distribution in the inversion channel can be calculated by Schrödinger-Poisson equations as expressed in equation (10), [16], [20].

$$\frac{\hbar^2}{2}\nabla\left(\frac{1}{m^*}\nabla\Psi_W\right) + (E_n - V)\Psi_W = 0 \qquad (10)$$

Where, m^* is the effective mass of the electron, ψ_W and E_n used to define the bound states, and Vdetermined by the band offsets at the interfaces in combination with the electrostatic potential:

$$\nabla(\varepsilon\nabla\varphi) = q\left(n_{QM} + n - p + N_D^+ + N_A^-\right) \qquad (11)$$

Here, $q$is the charge of the electron, ε is the permittivity of Si, φ is the electrostatic potential, n and p are the electron and hole concentrations, N_D^+ and N_A^- are the ionized donor and acceptor concentrations, $n_{QM}$is the electron gas in the inversion channel as expressed in equation (12):

$$n_{QM} = \sum_n \frac{m^*}{\pi\hbar^2}\Theta(E_F - E_n)\,ln\left[1 + exp\left(\frac{E_f - E_n}{kT}\right)|\psi_n|^2\right] \qquad (12)$$

Where, E_f is the Fermi level, Θ is the Heaviside step function, E_n is the eigen energy and $\psi_n$is the wave function of bound states, determined by the Schrödinger equation (12).

2.3 Current - Voltage characteristics

The energy band diagram of QDG-FET is shown in the Figure 4, the total of voltages across the gate stack is determined by [22]:

$$V_{GS} = V_{FB} + V_{QD} + V_{ox} + \varphi_s \qquad (13)$$

Where, φ_s is the surface potentials of semiconductor, V_{ox} is the voltage across the oxide layer, V_{QD} is the voltage across the quantum dot, V_{FB} is the flat-band voltage.

The work-function Φ_{MS} difference between the gate electrode and the semiconductor is defined by:

$$\Phi_{MS} = \Phi_M - \Phi_S = \Phi_M - \chi - \frac{E_g}{2q} - \varphi_t\,ln\left(\frac{N_{SUB}}{n_i}\right) \qquad (14)$$

Where, E_g is the semiconductor energy gap, χ the electron affinity ($q\chi = E_v - E_c$), q is the electron charge, φ_t the thermal potential, N_{SUB} is the substrate doping concentration and n_i is the intrinsic doping concentration (n ≈ 10^{10} cm^{-3} at T = 300 K).

The thermal potential is $\varphi_t = ln\frac{kT}{q}$; the bulk potential is $\varphi_F = \varphi_t\, ln\left(\frac{N_A}{n_i}\right)$. The capacitance on the oxide layer C_{ox} and the capacitance of the quantum dot C_{QD} are determined by:

$$C_{ox} = \frac{\varepsilon_o \varepsilon_{ox}}{T_{ox}} A;\ C_{QD} = \frac{Q_{QD}}{V_{QD}} \tag{15}$$

The total capacitance in gate stack is determined as:

$$C_{tt} = \left(\frac{1}{C_{QD}} + \frac{1}{C_{ox}}\right)^{-1} \tag{16}$$

Where, T_{ox} is the oxide layer thickness, Q_{QD} is the charge of the quantum dot layer, V_{QD} is the voltage in the quantum dot layer, and A is the contact surface area ($A = WL$).

Therefore, the gate-source voltage with quantum effect is given by [22]:

$$V_{GS} = V_{FB} + \varphi_s \pm \gamma \left[\varphi_t e^{-\frac{\varphi_s}{\varphi_t}} + \varphi_s - \varphi_t + e^{-\frac{2\varphi_{QD}}{\varphi_t}}\left(\varphi_t e^{-\frac{\varphi_s}{\varphi_t}} - \varphi_s - \varphi_t\right)\right]^{1/2} \tag{17}$$

$$\gamma \equiv \frac{\sqrt{2q\varepsilon_o \varepsilon_s N_{SUB}}}{C_{tt}} \tag{18}$$

Due to the tunneling effect, the charges from the conduction channel transfer to the quantum dot layer changing the threshold voltage. Using equations for conventional MOS device, the effect threshold voltage is expressed following:

$$V_T = V_{FB} + \frac{1}{C_{tt}} qN_A \sqrt{\frac{2\varepsilon\varepsilon_0(2\psi_B + V_x}{qN_A}} + 2\psi_B \tag{19}$$

$$\psi_B = \frac{kT}{q} ln\frac{N_A}{n_i}$$

When a gate voltage V_{GS} is applied to the QDG-FET, because of the tunneling effect of electrons through the layers of the quantum dot, the threshold voltage V_T will be varied between V_{GS1} and V_{GS2}. In addition, it depends on the thickness of the oxide layer, the size and number of quantum dots in the gate stack. This dependence coefficient is characterized by the parameter k. When $k = 0$, QDG-FET behaves as a classical MOSFET. When $k = 1$, the threshold voltage will be changed based on the gate voltage V_{GS} [23].

$$V_T = \begin{cases} V_T \\ k(V_{GS} - V_{G1}) + V_T \\ k(V_{G2} - V_{G1}) + V_T \end{cases} \quad \text{for} \quad \begin{cases} V_{GS} < V_{G1} \\ V_{G1} < V_{GS} < V_{G2} \\ V_{GS} > V_{G2} \end{cases} \tag{20}$$

Assuming that electron mobility is constant, μ_n is the electron mobility. Using the simple charge control model, the absolute value of the electron velocity is given by:

$$V_x = \mu_n \frac{dV}{dx} \tag{21}$$

The charge in the channel Q_x is determined as:

$$Q_x = -C_{tt}[V_G - V_T - V_x] \tag{22}$$

Therefore, the total current flowing in the channel from drain D to source S is given by [22]:

$$I_{DS} = -WLQ_x V = LW\mu_n C_{tt}(V_{GS} - V_T - V_x)\frac{dV}{dx} \tag{23}$$

$$I_{DS}dx = WL\mu_n C_{tt}(V_{GS} - V_T - V_x)dV \tag{24}$$

$$I_{DS}\int_0^L dx = WL\mu_n C_{tt}\int_0^{V_{DS}}(V_{GS} - V_T - V_x)dV \tag{25}$$

Where, μ_n is the electron mobility, q is the charge of carriers, L and W is the length and width of the channel, respectively.

Thus, from equations (20) and (25) we can represent the current - voltage characteristics in the QDG-FET by the mathematical equation following:

$$I_{DS} = 0 \quad \text{for} \quad V_{DS} < V_T, \tag{26}$$

$$I_{DS} = \frac{W}{L}\mu_n C_{tt}\left[\frac{1}{V_{DS}}(V_{GS} - V_T) - \frac{V_{DS}^2}{2}\right] \quad \text{for} \quad V_{DS} < V_{GS} - V_T, \tag{27}$$

and

$$I_{DS} = \frac{W}{L}\mu_n C_{tt}\left[\frac{(V_{GS} - V_T)^2}{2}\right] \quad \text{for} \quad V_{DS} > V_{GS} - V_T, \tag{28}$$

From equation (26), (27), (28), it is clear that the voltage-current characteristics in the device can be divided into three basic regions corresponding to three logic levels: logic 0 at 0 V, logic 1 at $V_{DS}/2$ and logic level 2 at V_{DS}. Therefore, QDG-FET is considered as a memory element and applied in the fabrication technology of nonvolatile memories at nanometer size.

3 SIMULATION RESULTS AND DISCUSSION

As shown in the equation (25), the current-voltage characteristic of QDG-FET depends on many different parameters such as the size of the device, the thickness of the material layers in the gate stack, the doping concentration of the semiconductor and electric field in the channel etc., These parameters are changed, the current - voltage characteristics on the component are also changed [24]. In order to simulate these characteristics of the device, the Matlab software was used. Here, QDG-FET is described based on 45 nm technology with typical parameters as shown in Table 1.

Table 1. Parameters used to simulate the QDG-FET

Parameters	Descriptions	Values used in simulation
L	Channel length	45 nm
W	Channel width	45 nm
ε	Effective dielectric constant	8
ε_{ox}	Dielectric constant of insulator	3,9
H	SiOx barrier height	3,25 eV
D	Si QD diameter	4 nm
T_1	Thickness of QD1	9 nm
T_2	Thickness of QD2	3 nm
T_{ox}	Thickness of insulator layer	2 nm
T	Total gate layer	14 nm
C_{ox}	Capacitance of insulator	15.10^{-6} C/cm^2
C_{QD}	Capacitance of QD gate	4.10^{-7} C/cm^2
V_{GS}	Gate-source voltage	0 ÷ 0.8 V
V_{DS}	Drain-source voltage	0 ÷ 0.8 V
μ	Electron mobility	600 cm^2/V.s
N_A	Substrate doping concentration (p-type)	10^{16} cm^{-3}

3.1 The V_{GS}, G_{DS} and VF characteristics

Figure 5 illustrates the relationship between gate voltage V_{GS}, the electric field E_{QD} and the velocity of electrons. It indicates that the more the V_{GS} gate voltage is increased, the more electric field E_{QD} in the channel also increases, the resistance in the channel decrease leading to the velocity of electrons increase. At gate voltage V_{GS} = 0.8 V, E_{QD1} is equal to 10 x10^{-5} (kV/cm^2), E_{Q2} is about 8.1 x 10^{-5} (kV/cm^{-2}). However, the change of the electric field is non-linear, it is not only influenced by the gate voltage V_{GS}, but also depended on the change of energy levels in the quantum dot and the tunneling of charges through the quantum dot in the gate stack.

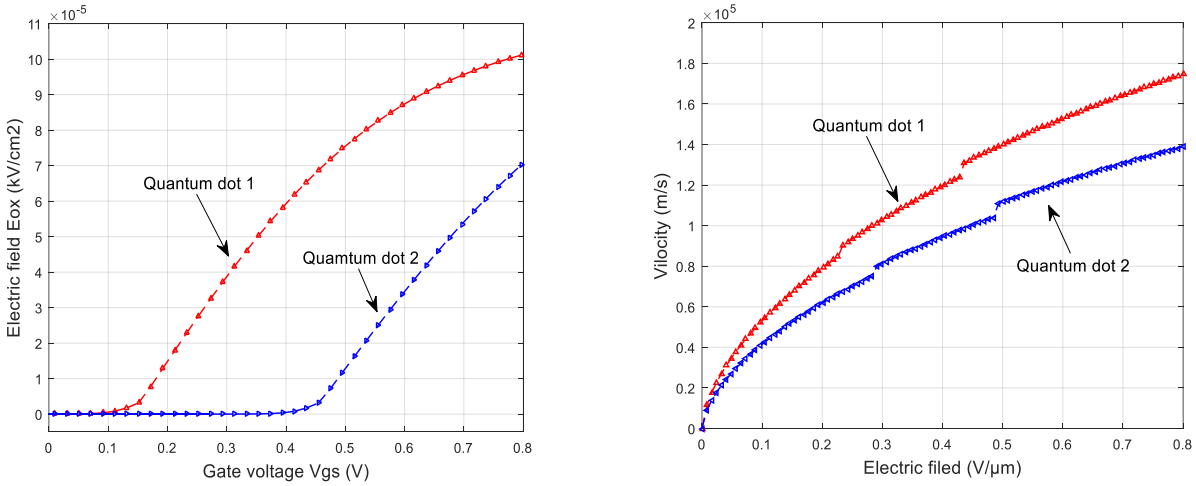


Figure 5. The EDS - VGS and VQD - VGS cures of QDG - FET

3.2 The IDS -VDS and IDS -VGS characteristics

As shown in equation (26), (27), (28), the drain - source current I_D in the channel of the QDG-FET depends on both V_{GS} and V_{DS}. When the gate voltage V_{GS} is constant, if the drain - source voltage V_{DS} increase, the drain - source current is nonlinear increase. It is clearly that when V_{DS} is less than V_T, the I_{DS} increase as exponential function laws. On the contrary, when V_{DS} is higher than V_{GS} - V_T, the I_{DS} reaches saturation value. When the V_{DS} is fixed and the V_{GS} is changed, the I_{DS} is also affected. At gate voltage V_{GS} = 0.8 V, the I_{DS} curve has two slopes at the potential barrier of the first and the second quantum dot layers, about V_{GS} = 0.39 V and V_{GS} = 0.58 V, respectively.

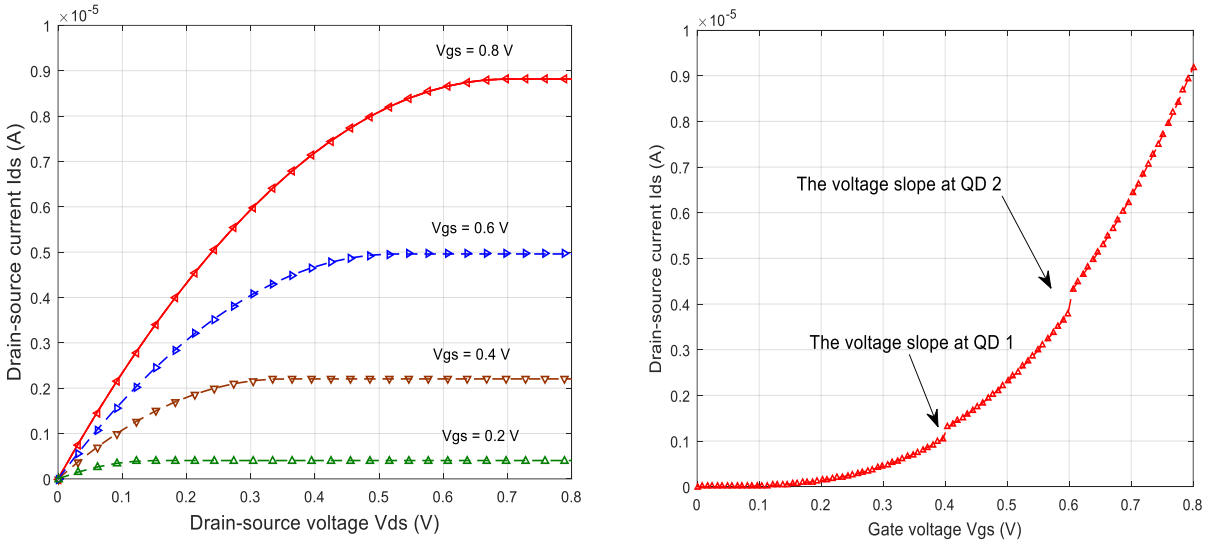


Figure 6. The IDS - V_{DS} and IDS - V_{GS} characteristics of QDG- FET

3.3 The effects of Oxide layer thicknesses

Figure 7 shows the dependence of the I_{DS} on the oxide layer thicknesses T_{ox}. As shown in equation (15), the oxide layer thickness affects to the electric field in the channel. At V_{GS} = 0.8 V, the more the oxide layer thickness increases, the more electric field decreases. However, in the case of the oxide layer is too thin, tunneling effect of charges occurs causing the leaked current in the gate. The I_{DS} curve has two slopes at the potential barrier of the first and the second quantum dot layer, the change of I_{DS} is almost linear with the changes of oxide layer thickness.

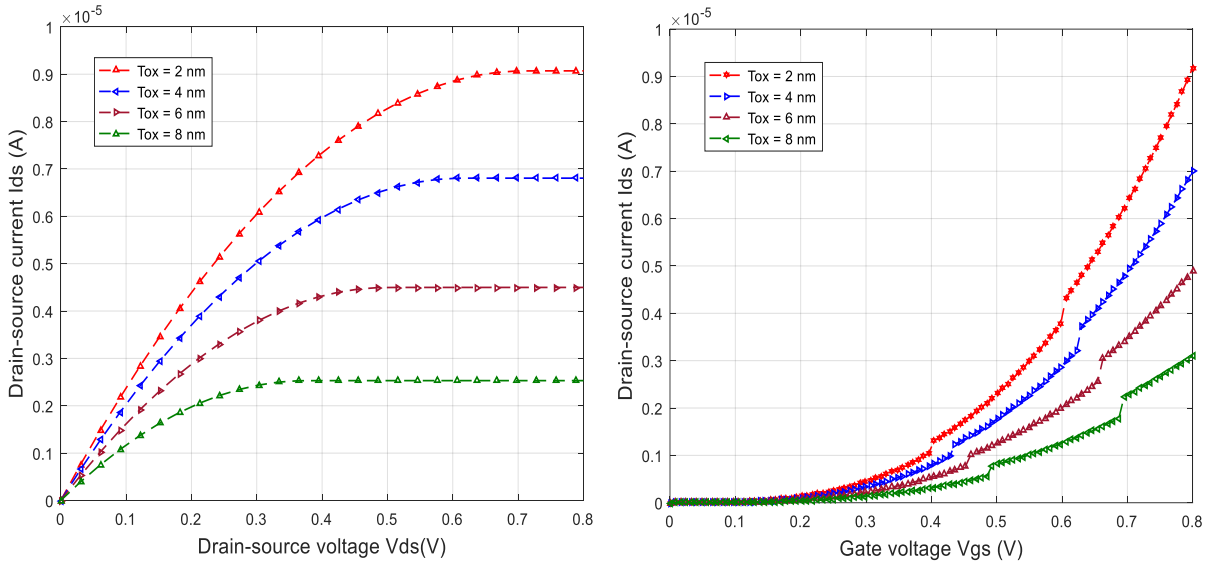


Figure 7. The I_{DS} - V_{DS} and I_{DS} - V_{GS} characteristics with the difference of oxide thicknesses

3.3 The effects of Quantum dot sizes

In the QDG - FET, the quantum dot size also affects on the electric field in the channel. At V_{GS} = 0.8V, the more the quantum dot size increase, the more electric field decreases making the I_{DS} more decreases. The distance between the two slopes of the I_{DS} curve remains constant (ΔV_{GS} = 0.02 V). However, the change of I_{DS} is nonlinear with the change of quantum dot sizes.

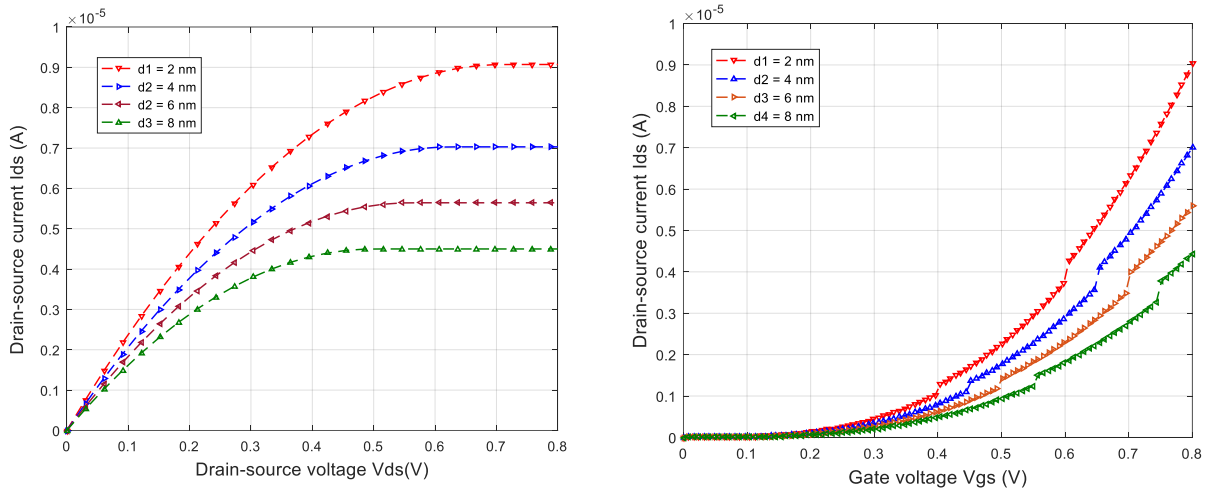


Figure 8. The I_{DS} -V_{DS} and I_{DS} -V_{GS} characteristics with the difference of quantum dot sizes

3.4 The effects of doping concentration in semiconductor

The dependence of the drain-source current I_{DS} on the doping concentration in semiconductor wafer N_{SUB} is shown in Figure 9. From the equation (13), it shows that the doping concentration affects on the conductivity in the channel. When V_{GS} is constant, the more the N_{SUB} increases the more I_{DS} increases. However, the semiconductor doping concentration also depends on the microelectronic fabrication technology. At the current research, N_{SUB} is about 10^{16} cm^{-3} to 10^{23} cm^{-3}.

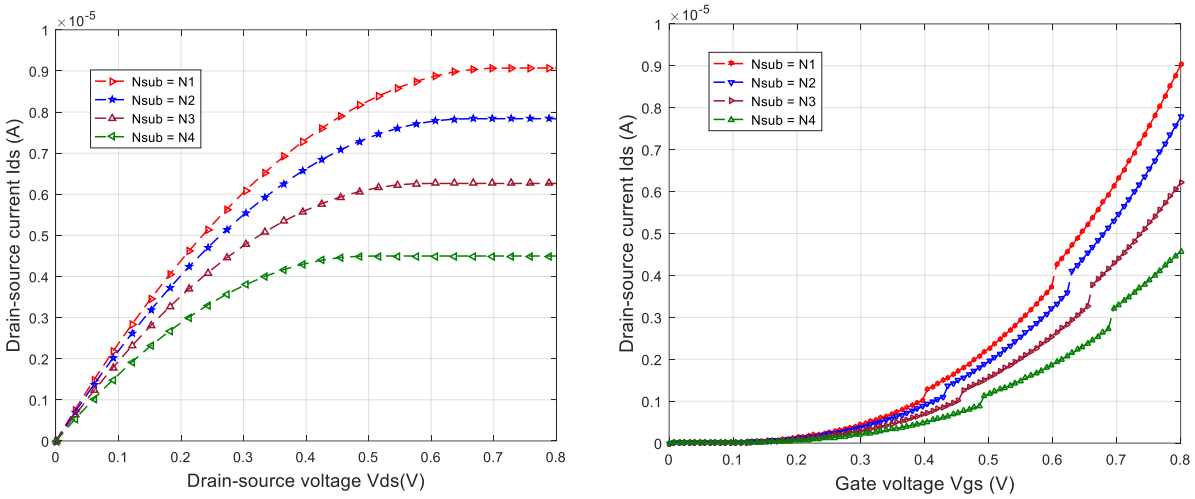


Figure 9. The I_{DS} -V_{DS} and I_{DS} -V_{GS} characteristics of the QDG-FET with the difference of doping concentration, where $N_1 = 10^{23}$ cm^{-3}, $N_2 = 10^{22}$ cm^{-3}, $N_3 = 10^{21}$ cm^{-3}, $N_4 = 10^{20}$ cm^{-3}

4 CONCLUSIONS

The study of the QDG-FET model gives a new approach in calculating the device's current - voltage characteristics. Accordingly, the effects of quantum dots on current - voltage in the device were analyzed and discussed. The current - voltage in the device were calculated by two step process. Firstly, the threshold voltage V_T was determined by using the Schrödinger - Poisson equations to find transmission probability of electrons through the quantum dots due to tunneling of charges. Secondly, the current - voltage characteristics was determined by calculus methods based on MOSFET model. It is shown that these characteristics have non-linear changes at the potential barrier of the first and the second quantum dot layer. Besides, such characteristics depend not only on the gate voltage V_{GS} but also on the material parameters such as channel sizes, doping concentration in semiconductor wafer, gate oxide thickness and quantum dot sizes, etc. In addition, this method can also be applied to calculate current - voltage in other nanoelectronic devices.

REFERENCES

[1] Hongwei Zhu, Zhiping Xu, Ying Fang, etc., "Graphene: Fabrication, Characterizations, Properties and Applications" Published by Elsevier Inc, 2018.

[2] Ali Eatemmadi, Hadis Daraee, Hamzeh Karimkhanloo, etc., "Carbon nanotubes: properties, synthesis, purification, and medical applications", Nanoscale Research Letters volume 9, article number 393, 2014.

[3] Mahdi Asgari, Leonardo, etc., "Semiconductor Nanowire Field-Effect Transistors as Sensitive Detectors in the Far-Infrared", Nanomaterial (MDPI), 2021.

[4] J. Kondo, M. Lingalugari, P. Mirdha, P.-Y. Chan, E. Heller, and F. Jain, "Quantum dot channel (QDC) field effect transistors (FETs) configured as floating gate nonvolatile memories (NVMS)," International Journal of high-Speed Electronics and Systems, 2017.

[5] R. Shankar, r. Velampati, el-sayed hasaneen, e. K. Heller, and faquir c. Jain, "Floating gate nonvolatile memory using individually cladded monodispersed quantum dots." IEEE transactions on very large scale integration (VLSI) systems, January 19, 2017.

[6] F. Jain, F.C. Suarez, E. Gogna, etc., "Novel quantum dot gate FETs and nonvolatile memories using lattice-matched II–VI gate insulators", J. Electron. Mater, 38, (8), pp. 1574–1578, 2009.

[7] N. Shaji, C.B. Simmons, etc., "Spin blockade and lifetime-enhanced transport in a few-electron Si/SiGe double quantum dot", Nat. Phys. 4, 540, 2008.

[8] S. Karmakar, etc., "Application of 25 nm quantum dot gate fets to the design of ADC and DAC circuits", International Journal of High Speed Electronics and Systems, 2011.

[9] F. Jain, S. Karmakar, P-Y. Chan, E Suarez, M.Gogna, J Chandy and E. Heller, "Quantum dot channel (QDC) field-effect transistors (FETs) using II-VI barrier layers" Journal of electronic materials, June 2012.

[10] J. Kondo, M. Lingalugari, P.-Y. Chan, E. Heller, and F. Jain, "Modeling and fabrication of quantum dot channel field effect transistors incorporating quantum dot gate", Proceedings of Tech Connect World, Washington, D.C, 2013.

[11] S. Karmakar, S., Chandy, J.A., Jain, et al.,: 'Design of ternary logic combinational circuits based on quantum dot gate FETs', IEEE Trans. Very Large Scale Integr Syst. (TVLSI), 21, (5), pp. 793–806, 2013.

[12] Dinesh Kumar Pandurangan, "Quantum dot aptamers-an emerging technology with wide scope in pharmacy", International Journal of Pharmacy and Pharmaceutical Sciences, 4 (3), pp.24-31, 2012.

[13] Yiming Zhao, "Quantum Dots and Doped Nanocrystals Synthesis", Optical Properties and Bio-applications, Ph.D.Thesis, Universiteit Utrecht, Nederlands, 2013.

[14] Hemant Kumar, Yogesh Kumar, Gopal Rawat, Chandan Kumar, Bratindranath Mukherjee, Bhola Nath Pal, and Satyabrata Jit, "Heating effects of colloidal ZnO quantum dots (QDs) on ZnO QD/CdSe QD/MoOx photodetectors", IEEE Transactions on nanotechnology, 16(6), pp. 1073-1080, 2017.

[15] Tom Hasell, "Synthesis of metal–polymer nanocomposites", Ph.D.Thesis, University of Nottingham, United Kingdom - China- Malaysia, pp. 18-21, 2008.

[16] J. Brandrup, E.H. Immergut, "Polymer Handbook (2 ed.)", New York, Wiley, pp. 240–246, 1966.

[17] R. Velampati, and F. C. Jain, "A Novel nonvolatile memory using Siox cladded Si quantum dots", NSTI Nanotech, Santa Clara, CA, May 20-24, 2007.

[18] S. Karmakar, "Novel Three-state Quantum Dot Gate Field Effect Transistor", Hand book, Springer India, 2014.

[19] E.K. Heller, S.K. Islam, G. Zhao, and F.C. Jain, "Solid-State Electron", 42, 901, 1999.

[20] C. Duke, "Tunneling in Solids". New York, NY, USA: Academic, 1969.

[21] S. L. Chuang and N. Holonyak, Jr., "Efficient quantum well to quantum dot tunneling: Analytical solutions," Appl. Phys. Lett., vol. 80, no. 7, pp. 1270–1272, 2002.

[22] Y. Tsividis, "Operation and Modeling of the MOS Transistor", McGraw-Hill, 1999.

[23] S. Karmakar, S., Suarez, E., Jain, etc., "Three-state Quantum dot gate field-effect transistor in silicon-on-insulator", Published in IET circuits, devices & systems, 2014.

[24] An Nguyen Van, etc., "Application of the Non-equilibrium Green's function in calculating current-voltage characteristics of the Graphene field effect Transistor", International Symposium on Advanced Manufacturing Technology & Applied Energy, Inductrial University of HCM city, 2016.

ICATSD2F.115

ATTENTION MODELS FOR COVID-19 DETECTION BASED ON LUNG ULTRASOUND IMAGES

HOA THANH LE[1,2], THAO DANH NGUYEN[2], LINH NGUYEN[3]

[1]*Ho Chi Minh University of Movie and Stage, Vietnam*

[2]*Industrial University of Ho Chi Minh City, Vietnam*

[3]*Department of Mathematics, University of Idaho, Moscow, ID 83844-1103, United States*

lethanhhoait@gmail.com, ndthao822002@gmail.com, lnguyen@uidaho.edu

Abstract. We implemented an attention mechanism for different models to detect better COVID-19 infection based on lung ultrasound images to improve model accuracy. We used a lung ultrasound dataset containing three categories: COVID-19 infected, pneumonia infected, or healthy lungs. First, we considered two models: VGG and ResNet, as baselines. Second, our two models introduced the convolutional block attention module (CBAM) as an attention mechanism. We found that the models with attention significantly outperformed the corresponding base models in the experimental results.

Keywords. attention mechanism, COVID-19, convolutional block attention module (CBAM), deep learning, lung ultrasound, pneumonia, VGG, ResNet.

1 INTRODUCTION

COVID-19 is a disease caused by a virus called SARS-CoV-2 and was discovered in December 2019 in Wuhan, China. The condition is highly contagious and has quickly spread around the world. In March 2020, the World Health Organization declared it a pandemic [1]. COVID-19 often causes fever, dry cough, and shortness of breath [2], which feel like a cold, flu, or pneumonia. COVID-19 can attack the lungs, respiratory system, and other body parts. Most people with COVID-19 have mild symptoms, but some become severe and even die. Older adults or people with certain underlying medical conditions are at increased risk for severe illness from COVID-19. Given the rapid spread of COVID-19, a quick and accurate diagnosis is essential and urgent. The RT-PCR test is considered a highly accurate method to confirm COVID-19 infection. However, this test is expensive and time-consuming and can take more than 24 hours to get results. In addition, RT-PCR testing is difficult to perform widely, and only a few clinical laboratories are deployed under the accreditation of health authorities. Alternative methods for RT-PCR testing, such as computed tomography (CT) chest [3], chest X-ray (CXR), and lung ultrasound [5,6], all showed potential in diagnosing COVID-19.

Computed tomography can assess the severity and monitor the progression of the disease. However, computed tomography has significant limitations: prohibitive cost, patients must be exposed to radiation, risk of cross-infection requires disinfection and cleaning after large-scale scanning. X-ray is also suggested; although studies have shown that for 636 CXRs from COVID-19 patients, 58.3% were reread as normal, and 41.7% were reread as abnormal [7], the sensitivity is relatively low. Compared with computed tomography and X-ray, ultrasound is a non-invasive, non-radiative, low-cost, ubiquitous, repeatable, and readily available method in most health facilities [8-10]. Also, due to the portability of the ultrasound device, ultrasound does not require patient relocation, reducing the risk of cross-contamination. Basic principles of lung ultrasound: air cannot be seen on ultrasound, and two interfaces create false images with different acoustic impedances. The thorax is an area where air and fluid are mixed. Diseases lungs are based on spoofing image analysis, and all signs start from the pleural line. Most characters are dynamic, and acute chest signs are related to the surface [11]. Lung ultrasound can diagnose chest wall diseases such as trauma, pneumonia, and other respiratory infections. Figure 1 illustrates the standard features for detecting these diseases in lung ultrasound. Line A is horizontal lines showing normal lung surface (red arrow in figure 1(a)), line B are longitudinally enhanced lines, starting from pleural line to screen edge, moving

synchronized with the sliding lung sign, delete lines A [12]. Figure 1(b) image of 2 individual B lines (red arrow) originating from 2 discontinuities on the pleural line (white arrow). Figure 1(c) image of multiple B lines consolidating to form a white lung; the pleural line is interrupted (white arrow). Figure 1(d) image of lung consolidation, inside a bronchial tree containing air (yellow arrow), the pleural line is no longer recognizable.

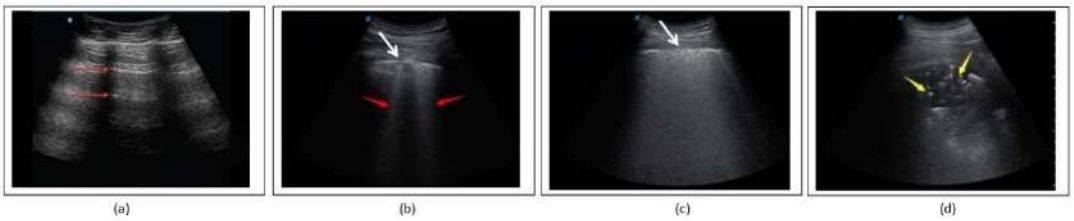


Figure 1. Illustrating the imaging features in lung ultrasound

For COVID-19, the most common abnormality is related to the interstitial syndrome described as B lines (i.e., three or more B lines present in one lung region, confluent B lines, or lung shape white). The presence of several B lines signals the onset of the pneumonia process. As the disease progresses, pneumonia becomes more widespread and more severe, and ultrasound shows more B-lines appearing. They cluster together to form a "white lung" image. Usually, to this degree, the patient requires mechanical ventilation. The most severe degree is pulmonary condensate lesions appearing and spreading, gas exchange function is almost no longer, it is necessary to appoint an extracorporeal oxygen exchange system (ECMO - Extra Corporeal Membrane Oxygenation). Progression improved when the repeated ultrasound showed a decrease in the number of B lines, the reappearance of A-lines, and a regular pleural line.
Our goal in this paper is to focus on implementing an attention mechanism for deep learning techniques to diagnose COVID-19 based on lung ultrasound images. Our work is novel because previous studies only considered basic models without the attention mechanism. The paper is organized as follows. Section 1 introduces the overview, why we choose lung ultrasound images to diagnose COVID-19, and related studies. In section 2, we introduce the image dataset we use for training and testing, and we conduct a 5-fold cross-validation from the image dataset. In section 3, we present the approach based on VGG and ResNet models and introduce the convolutional block attention module (CBAM) and how we integrate it into the two models above. Section 4 compares and presents positive results when combining the CBAM attention channel into the models. Section 5 contains our conclusion.

1.1 Related Works

Deep learning algorithms show impressive accuracy in analyzing medical images, from recognition and object detection to semantic segmentation. The recent literature on exploiting medical image analysis and deep learning to classify lung ultrasound imaging data for COVID-19 patients has gained momentum. Born et al. [13] introduced a lung ultrasound imaging dataset with 202 videos and 59 images, including samples of COVID-19 patients, patients with bacterial pneumonia, and normal health. The authors performed classification based on both frames and videos. For the framework-based classification, they compared NaNET Mobile, VGG-Segment, and Segment-Enc with two VGG16-based architectures named VGG and VGG-CAM. Born et al. [13] compared VGG-CAM and Genesis models for video-based classification. Roberts and Tsiligkaridis [14] also presented their discoveries using deep learning networks CNNs to diagnose COVID-199 based on lung ultrasound images by applying adversarial training. Two networks, VGG16 and ResNet18, were used for training with 3119 frames from 195 ultrasound videos for the training process. The above articles use base deep learning networks to perform the training process. We use the attention mechanism on the base networks and compare them, which the previous authors have not performed.

2 DATASETS

The dataset consists of multiple lung ultrasound images and videos [15]; the videos and images are recorded using a transducer that emits high-frequency sound waves to capture images directly inside the body, which

are reflected and expressed through medical images. Ultrasound images are created based on the reflection of sound waves when touching organs or the body's internal organs. The intensity of this reflected sound wave combined with the time of receiving the reflection will be processed by the computer and shown through images. There are two main types of transducers: linear probes and convex probes. Linear probes: straight, high-frequency, high-resolution transducer, suitable for evaluating superficial areas: such as skin, thyroid, mammary gland, and blood vessels. Convex probes: lower frequency, lower resolution, suitable for evaluating deep organs, often used for ultrasound of the abdomen, fetus, and deep blood vessels.
Data from the convex probes include:
- 182 videos (64x COVID-19, 49x bacterial pneumonia, 66x healthy, 3x viral pneumonia).
- 53 images (18x COVID-19, 20x bacterial pneumonia, 15x healthy).
Data from the linear probes include:
- 20 videos (6x COVID-19, 2x bacterial pneumonia, 9x healthy, 3x viral pneumonia).
- 6 images (4x COVID-19, 2x bacterial pneumonia).
- 45 videos of possible patients with COVID-19 were collected in Piacenza - Italy; There were not enough PCR tests available, so the label is unclear.

- **Train/Validation/Test Split**

The dataset consists of images (obtained from video) labeled into the following three categories:
- COVID-19 infected
- Pneumonia infected
- Healthy

The dataset will be split into five folds for cross-validation. Frames taken from the video will only be included in a single fold to ensure that the training and test sets are entirely different.

3 METHODOLOGIES

3.1 VGG-Style Base Network

VGG is a typical convolutional neural network model, considered one of the best model architectures until now. The model proposed by Karen Simonyan and Andrew Zisserman, VGG, is an acronym for the group name: Visual Geometry Group from the University of Oxford in 2014 in the paper "Very Deep Convolutional Networks for Large-Scale Image Recognition.". All configurations of VGG have a block structure. Each VGG block consists of a sequence of convolutional layers followed by a max pooling layer. The same kernel size (3×3) is applied to all convolutional layers. The VGG network can be divided into two parts: the first consists mainly of convolutional and pooling layers, and the second consists of fully connected layers. VGG16 includes 13 convolutional layers, 5 max-pooling layers, and 3 fully connected layers. Thus, the number of layers with tunable parameters is 16 (13 convolutional layers and 3 fully connected layers). That is why the model's name is VGG16. The input image has a fixed size of 224x224, two convolutional layers of the first block (see conv1 figure 2), each with 64 kernels of size 3x3, followed by a max pooling layer of size 2x2. Two convolutional layers of the next block (see conv2 figure 2), each with 128 kernels of size 3x3, followed by a max pooling layer of size 2x2, the same as the previous one. Three convolutional layers of the next block (see conv3 figure 2), each with 256 Kernels of size 3x3, followed by a 2x2 max pooling layer. Subsequent, three convolutional layers of the next block (see conv4 figure 2), each with 512 kernels of size 3x3, followed by a max pooling layer of 2x2. The last block (see conv5 figure 2) has three convolution layers, each with 512 kernels of size 3x3 and followed by a max pooling layer of size 2x2. Finally, this model is completed by two fully connected hidden layers (see fc6, fc7 figure 2) and an output layer (see fc8 figure 2). The two fully connected layers have the same number of neurons as 4096. The output layer consists of 1000 neurons corresponding to the type of number of the Imagenet [16] dataset.
VGG19 is a variant of the VGG model, consisting of 19 layers (16 convolutional layers, 3 fully connected layers). VGG19 architecture is like VGG16; the input image is 224x224, using a 3x3 kernel, using a max-pooling layer of size 2x2 but with one more convolution layer in the last three convolution blocks. VGG19 also has 3 layers of fully connected. Out of the three layers, the first two layers have 4096 neurons each, and the third layer has 1000 neurons corresponding to the type of number of the Imagenet dataset.

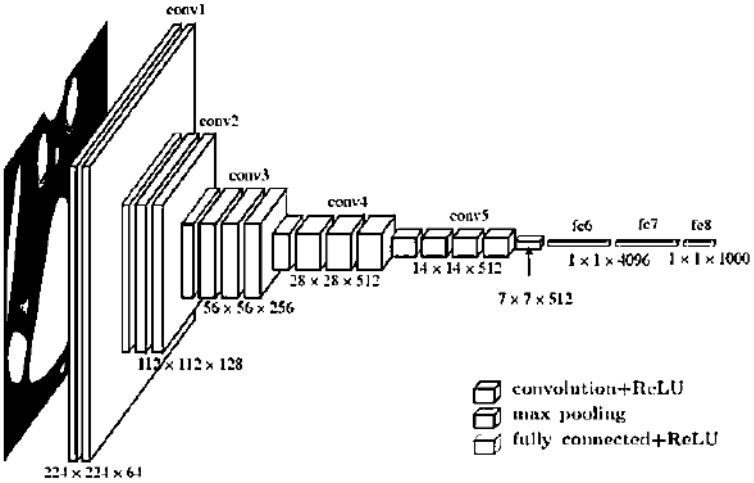


Figure 2. The architecture of VGG16

3.2 RESNET – Style base network

ResNet was introduced to the public in 2015. It won the ImageNet ILSVRC 2015 with an error rate of 3.57%. ResNet has a structure like VGG with many stacking layers, making the model deeper, and also a convolutional neural network consisting of many small residual blocks (see Figure 3). The main idea is to skip one or more layers by adding connections to the previous layer. The idea of residual block is that when input x passes through some layers, we get F(x) then add x to H(x) = F(x) + x. The model is easier to learn by adding features from the previous layer. Currently, there are many variants of ResNet architecture with different number of layers such as ResNet18, ResNet34, ResNet50, ResNet101, ResNet152, etc. With the name ResNet followed by a number only ResNet architecture with a certain number of layers.

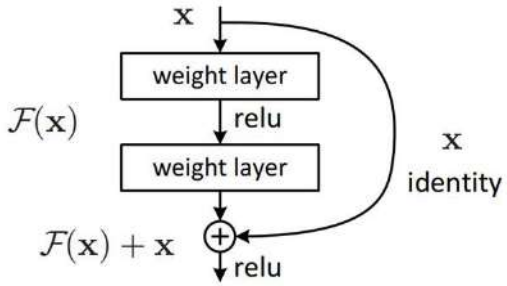


Figure 3. Residual block

ResNet18 is a convolutional neural network model consisting of 18 layers (there are four convolution layers in each layer block (see layer1, layer2, layer3, layer4 in figure 4)), the first convolution layer (with kernel size 7x7), and the last fully connected layer). The network has an image input size of 224x224.

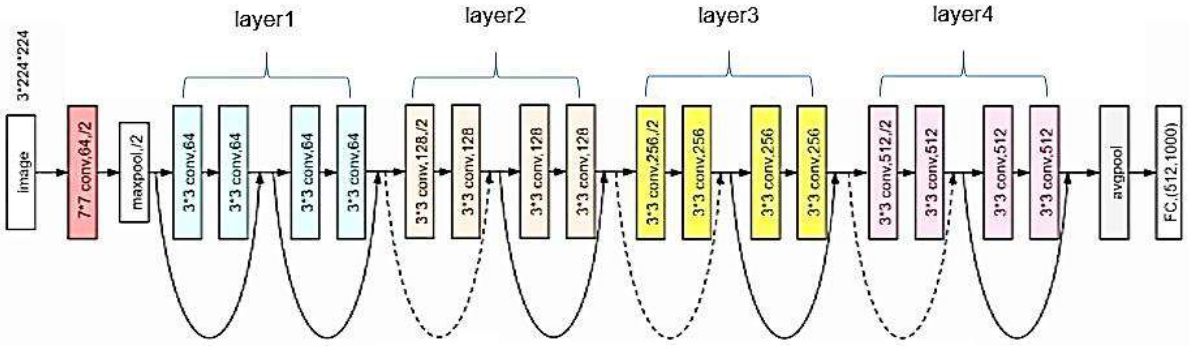


Figure 4. The architecture of ResNet18

3.3 CONVOLUTIONAL BLOCK ATTENTION MODULE (CBAM)

The attention module is used to make the convolutional neural network model learn and focus more on important information instead of learning basic information that is not useful. In the case of object detection, the useful information is the target objects or classes we want to classify and localize in the image. Recently, Woo et al. [17] proposed a convolution block attention module, which mainly features in the attention mechanism. Attention not only tells the neural network model where to focus, it also improves the representation of interests. Before diving into CBAM details, let's take a look at the general structure of the module (see figure 5):

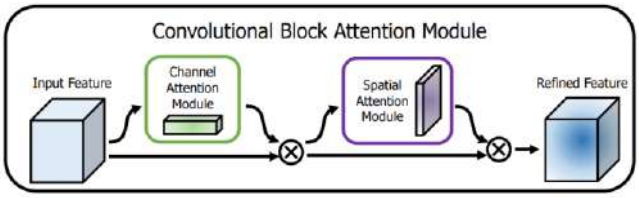


Figure 5. Convolutional Block Attention Module layout

CBAM contains two successional submodules called Channel Attention Module (CAM) and Spatial Attention Module (SAM), which are applied in that order. Channel attention focuses on 'what' is meaningful given an input image, and spatial attention focuses on 'where', an informative part, which is complementary to channel attention. The authors show that CBAM is applied at every complication block in deep networks to get a subsequent "Refined Feature Map" from the "Input Feature Map." The overall attention process can be epitomized as follows, see equation (1):

$$\begin{aligned} \mathbf{F}' &= \mathbf{M_c}(\mathbf{F}) \otimes \mathbf{F}, \\ \mathbf{F}'' &= \mathbf{M_s}(\mathbf{F}') \otimes \mathbf{F}', \end{aligned} \tag{1}$$

Where $\otimes$ stands for element-wise multiplication? If the two operands are not in the same dimension, the valuations are broadcasted (copied) similarly. The spatial attentional valuations circulate along the channel dimension, and the channel attention values are broadcasted along the spatial dimension. The following describes an overview of each attention module.

First, we define the channel attention module. Both max pooling and average pooling are used to generate two different spatial context descriptors: F^c_{avg} and F^c_{max}, representing the max pooling features and average pooling features. Both are then forwarded to a shared multi-layer perception (MLP) to generate the output features, which are then merged using element-wise summation $\oplus$. The merged sum is finally sent to the sigmoid function σ. To reduce the parameter resources, the hidden size of MLP is set to $\mathbb{R}^{C/r \text{ x1x1}}$, where r is the reduction ratio. Suppose $W_0 \in \mathbb{R}^{C/r \text{ x } C}$ and $W_1 \in \mathbb{R}^{C \text{ x } C/r}$ stand for the MLP weights, respectively, see equation (2):

$$\begin{aligned} \mathbf{M_c}(\mathbf{F}) &= \sigma(MLP(AvgPool(\mathbf{F})) + MLP(MaxPool(\mathbf{F}))) \\ &= \sigma(\mathbf{W_1}(\mathbf{W_0}(\mathbf{F^c_{avg}})) + \mathbf{W_1}(\mathbf{W_0}(\mathbf{F^c_{max}}))), \end{aligned} \tag{2}$$

Note that W_0 and $\mathbf{W_1}$ are shared by both max pooling (MaxPool) and average pooling (AvgPool). Figure 6 shows the flowchart of the channel attention module. Note that the squeeze-and-excitation (SE) [18] method is like the channel attention module.

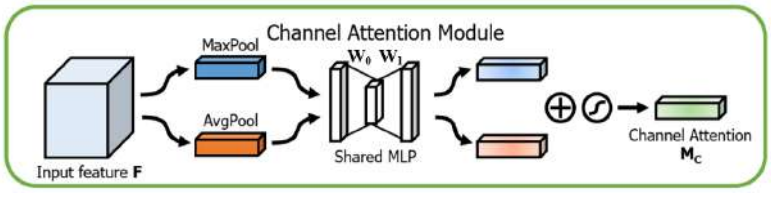


Figure 6. Flowchart of the channel attention module

Second, we describe the detailed procedures of the spatial attention module. The spatial attention module is a complementary step to the previous channel attention module. The average pooling and max pooling are applied to the channel-refined activation map $F^{/}$. Both average pooling and max pooling are two-

dimensional activation maps: $F^s_{avg} \in \mathbb{R}^{1xHxW}$ and $F^s_{max} \in \mathbb{R}^{1xHxW}$. The concatenated activation map is then passed into a standard 7×7 convolution f^{7x7} and followed by a sigmoid function σ. In all, we ha tove see equation (3):

$$\begin{aligned} \mathbf{M_s}(\mathbf{F}) &= \sigma(f^{7\times7}([AvgPool(\mathbf{F}); MaxPool(\mathbf{F})])) \\ &= \sigma(f^{7\times7}([\mathbf{F^s_{avg}}; \mathbf{F^s_{max}}])), \end{aligned} \tag{3}$$

The flowchart of the spatial attention module is drawn in Figure 7.

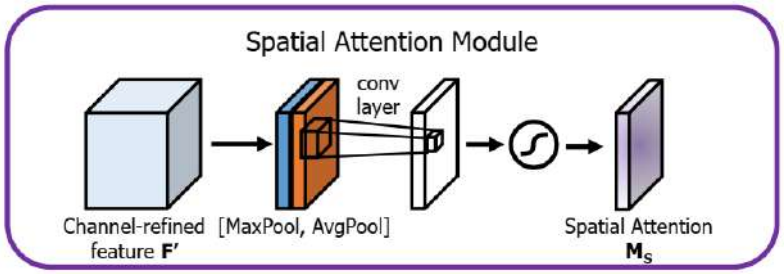


Figure 7. Flowchart of the spatial attention module

The attention mechanism CBAM is embedded into our VGG-style base network. The integration is shown in Figure 8. The two consecutive attention modules (channel and spatial) are added for the activation map F before the last convolutional layer. The refined features are sent to the final convolution block (CB). Here AACB means the attention attached convolution block, composed of one convolution block and the following attention modules.

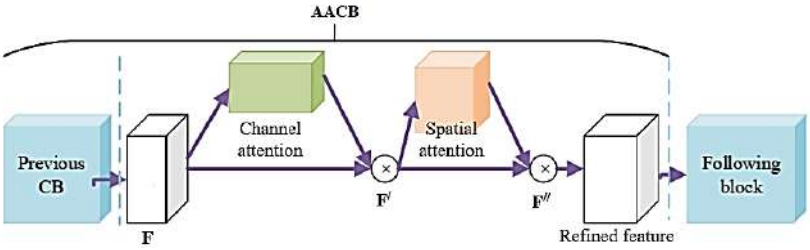


Figure 8. Illustration of AACB: Integration of CBAM with VGG-style base network

For the ResNet-style base network, the CBAM integration is shown in Figure 9. This figure shows the accurate position of our module when absorbed within a ResBlock. We refer to CBAM on the convolution outputs in each block.

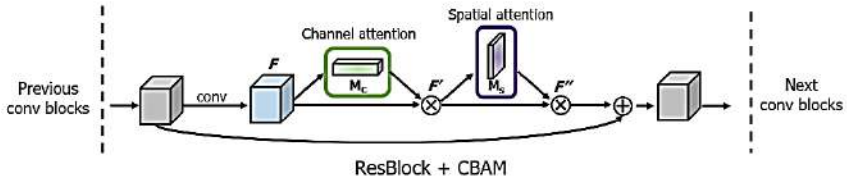


Figure 9. Illustration of AACB: Integration of CBAM with ResNet-style base network

4 RESULTS

This study implements the Windows 11 operating system model, Intel i5-11400F 6*core, 4.40GHz CPU, 32GB memory, and NVIDIA GTX 3060 12GB VRAM platform. We use Python programming language and use deep learning framework PyTorch.

We got the final performance by the max metrics of 5-fold. We used the metrics of sensitivity, specificity, precision, F1 score, and accuracy to evaluate the performance of our method. The training process terminates at 20 epochs, and the batch size is 20. The Adam optimizer is used in each epoch, and the learning rate is 1e-4.

We compare different configurations of attention here. From Table 1, we can see that the base VGG16 method achieves a lower accuracy of 0.87 than the VGG16-CBAM method with an accuracy of 0.93. For the base method, VGG19 also gives lower accuracy of 0.91 results than VGG19-CBAM accuracy of 0.93. Similarly, the base ResNet18 method also provides an accuracy of 0.87, lower than the ResNet18-CBAM accuracy of 0.93. It indicates that the model with integrated CBAM will achieve the effect with higher accuracy. In addition, we also present the results using the confusion matrix and Receiver Operating Characteristic (ROC), see Figures 10 and 11.

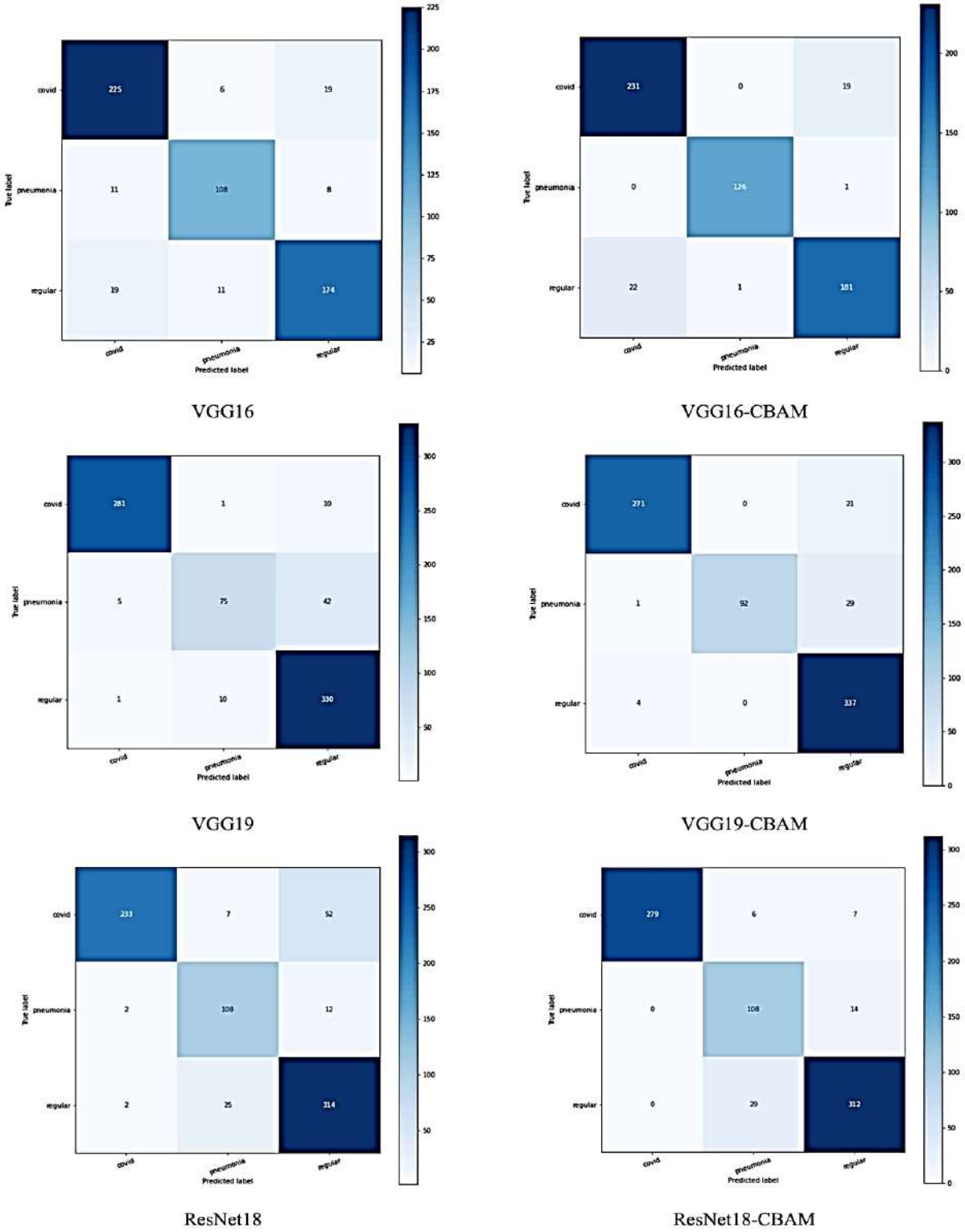


Figure 10. Confusion matrix for the based model and CBAM integration

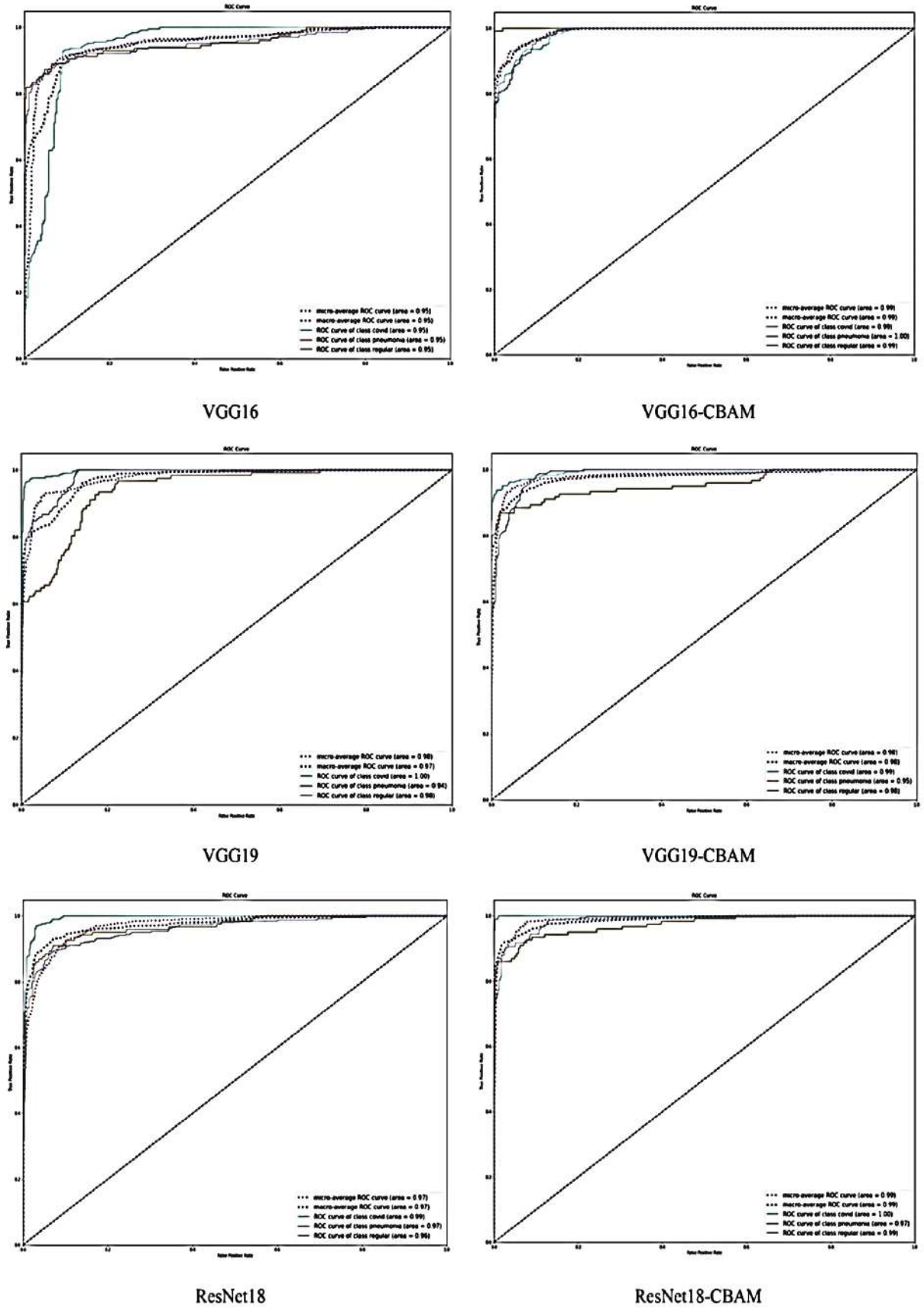


Figure 11: ROC Curves for the based model and CBAM integration

Table 1. Comparison of different configurations of attention

	VGG16			VGG16-CBAM		
	Precision	Recall	F1-score	Precision	Recall	F1-score
covid	0.88	0.90	0.89	0.91	0.92	0.92
pneumonia	0.86	0.85	0.86	0.99	0.99	0.99
regular	0.87	0.85	0.86	0.90	0.89	0.89
accuracy	0.87			**0.93**		
	VGG19			VGG19-CBAM		
covid	0.98	0.96	0.97	0.98	0.93	0.95
pneumonia	0.87	0.61	0.72	1.00	0.75	0.86
regular	0.86	0.97	0.91	0.87	0.99	0.93
accuracy	0.91			**0.93**		
	ResNet18			ResNet18-CBAM		
covid	0.98	0.80	0.88	1.00	0.96	0.98
pneumonia	0.77	0.89	0.82	0.76	0.89	0.82
regular	0.83	0.92	0.87	0.94	0.91	0.93
accuracy	0.87			**0.93**		

5 CONCLUSIONS

This study presents a COVID-19 detection method based on attention and transfer learning mechanisms. Experimental results show that in the condition of small sample size, our model shows the highest performance for the classification task.

However, some limitations of this study remain and will be addressed in future work. The interpretation of the proposed model is complex, and the reason for this precise classification result is unknown. Besides, the dataset we use is small.

We plan to visualize every phase of the model in the future, further exploring how the model works. At the same time, we will extend our model to other datasets and improve performance based on the results. In addition, we will collect more COVID-19 images and further improve our model to detect more specific diseases.

ACKNOWLEDGMENT

Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number P20GM104420. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES

[1] WHO, "WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020", https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-COVID-19---11-march-2020

[2] W. Guan, Z. Y. Ni, Y. Hu et al., "Clinical characteristics of coronavirus disease 2019 in China," New England Journal of Medicine, vol. 382, no. 18, pp. 1708–1720, 2020

[3] A. Bernheim, X. Mei, M. Huang, et al., "Chest CT findings in coronavirus disease-19 (COVID-19): relationship to duration of infection," Radiology, vol. 295, no. 3, pp. 200463–200463, 2020.

[4] M.-Y. Ng, E. Y. P. Lee, J. Yang et al., "Imaging profile of the COVID-19 infection: radiologic findings and literature review," Radiology: Cardiothoracic Imaging, vol. 2, no. 1, article e200034, 2020

[5] N. Buda, E. Segura-Grau, J. Cylwik, and M. Wełnicki, "Lung ultrasound in the diagnosis of COVID-19 infection - a case series and review of the literature," Advances in Medical Sciences, vol. 65, no. 2, pp. 378–385, 2020.

[6] G. Soldati, A. Smargiassi, R. Inchingolo, et al., "Is there a role for lung ultrasound during the COVID-19 pandemic?," Journal of Ultrasound in Medicine, vol. 39, no. 7, pp. 1459–1462, 2020.

[7] M. B. Weinstock, A. N. Echenique, J. W. Russell, et al., "Chest X-ray findings in 636 ambulatory patients with COVID-19 presenting to an urgent care center: a normal chest X-ray is no guarantee," The Journal of Urgent Care Medicine, vol. 14, no. 7, pp. 13–18, 2020

[8] N. Xirouchaki, E. Magkanas, K. Vaporidi, et al., "Lung ultrasound in critically ill patients: comparison with bedside chest radiography," Intensive Care Medicine, vol. 37, no. 9, pp. 1488–1493, 2011.

[9] D. Lichtenstein, I. Goldstein, E. Mourgeon, P. Cluzel, P. Grenier, and J.-J. Rouby, "Comparative diagnostic performances of auscultation, chest radiography, and lung ultrasonography in acute respiratory distress syndrome," The Journal of the American Society of Anesthesiologists, vol. 100, no. 1, pp. 9–15, 2004.

[10] Y. Yang, Y. Huang, F. Gao, L. Yuan, and Z. Wang, "Lung ultrasonography versus chest CT in COVID-19 pneumonia: a two centered retrospective comparison study from China," Intensive Care Medicine, vol. 46, no. 9, pp. 1761–1763, 2020.

[11] Lichtenstein, Daniel A. (2010), "Whole Body Ultrasonography in the Critically ill."

[12] Lichtenstein, Daniel A. (2012), "General ultrasound in the critically ill book," Springer, pp.33

[13] J. Born, N. Wiedemann, M. Cossio, et al., "Accelerating detection of lung pathologies with explainable ultrasound image analysis," Applied Sciences, vol. 11, no. 2, p. 672, 2021.

[14] J. Roberts and T. Tsiligkaridis, "Ultrasound diagnosis of COVID-19: robustness and explainability," 2020, https://arxiv.org/abs/2012.01145.

[15] https://github.com/ar-ambuj23/covid19_pocus_ultrasound_pytorch

[16] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. Imagenet: A large-scale hierarchical image database. In Proc. CVPR, 2009.

[17] S. Woo, J. Park, J.-Y. Lee, and I. So Kweon, "CBAM: Convolutional block attention module," in Proceedings of the European conference on computer vision (ECCV), Munich, Germany, 2018, pp. 3-19.

[18] J. Hu, L. Shen, S. Albanie, G. Sun, and E. H. Wu, "Squeeze-and-Excitation Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, pp. 2011-2023, Aug. 2020

ICATSD2F.116

HAND-WRITTEN CHARACTER RECOGNITION BY WRITING SOUND USING A ONE-DIMENSIONAL CONVOLUTIONAL NEURAL NETWORK

IORI TAJIMA, HIROAKI KUROKAWA
School of Engineering, Tokyo University of Technology,
hkuro@stf.teu.ac.jp

Abstract. The Convolutional Neural Network (CNN) is one of the most successful models in recent Artificial Intelligence and machine learning models. A lot of applications are developed based on CNN. Image processing is a major application of CNN, and the feature extraction layer in CNN has a 2D structure for image processing. In this study, we try to make hand-written character recognition by writing sound using one-dimensional CNN (1DCNN). The target characters to recognize are Japanese characters called Hiragana. For the experiments, we prepare the character writing sound data of 5 Hiragana, 200 data for each character, total of 1000 sound data from the unique subject. Here, all sound data is quantized as one-dimensional data. The result of our trial to classify 5 characters shows an average accuracy of 93%. On the other hand, the result of another trial using data from unlearned subjects shows an average accuracy of 37%. These results of these toy problems suggest that the 1DCNN is sufficiently effective to recognize writing sounds from learned subjects, however, it does not show the expected effect for unlearned subjects. This means that it is difficult to obtain proper results, especially in the aspect of the ability of universal use. However, these unexpected characteristics can be used for personal authentication as a kind of biometrics technique. Therefore, we show the performance of personal authentication using character writing sound by 1DCNN. The result of our test to recognize four subjects using their writing sound shows an average accuracy of 82%. Although improvement of the accuracy is our future challenge, these results suggest the feasibility of personal authentication by writing sounds.

Keywords. convolutional neural network, character writing sounds, classification

1 INTRODUCTION

A neural network is a mathematical model that can be described as a network of artificial neurons and has shown various applications in a very wide range of technical fields. In particular, the multi-layer neural network is essentially the model that can acquire the ability to solve regression problems by machine learning and are extremely powerful in function approximation, multi-class classification, and so on. Multi-layer neural networks had been recognized as a major model in the second artificial intelligence boom around the 1980s. A reason for the attention was the proposal of a machine learning algorithm called backpropagation. Backpropagation is a method that enables self-organization only by presenting a training data set including teacher data and can be considered as a method to configure mathematical models that do not require programming based on expert knowledge and experience.

In the recent decade, after the AI winter, research in the field of artificial intelligence has again experienced a major boom and rapid development. The main trigger of this third AI boom was "Deep Learning", which had been established around 2010[1-5]. Here "Deep Learning" means any machine learning technique applicable to highly multi-layered neural networks, and this technique enables the configuration of more accurate and versatile mathematical models using big data. In other words, the proposal of this new machine learning method has once again brought a breakthrough in artificial intelligence technology.

In recent studies of AI, convolutional neural networks (CNN) have made significant achievements in the field of visual information processing. The CNN used for image classification is similar in structure to the Neocognitron [6] proposed in 1980. The significant difference between CNN and Neocognitron is the method of learning, that is, the back-propagation-based learning algorithm dedicated to Deep learning is applied to CNN. The major application of CNN is visual information processing. However, in general,

CNN provides solutions that integrate feature extraction and classification based on those features for given data and have a variety of possible applications [7-11].
In this study, we show a procedure for hand-written character recognition by extracting features of the writing sounds of the characters. Here, the 1DCNN is used for character recognition. The characters to be recognized are Japanese characters called Hiragana. We also show that the same system can be used to authenticate the writer using the features of writing sounds.
While most studies of audio data classification use methods that classify audio data as image data such as spectrograms, this study simply uses time-series data of quantized sound data. This enables a simple system configuration without any pre-processing. Since time series sound data is one-dimensional, the input layers of the CNN, *i.e.*, the convolution layer and pooling layer, should have a one-dimensional structure. Although the structure of the CNN is changed according to the shape of the data, the learning method can be applied without any changes. In this study, the CNN was implemented using TensorFlow [12].
The rest of the paper is organized as follows. Section 2 describes the structure of the 1D CNN, Section 3 describes character recognition using writing sound data, and Section 4 describes personal authentication using writing sound data.

2 ONE-DIMENSIONAL CONVOLUTIONAL NEURAL NETWORKS

Since CNN is a model proposed for visual information processing and has attracted attention mainly for its application to image classification, many CNNs used in various studies so far have a structure that allows the input of two-dimensional data such as images. The input data is propagated through a convolutional layer and a pooling layer to the affine layer. In the convolution layer, multiple filters generate channels to extract multiple features, and the data processed in each channel becomes the entire feature vector by the pooling layer. The affine layer works as a multi-layer neural network and performs necessary processing based on the given feature vector.
For example, in the case of a multi-class classification problem, neurons corresponding to the category to be classified are placed in the output layer, and the classification result of the input data into each class is determined according to the output of those neurons. The activation function of each neuron is rectified linear unit (ReLU), while the activation function in the output layer is a softmax function to match its purpose.

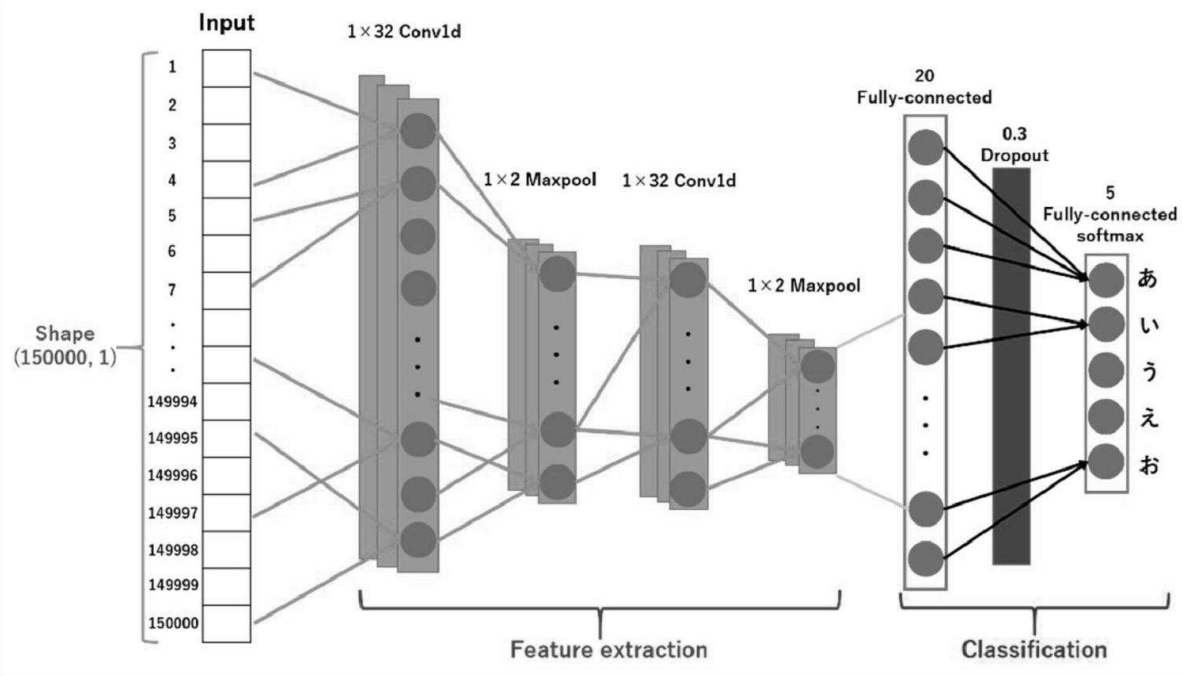


Figure 1. Schematic of one-dimensional convolutional neural network. Configuration for the application to classify the five kinds of Japanese Hiragana characters listed on the right side of the affine layer.

The internal processing of the CNN is essentially the same whether the dimension of the input data is 1D or 2D. The processing in the convolutional and pooling layers is performed by expanding the data into a 2D matrix whose size depends on the number of channels and batch size, so only the size of the 2D matrix changes. In addition, the input data to the affine layer is originally expanded to one dimension and sent to the affine layer, so the dimension does not affect the input data.
Figure 1 shows a schematic diagram of the 1DCNN used in this study. Here, the output layer is configured to identify five characters, and the Japanese Hiragana characters shown on the side of output neurons are the five characters to be classified. The input data is discretized one-dimensional audio data, which is time series data with 150000 elements. The number of filters used in the convolution layer is 32, and the filter size is 1 × 5. The model has two convolution layers with a 1 x 2 max-pooling layer. The affine layer consists of 2 layers with a dropout rate of 0.3. The batch size varied from 5~10 depending on the number of training data. The loss function was determined by categorical cross-entropy loss.

3 CLASSIFICATION OF HIRAGANA CHARACTERS BY WRITING SOUNDS

In this study, we show a method for classifying hand-written characters by writing sounds using a 1DCNN. The following section describes the method of collecting writing sound data and shows the results of classifying five different Hiragana characters. Then, we evaluate the trained model from the viewpoint of versatility.

3.1 Preparation of writing sound data and the character classification results

The writing sound data used in this experiment were collected using a microphone. A piece of paper is placed on a plastic board in a quiet room and the subject wrote the requested letters with a ballpoint pen. Here, there were no strict regulations on the time taken to write or the distance between the microphone and the pen tip. This gives the data randomness, and it ensures that the learning performs well. The sampling frequency was 44.1 kHz, and the quantization bit rate was 16 bits. Figure 2 shows an example of recorded audio data. Each represents the writing sounds of five Hiragana characters. These data were recorded in WAV format. The recorded data were converted to time series data of volume and used as input data to the 1DCNN. The number of elements in the input data was set to 150,000 for all the data, and the parts with no data were filled with zeros.

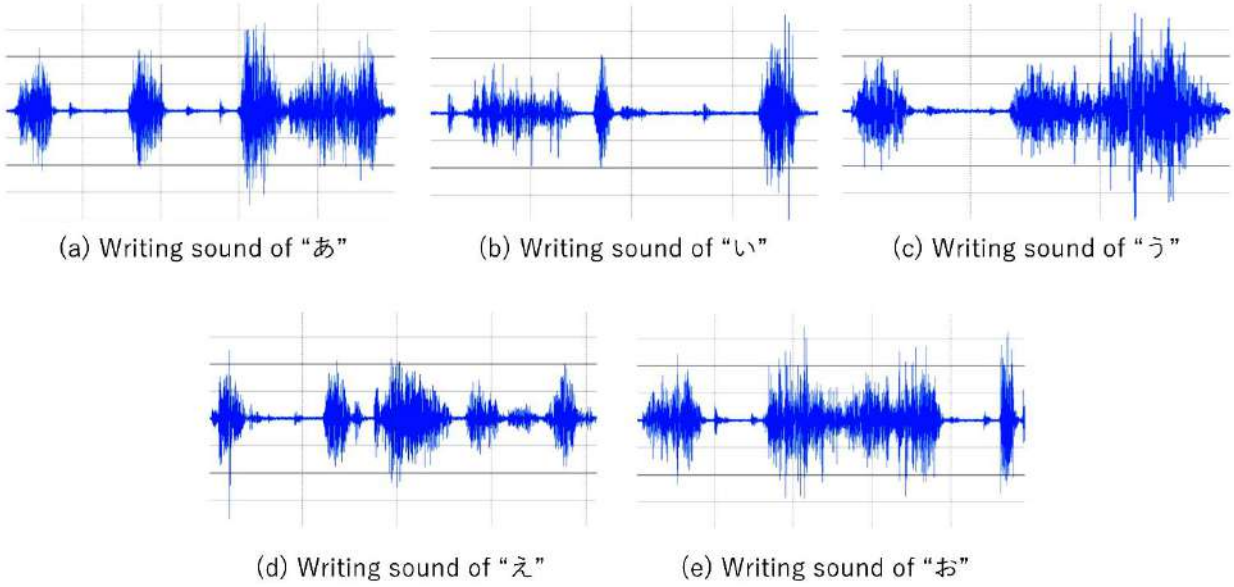


Figure 2. Examples of writing sound data of five Hiragana characters.

Under these conditions, 200 writing sound data were collected per character, for a total of 1000 data for five different characters for the multi-class classification. All the writing sound data are provided by one subject. Here, the configuration of the 1DCNN is as shown in the previous section. In these writing sound

data, 80% was used as training data and 20% as test data to evaluate the training model. And the number of epochs in training was set to 200.
Figure 3 shows the results of the training and validation. From the results shown in Figure 3(a), the trained 1DCNN model was able to classify the training data with almost 100% accuracy and the test data with about 93% accuracy. The confusion matrix in Figure 3(b) indicates the accuracy of classification for each character. Although the classification accuracy differs depending on the characters, it is found to be generally highly accurate (82.5%~100%).

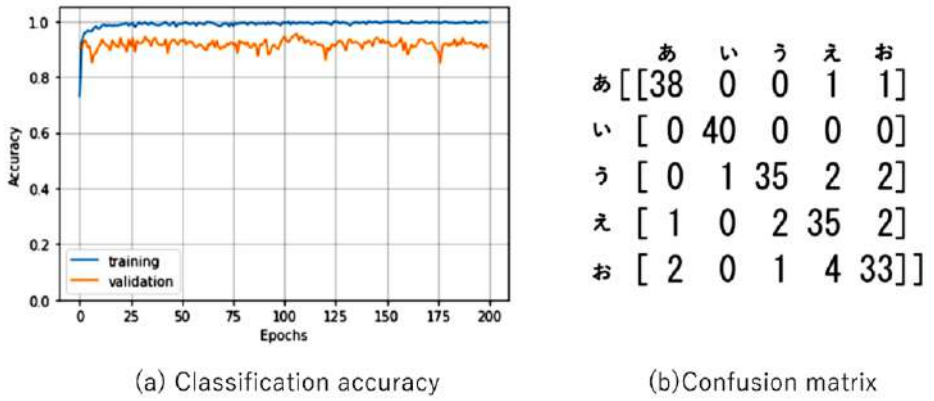


Figure 3. Results of classification of five Hiragana characters. (a) Classification accuracy of training data and test data. (b) Confusion matrix of test results of 5 x 50 test data.

Incidentally, when the same multi-class classification was performed using a simple DNN without a convolutional layer or pooling layer, the classification accuracy becomes 33.5%. The results indicate the superiority of the 1DCNN. Note that the DNN used here has three hidden layers in addition to the input and output layers, and the number of hidden layer neurons is 150, 100, and 50, respectively. With this configuration, the number of training parameters is almost the same as that of the 1DCNN used in the above experiment.

3.2 Evaluation of the versatility

In the experiments in the previous section, only one subject provided all the data, and these data were used to train and test the 1DCNN. Assuming that the goal of this research is to develop a system for character recognition by writing sounds, the accuracy of character classifications independent of the writer would be required. Therefore, to test the versatility of the trained CNN model, we collected data on writing sounds from three new subjects and show the results of a character classification test.
Figure 4 shows the writing sounds of the same letter written by new three different subjects. As shown in the figures, these appear to be distinctly different. Then it is expected that a 1DCNN trained only on the original subjects' data would have difficulty accurately classifying the letters from these new sound data.

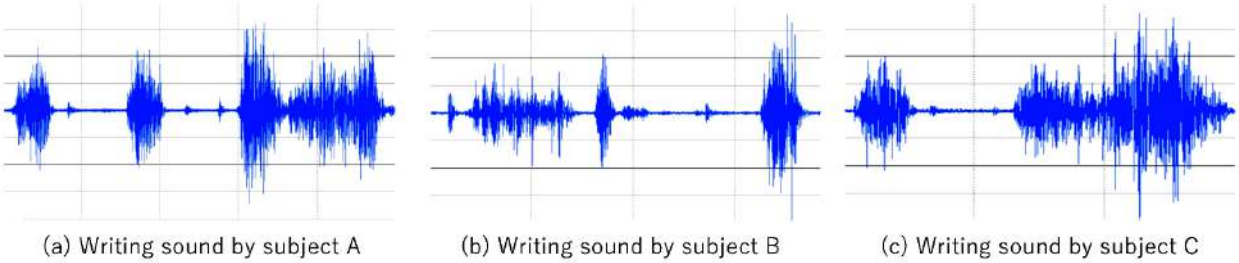
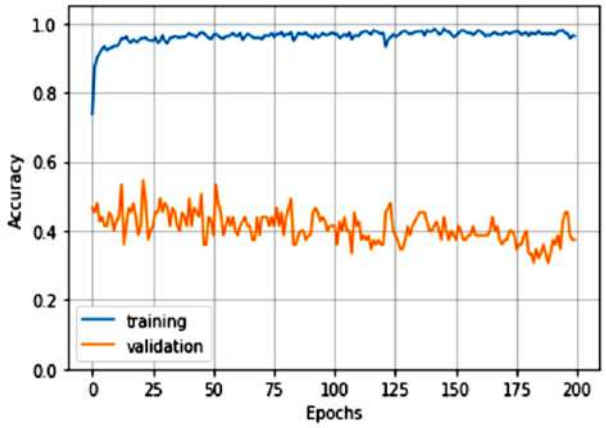


Figure 4. Examples of writing the sound of the same HIRAGANA character by subjects A, B, and C.

Practically, the results of the classification test using the writing sound data provided by three new subjects are shown in Figure 5. In this experiment, 1000 writing sound data from the original subjects were used as training data, and the three new subjects' writing sound data were used as test data to evaluate the trained 1DCNN model. Here the number of training epochs was set to 200.

The number of writing sound data from the new subject was five for each Hiragana character, for a total of 25 data. Thus, a total of 75 writing sound data from new three subjects were used for the test. Figure 5(a) shows that the accuracy of multi-class classification for the new subjects' writing sound data was significantly lower. The confusion matrix in Figure 5(b) also shows that the classification accuracy is entirely lower, although there are differences depending on the kind of letter.



(a) Classification accuracy

	あ	い	う	え	お
あ	4	0	0	1	10
い	0	10	1	1	3
う	1	1	8	3	2
え	0	0	8	4	3
お	0	0	8	5	2

(b)Confusion matrix

Figure 5. Results of classification of unlearned writing sound by other subjects A, B, and C. (a) Classification accuracy of training data and test data of other subjects. (b) Confusion matrix of test results of 5 x 15 test data.

To classify the data shown in Figure 4 as the same class, it is necessary to learn the features that are common to all of them. The CNN and its learning algorithm have the potential to achieve such classification. However, it requires training with huge data sets from many people. Currently, we do not have an opportunity to collect such "Big Data", so solving this problem with this approach is a future challenge.

4 PERSONAL AUTHENTICATION BY CHARACTER WRITING SOUND

In the previous section, we show the classification results of the five Hiragana characters by 1DCNN using writing sound data. The results also suggest that the characteristics of writing sounds extracted by CNN are different depending on the writer, although the same character is written. Therefore, in this section, we discuss personal authentication using writing sounds. Although personal authentication using the writing sound of signatures could be considered for practical applications, in this study, we try to verify whether it is possible to identify a writer by writing sound using a 1DCNN. Here we consider the case where different writers write the same letters.

In this verification, we use 50 new writing sound data per letter from the original subject and three subjects, A, B, and C who provided the writing sound data in the previous section. Therefore, we used a total of 200 data from the four subjects, 120 of which were used for training, and 80 of which were used for a personal authentication test. A similar five tests were performed on five different letters with the 1DCNN shown in Figure 1.

(a) “あ”	(b) “い”	(c) “う”	(d) “え”	(e) “お”
[20 0 0 0]	[19 0 1 0]	[20 0 0 0]	[20 0 0 0]	[19 0 0 1]
[1 12 6 1]	[2 16 1 1]	[0 15 5 0]	[0 19 0 1]	[0 19 1 0]
[1 4 10 5]	[0 1 13 6]	[0 1 19 0]	[0 0 20 0]	[0 0 19 1]
[1 4 5 10]]	[0 6 0 14]]	[0 1 0 19]]	[0 4 4 12]]	[0 4 0 16]]

Figure 6. Results of the personal authentication. A confusion matrix of test results of 4 x 20 test data is shown for five Hiragana characters.

In Figure6, (a)~(e) show the confusion matrices for the personal authentication test for each of the different Hiragana characters. The accuracy was from 65% to 91% depending on the character. This result suggests that it is possible to identify the writer from the writing sound if the character is chosen appropriately. Although the accuracy obtained in this experiment is not at the level of practical use, we believe there is potential for improvement in accuracy due to the small number of data used in the study.

5 CONCLUSIONS

In this study, we show a 1DCNN procedure for multi-class classification by extracting features of the writing sounds of Japanese Hiragana characters. Possibilities of two applications were presented: recognition of characters by writing sounds and personal authentication by writing sounds. First, we showed the results of multi-class classification for five characters, and we obtained an average accuracy of 93%. However, another trial using data from unlearned subjects showed an average accuracy of 37%. Therefore, it suggested the difficulty to obtain adequate results in terms of versatility. Next, based on these results, we tried personal authentication using writing sounds with 1DCNN. As a result, we obtained an average accuracy of 82% for the personal authentication for four subjects.

The results of this study indicate that this application of 1DCNN is far from practical use currently. For practical use in the future, development for high classification accuracy is required. For example, preprocessing to avoid aliasing would be one of the effective solutions for the improvement of accuracy. Also, personal authentication by writing sounds using a hidden Markov model has been proposed [13], and it would be interesting to compare its identification performance with that of the proposed model. In such a case, using the writing sound of a signature, rather than a single letter, would make an application more practical. However, since we were able to demonstrate the basic performance of the 1DCNN in this study, we believe that we have shown basic knowledge for future research.

ACKNOWLEDGMENT

This study was performed under the “Cooperative Research Project (No. R02/A31) of the Research Institute of Electrical Communication, Tohoku University.

REFERENCES

[1] Krizhevsky, A.; Sutskever, I.; Hinton, G. ImageNet classification with deep convolutional neural networks. *In Proceedings of the NeurIPS 2012*, Lake Tahoe, USA, 3 December 2012.

[2] Schulz, H.; Behnke, S. Deep Learning. *Künstl Intell*, 2012, 26(4), 357-363.

[3] Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Nets. *In Proceedings of the NeurIPS 2014*, Montreal, Canada, 9 December 2014.

[4] Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; The MIT Press: Cambridge, USA, 2016.

[5] LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature*, 2015, 521, 436–444.

[6] Fukushima, K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 1980, 36(4), 93–202.

[7] Silver, D.; Huang, A.; Maddison, C. et al. Mastering the game of Go with deep neural networks and tree search. *Nature*, 2016, 529, 484–489

[8] Yakovenko, N., Cao, L., Raffel, C., Fan, J. Poker-CNN: A Pattern Learning Strategy for Making Draws and Bets in Poker Games. *arXiv:1509.06731 [cs.AI]*, https://doi.org/10.48550/arXiv.1509.06731

[9] Raví, D., Wong, C., Lo, B., Yang, G. A Deep Learning Approach to on-Node Sensor Data Analytics for Mobile or Wearable Devices. *IEEE Journal of Biomedical and Health Informatics*, 2017, 21(1), 56-64

[10] Kim, Y. Convolutional Neural Networks for Sentence Classification, *arXiv:1408.5882 [cs.CL]*, https://doi.org/10.48550/arXiv.1408.5882

[11] Mei, Z., Ivanov, K., Zhao, G., Wu, Y., Liu, M., Wang, L. Foot type classification using sensor-enabled footwear and 1d-CNN. *Measurement*, 2020, 165, 108184

[12] TensorFlow, https://www.tensorflow.org/

[13] Asakura, D.; Nishino, T.; Naruse, H. Writer recognition with a sound in hand-writing, The Journal of the Acoustical Society of America 140, 3374, 2016; https://doi.org/10.1121/1.4970786

ICATSD2F.117

A NEW CABLE DRIVE SOLUTION FOR TORQUE CONTROL OF EXOSKELETON MOTORIZED JOINTS

TAN HUNG HUYNH[1], MINH THONG NGUYEN[1], HAU TRAN VAN[1], HUNG TINH VU NGUYEN[1], HUY THIEN NGUYEN[1], VIET ANH DUNG CAI[1], VIET THANG NGUYEN[1], LONG TRIET GIANG HUYNH[2]

[1]*Mechatronics Department, School of Engineering Eastern International University, Binh Duong, Vietnam*

[2]*Mechatronics Department, School of Mechanics, Saigon Technology University, Ho Chi Minh, Vietnam*

hung.huynh.set15@alumni.eiu.edu.vn, thong.nguyen.set17@eiu.edu.vn, hau.tran.set17@eiu.edu.vn, vu.nguyen.set15@alumni.eiu.edu.vn, vu.nguyen.set15@alumni.eiu.edu.vn, thien.nguyen.set15@alumni.eiu.edu.vn, dung.cai@eiu.edu.vn, thang.nguyen@eiu.edu.vn, gianghuynhlong.cdt@gmail.com

Abstract. Robotic exoskeleton technology has shown great potential in the development of assistive devices for human neuro-motor rehabilitation. In this paper, the authors present an innovative design of a push-pull cable-driven actuation system including two DC motors and three torque sensors for the control of a human anatomical joint. Two torque sensors are put at the motor shafts to measure the motors' output torques and one is placed at the human targeted joint level to detect the user's intention of motion via the "human/machine" torque of interaction. By using two DC motors, the actuation unit acts like human tendons and provides more natural assistive motion. Experimental results of user's motion tracking are provided and discussed.

Keywords. Rehabilitation Robotics, Knee Exoskeleton, Impedance Control, Interactive Torque Control.

1 INTRODUCTION

Stroke is one of the leading causes of death and disability. Every year, stroke causes 5.7 million of deaths around the world that makes it become the second cause of death after ischemic heart disease [1]. Those who survived from stroke have to suffer from the loss of physical strength and are preferred to undergo therapy to recover the motion ability. In the last 20 years, robotic rehabilitation systems have been continuously developed as the latter are known to have certain advantages in assisting patients during rehabilitation exercises, in a way that these devices can not only provide motion and assistive force but also collect kinematic and dynamic data that are helpful for diagnosis and therapy. Among numerous types of robotic rehabilitation devices, exoskeletons show a great potential. Exoskeletons are devices that could be attached to human limbs through multiple physical contacts so that they could re-create the human's postures and motion and therefore provide assistance to various types of therapy. These devices are divided into two groups according to the rehabilitation purposes: upper limb exoskeletons and lower limb exoskeletons. Since the last decade, many works on rehabilitation exoskeletons have been introduced in the literature. In 2007, Joel C. Perry et al. presented a 7 Degree of Freedoms (DOFs) upper arm exoskeleton which was called CADEN-7. This device was designed to support the whole arm including shoulder, elbow, and wrist [2]. In 2008, an upper limb exoskeleton named RUPERT was introduced by Sivakumar Balasubramanian and his team. The device was wearable and light weight so the patients could perform therapy both while standing and sitting [3]. Besides upper limb exoskeleton, lower limb exoskeleton also caught lots of attention of researchers. In 2015, Shiqian Wang et.al represented a lower limb exoskeleton named MINDWALKER. This exoskeleton had 10 DOFs and all the active joints were actuated by series elastic actuators which could deliver a torque of 100 Nm and a power of 1 kW [4]. In the same year, H2 exoskeleton was introduced by Magdo Bortole1 et al. This device had 6 powered joints and was

programmed to apply torques when the patient walked out of a certain movement pattern [5]. In 2018, Miguel Sanchez-Manchola et al. presented AGoRa, an exoskeleton that was designed for stroke rehabilitation with a passive add/abduction DOFs for improving walking balance [6].
Despite promising experiment results, these types of exoskeleton devices can be bulky with high energy consumption in case that the electric actuators are directly fixed onto the mechanical frame. To overcome this issue, cable-drive actuators can be a good choice. Cable drive solutions not only allow the separation of the motor bloc from an exoskeleton's mechanical mainframe, but also facilitate force control through series-elasticity technique [7] (the spring at the output can be replaced using flexible cables). Several authors have been working on cable-drive actuators developed for exoskeletons. Research on transmission using Bowden Cable were realized by numerous teams in the 2000s [8][9]. This solution allows the control of a distal/non-parallel axis of motion by using 1 motor and 2 cables operating in push-pull mode. This tension of the cable needs to be manually adjusted before use and its value must be large enough to avoid sliding caused by high motor's torque. Researchers [10] also considered the difference of the motor's torque and the output torque at the controlled axis and proposed the measure of these values using torque sensors in order to compensate this difference. The dynamic performance at the output axis depended on the motor's dynamic, which can be limited. One should always find a tradeoff between the motor's maximal torque output versus its output impedance and rise-time. Coupling a powerful motor with a smaller one can help overcome this issue [11]. The small motor then provides small torque resolution meanwhile the bigger motor assures that the required maximum torque can be reached. Furthermore, this solution also allows to compensate the torque output offset each time the rotational movement sense is reversed.
In this paper, we introduced an actuation solution where each motorized joint is actuated by 2 cable driven actuators. The system uses 2 DC motors to control the internal tension of the cable, thus providing push-pull motion by controlling this tension. Torque sensors are mounted not only on the output axis of the targeted joint (to directly measure the interaction with user's limb) but also at each motor axis, that allow the control of the motor output torque in close loop. The design and the control of the system are presented in section 2 and 3. Some first experimental results as well as an example of application on a lower limb exoskeleton are presented and discussed in section 4.

2 DESIGN METHODOLOGY

2.1 Idea of the design

The motorized joints of the exoskeleton (located at the user's knees) are controlled by cable driven actuation units. By using cable driven actuation, the DC motors do not have to be put on the exoskeleton's mechanical frame. As a result, it greatly reduces the weight and the compactness of the frame. The actuation unit is a combination of two DC geared motors and three torque sensors. The structure of the actuation unit is described in Figure 1.

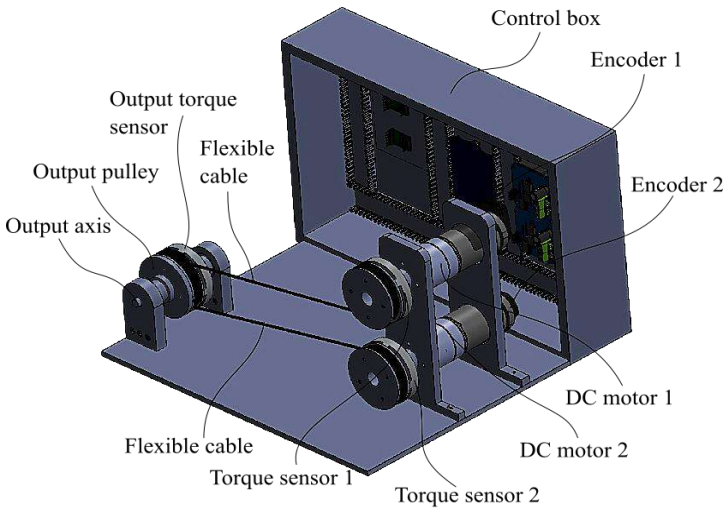


Figure 1. Functioning principle of the cable driven actuation unit

In this actuation unit, there are two torque sensors, which are placed at the motor shafts for measuring the torques generated by the tension of the cable and one at the anatomical joint level to detect the wearer's intention of motion. It can be seen from the figure that the actuator does not have a spring system to generate initial tension inside the cables as this tension can be controlled directly by controlling the torques of the 2 motors, with the help of torque sensors. In our system, we used flexible cables instead of Bowden cable that allow the system to adapt to a larger dynamic range, thus enhancing its stability. For the design of the exoskeleton, the output axis can be placed in any direction and is not necessarily in parallel with the 2 motors' axes, that is a significant advantage comparing to other types of transmission (Belt/chain drives).

2.2 Control Solution

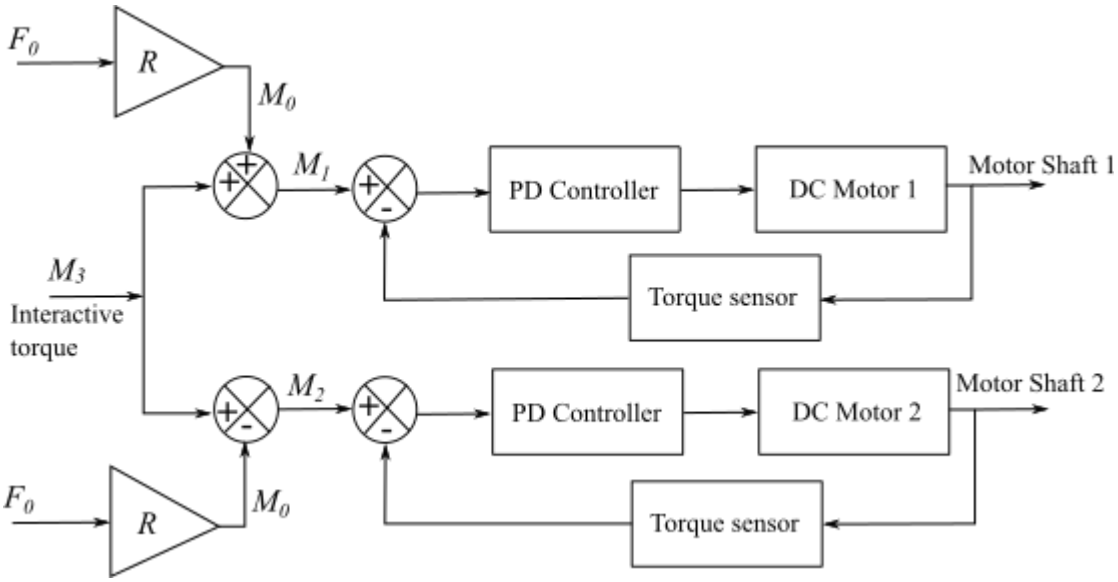


Figure 2. The interactive torque control scheme (for the control of anatomical joint torque)

Figure 2 illustrates the control of the system. Here, the user's intention of motion can be detected by mean of the the torque M_3, that is measured at the output axis. The torque set points M_1 and M_2 of the motors can then be determined following the system of equations below:

$$\begin{cases} M_1 = \left(F_0 + \frac{M_3}{R_3}\right).R_1 \\ M_2 = \left(F_0 - \frac{M_3}{R_3}\right).R_2 \end{cases} \quad (1)$$

in case that $R_1 = R_2 = R_3 = R \Rightarrow \begin{cases} M_1 = F_0.R + M_3 = M_0 + M_3 \\ M_2 = F_0.R - M_3 = M_0 - M_3 \end{cases}$ (2)

In this system of equations, M_0 is the initial torque required to generate the tension F_0 inside the cables when there is no motion (this initial cable tension is set by the operator and is controlled by the actuation unit), M_3 is the interactive torque measured at the anatomical joint, M_1 and M_2 are the torque set-points at the 2 motors' shafts. R is the radius of the pulleys. Here all the 3 pulleys (at the 2 motors' axes and at the output axis) are similar.

When there is no motion ($M_3 = 0$), the torque set-points are equal to the initial torque M_0. In this case, the system only controls the tension inside the cables. If there is intention of motion, that is signaled by an additional torque appeared at the output axis ($M_3 \neq 0$), there will immediately be difference between the two desired torques M_1 and M_2. As a result, one motor will pull, and the other motor will release the cable causing a movement following the user's intention of motion.

No encoder was used since the prototype was built to verify the possibility using the system of equation (1) to control the output torque. For this reason, we focused on the behaviour of motors torques rather than the control in position. The control scheme of the system is shown in figure 2, illustrating the use of 2 PD controllers to control the 2 DC motors' torque. The PD coefficient were chosen experimentally. The idea

is, when there is no interactive torque, the set-points of both PD controllers were equal to the initial torque M_0 which is computed by the product of the tension force F_0 in the transmission cable and the radius R of the pulley. When there is motion's intention, the set-points will be changed, and the actuators would move accordingly and follow the intention. The results of the experiment are shown in figures 4 and 5.

3 RESULTS

3.1 First experimental results

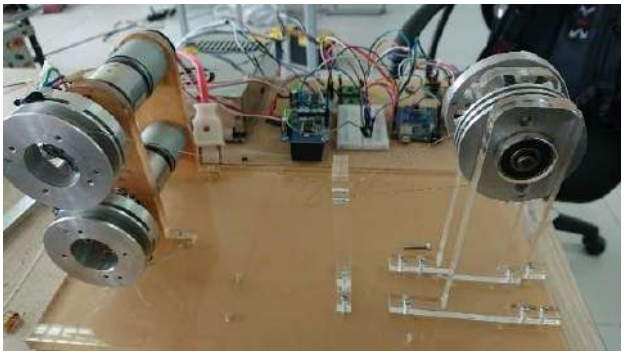


Figure 3. Prototype of the actuation unit

To verify the new actuation solution, a prototype was built as shown in figure 3. The prototype includes 2 DC motors that could be placed randomly, and the motion was transmitted to the output axis via cable transmission. A mechanical part was used to create rough edges that generated frictions when the cable passed through, with the intention to simulate more realistically the situation when the cables slide inside the guiding pipes in an exoskeleton system.

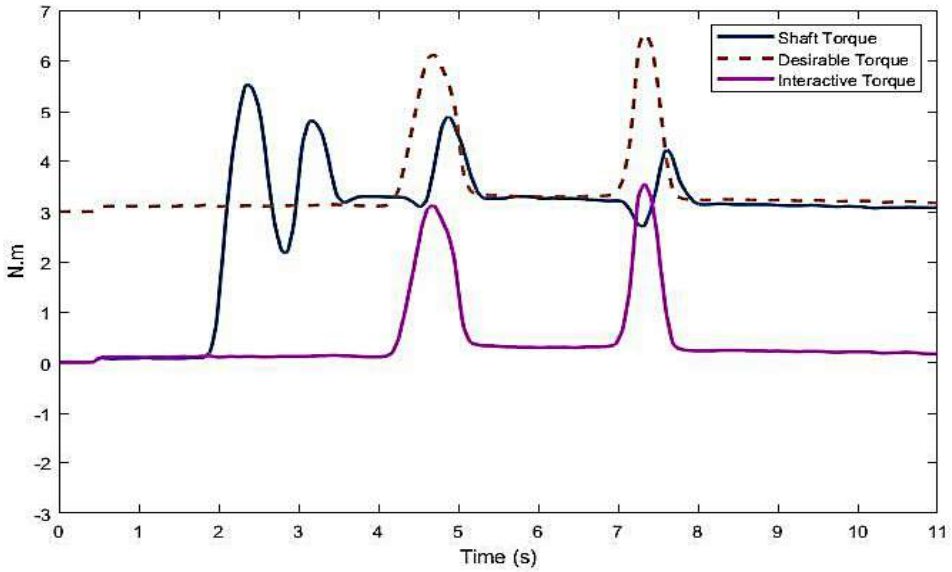


Figure 4. Torque behavior on motor 1

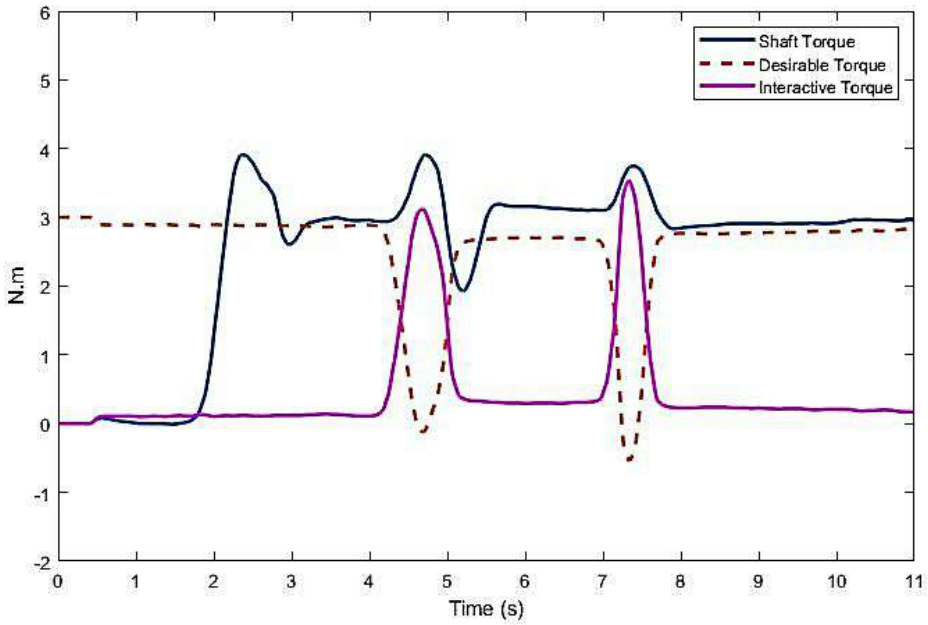


Figure 5. Torque behavior on motor 2

In figures 4 and 5, it can be seen that when the system started up, it took approximately 3 seconds to reach the desired level of tension inside the cable. Then when the interactive torque took place, from the 4th to the 5th second, the shaft's torque of motor 1 slightly dropped. This phenomenon could be observed better between the 7th and 8th second when again the interactive torque increased a second time. However, according to the system of equations (1), the set-point value increased accordingly causing the motor to immediately pull the cable. An opposite phenomenon can be observed in motor 2. The shaft's torque increased when there was interaction at the controlled joint as the output axis motion pulled the cable. Therefore, the motor had to release the cable as the set-point value decreased (see Eq. (1)). This caused a fall in the shaft's torque value just after. We can clearly observe in these 2 figures that one motor pulls while the other one releases the cable so the joint would be able to move following the intention. When the system caught the user's movement (at the output-controlled pulley), the interactive torque dropped to 0 then each shaft torque returned to its initial values.

3.2 Example of application – Assistance of the knee joints

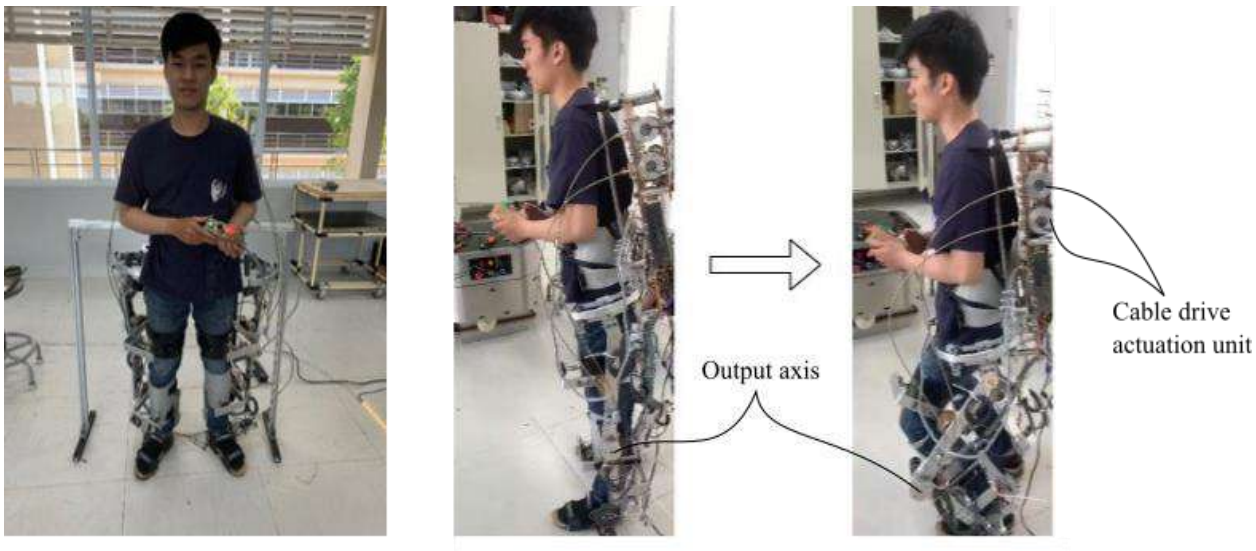


Figure 6. Lower limb exoskeleton using the new cable-drive transmission solution at the 2 knee joints

Figure 6 presents an example of application of this new cable drive transmission solution, integrated into a lower limb exoskeleton. The exoskeleton was designed according to the principles described in [12], with passive joints added into the mechanical linkage to free residual stress resulted from the misalignment between the motorized axis and the human joint, hence improving the transparency and allowing a more natural interaction between the user and the system. Here, the motors are mounted on the back of the subject and only transmit the assistive torques to the user's knee joints and other joints (hips, ankles) are free to move. This design solution reduces the weight and greatly improves the compactness of the whole mechanical structure.

In the very first experiment, the user was asked to simulate stand to sit and sit to stand (STS) motions. Before using the exoskeleton, the wearer has to stand straight up so that the system can record the initial positions of the joints (calibration procedure). The system was controlled following the motion intention using the controller presented in figure 2. The motors shaft torques were defined as in the system of equation (1). Their values can be observed from the results shown in figures 7 and 8. When the user was standing straight, the interactive torques and the extension angles of the knee were 0. When the wearer started sitting down, the negative interactive torque was generated, and the system started moving follow the motion intention.

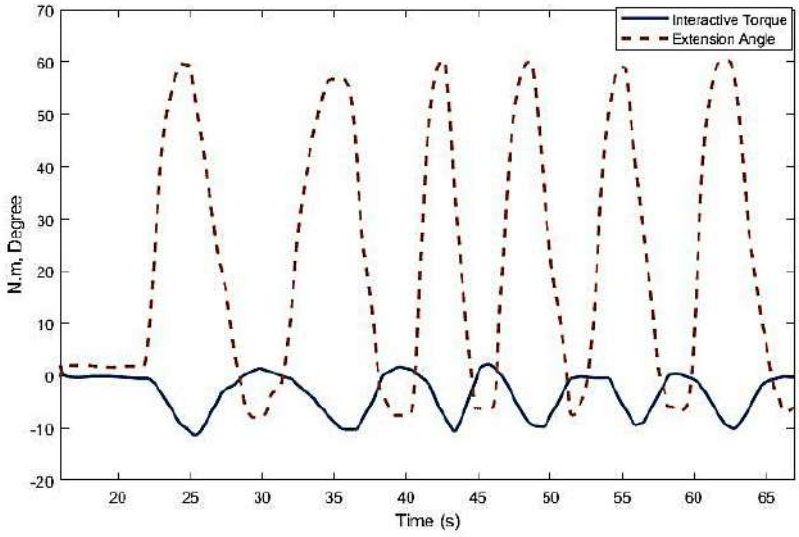


Figure 7. Experimental result on the user's left leg

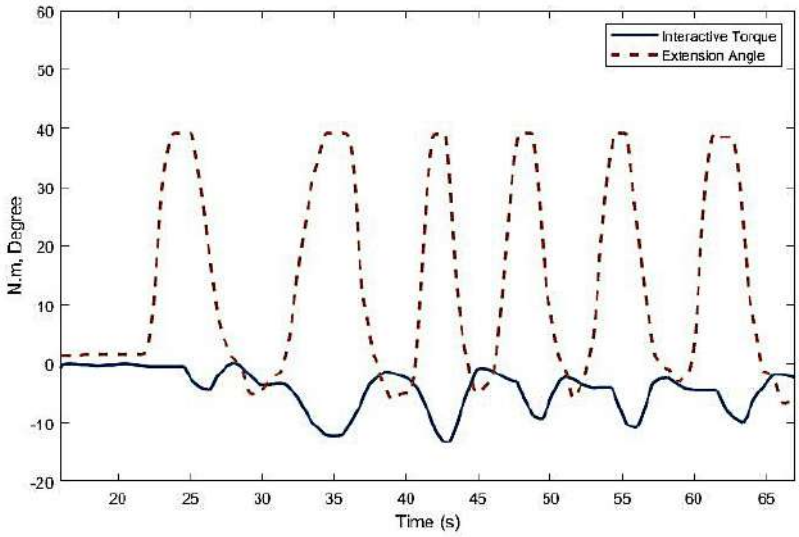


Figure 8. Experimental result on the user's right leg

4 CONCLUSIONS

In this paper, the authors present an innovative cable driven transmission system that includes 2 DC motors and 3 torque sensors for one anatomical joint. Using this actuation solution, the design of powered exoskeletons become simpler as it allows a separation of DC motors from the exoskeleton's mechanical frame. This solution allows a significant reduction of the total weight and dimension of the whole structure. The use of 2 different motors allows an easy control of the transmission cable's tension, which can be measured in real-time by torque sensors. It can be noticed that the combination of 2 motors of different sizes should improve the dynamic range of the system. A simple control algorithm with PD torque controller was developed and tested with success.

The purpose of the development of this actuation solution is to create a "fully transparent robotic exoskeleton". The system was tested on a wearable knee exoskeleton that assists the 2 knee joints of the user by 2 cable driven actuation units. First experimental results have confirmed the capacity of the system to detect the wearer's intention of motion and then to follow this motion, meanwhile cancelling the interactive torques measured at the knees. In a near future, different control strategies should be tested on the system in order to verify the capacity of the system to provide assistance/resistance to the users at high frequency and under different load conditions. This novel actuation solution will be integrated into the team's future design of powered exoskeletons.

REFERENCES

[1] Antonio Di Carlo, Human and economic burden of stroke, *Age and Ageing 2009*, vol 38: 4–5.

[2] Joel C. Perry, Jacob Rosen and Stephen Burns, Upper-Limb Powered Exoskeleton Design, *IEEE/ASME Transactions on Mechatronics*, vol. 12, no. 4, 2007.

[3] Sivakumar Balasubramanian et al., RUPERT: An Exoskeleton Robot for Assisting Rehabilitation of Arm Functions, *Virtual Rehabilitation* , 2008

[4] Shiqian Wang, Design and Control of the MINDWALKER Exoskeleton, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, 2015.

[5] Magdo Bortole1, Anusha Venkatakrishnan, Fangshi Zhu, Juan C Moreno, Gerard E Francisco, Jose L Pons and Jose L Contreras-Vidal, The H2 robotic exoskeleton for gait rehabilitation after stroke: early findings from a clinical study, *Journal of NeuroEngineering and Rehabilitation*, vol.12, no.54, 2015.

[6] Miguel Sanchez-Manchola, Daniel Gomez-Vargas, Diego Casas-Bocanegra, Marcela Munera and Carlos A. Cifuentes, Development of a Robotic Lower-Limb Exoskeleton for Gait Rehabilitation: AGoRA Exoskeleton, *IEEE ANDESCON*, 2018.

[7] D.W.Robinson, Design and Analysis of Series Elasticity in Closed loop Actuator Force Control, *Ph.D. dissertation*, Massachusetts Institute of Technology (MIT), 2000.

[8] J. F. Veneman, R. Ekkelenkamp, R. Kruidhof, F. C. T. van der Helm, H. van der Kooij, Design of a Series Elastic- and Bowdencable-based actuation system for use as torque-actuator in exoskeleton-type training, *Proceedings of the 2005 IEEE 9th International Conference on Rehabilitation Robotics*, Chicago, IL, USA, 2005.

[9] A. Schiele, P. Letier, R. van der Linde and F. van der Helm, Bowden Cable Actuator for Force-Feedback Exoskeletons, *International Conference on Intelligent Robots and Systems (IROS)*, 2006.

[10] S. Grosu, C. Rodiguez-Guerrero, V. Grosu, B. VanderBorght and D. Leferber, Evaluation and Analysis of Push-Pull Cable Actuation System Used for Powered Orthoses, *Frontiers in Robotics and AI*, vol. 5, 2018.

[11] G. Millet, S. Haliyo, S. Regnier and V. Hayward, The Ultimate Haptic Device: First Step, *Proceedings of the Third Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pp. 273-278.

[12] V. A. D. Cai, A. Ibanez, C. Granata, V. T. Nguyen and M. T. Nguyen, Transparency enhancement for an active knee orthosis by a constraint-free mechanical design and a gait phase detection based predictive control, *Meccanica*, 2017.

ICATSD2F.118

PERFORMANCE ANALYSIS 16QAM-OFDM SIGNAL EMPLOYING DVB-T2 STANDARD IN 8K MODE

NGUYEN HOANG VIET

[1]*Faculty of Electronics Technology, Industrial University of Ho Chi Minh City, Ho Chi Minh City, Vietnam*

nguyenhoangviet@iuh.edu.vn

Abstract. We have simulated about modulation and demodulation OFDM signal in DVB-T2 for 8k mode. The application of 16-QAM modulation technique to increase the bit rate of transmission and use IFFT / FFT transform to show orthogonality between subcarriers before digitizing the signal to broadcast long distances. In addition, this paper is also concerned with the influence of AWGN channel on OFDM system through signal to noise (SNR) compared to the bit error rate (BER) by Matlab software.

Keywords. digital video broadcasting - terrestrial 2 (DVB-T2), additive white Gaussian noise (AWGN), orthogonal frequency division multiplexing (OFDM).

1 INTRODUCTION

During the past three decades, the demands for the quantity and quality of television programs have increased continuously. There are currently three standards for digital terrestrial television: ATSC [1], DVB-T [2] and ISDB-T [3]. The three standards above have the same point of using the MPEG-2 compression standard for video signals. ATSC uses 8-VSB modulation while DVB-T and ISDB-T use OFDM orthogonal multiplexing, component carriers using QPSK, 16-QAM or 64-QAM modulation techniques [4] [5]. Recently, OFDM signals will become a strong candidate for transmission. OFDM systems have much greater spectrum efficiency than conventional modulation techniques [6], [7], [8], [9]. OFDM is applied in many different fields such as broadband internet, 5G, DVB-T2 digital terrestrial television system [10], [11], [12], [13]. This paper only mentions the application of OFDM technique in the DVB-T2 terrestrial digital television system, in 8K mode and the application of 16-QAM to increase the transmission bit rate.

2 SYSTEM DESIGN AND CONFIGURATION

2.1 OFDM transmission

The process of digital 16-QAM modulation is shown in Figure 1:

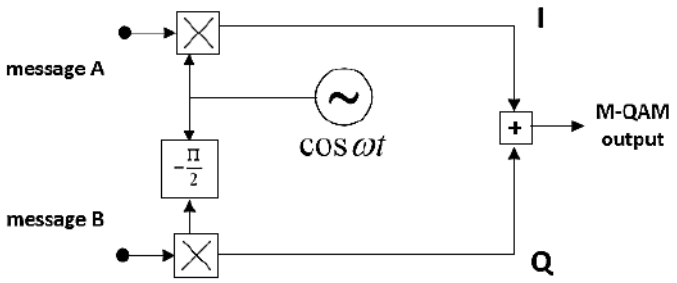


Figure 1. 16-QAM modulation scheme

The 16-QAM modulation equation is shown in figure 1:

$$X_k = \sqrt{\frac{2E_0}{T}}\, a_k\, cos(\,\omega t) - \sqrt{\frac{2E_0}{T}}\, b_k\, sin(\,\omega t)\ (0 \le t \le T) \quad (1)$$

E_0: energy of the signal with the lowest amplitude

a_k, b_k: pair of independent integers depending on constellation position

Then, the constellation 16-QAM will have the form:

$$X_k = \begin{bmatrix} (-3,3j) & (-1,3j) & (1,3j) & (3,3j) \\ (-3,j) & (-1,j) & (1,j) & (3,j) \\ (-3,-j) & (-1,-j) & (1,-j) & (3,-j) \\ (-3,-3j) & (-1,-3j) & (1,-3j) & (3,-3j) \end{bmatrix}$$

$$x(t) = \frac{1}{N_s}\sum_{k=k_{MIN}}^{k=k_{MAX}} X_k(f).\, e^{\frac{j2\pi kt}{T}} \quad (2)$$

$x(t), 0 \le t \le T_U$ discreate signal series in the time domain

N_s: carrier number

Output in OFDM transmitter system:

$$s(t) = e^{j2\pi f_c t} \sum_{k=K_{min}}^{K_{max}\Sigma} X_k(f) e^{j2\pi k'(t-T_g)/T_U} \quad (3)$$

$e^{j2\pi f_c t}$ high frequency components for OFDM signals to be transmitted in the VHF frequency range from 30MHz to 300MHz

The block diagram of OFDM transmitter modulation for 8k mode is shown in Figure 2:

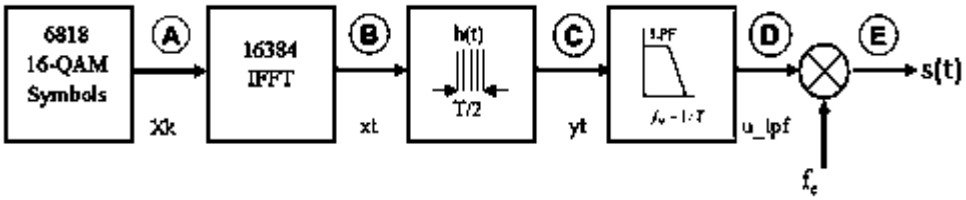


Figure 2. The block diagram of OFDM transmitter modulation for 8k mode

2.2 OFDM Reception

In this paper, OFDM signal reception is just the opposite of transmission. The block diagram of OFDM receiver demodulation for 8k mode is shown in Figure 3:

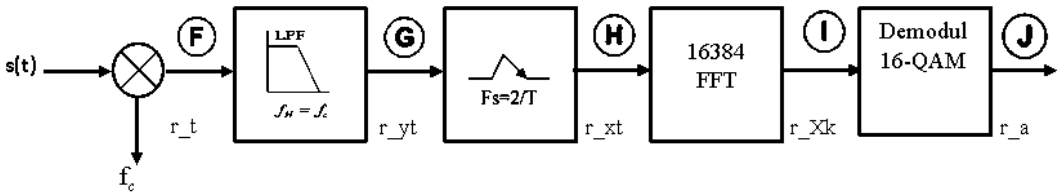


Figure 3. The block diagram of OFDM receiver demodulation for 8k mode

2.3 AWGN channel

The formula for OFDM signal over AWGN channel as follows:

$$y'(t) = x(t) \otimes h(t) + n(t) \quad (4)$$
$$h(t) = rect_N(t)$$

$n(t)$ white Gaussian noise

The influence of the AWGN channel on the M-QAM digital modulation is also shown in [10] through the formula:

$$BER_M = 1 - (1 - BER_{\sqrt{M}})^2 \quad (5)$$

To accurately calculate the bit error rate on the AWGN channel, in [10], the operating range is limited as follows:

$$BER_M \le 4 \cdot Q\sqrt{\frac{3 \cdot k \cdot SNR}{(M-1)}} \quad (6)$$

3 SIMULATION RESULTS

3.1 The simulation of OFDM signal transmission

The simulation of OFDM signal generation is performed based on Figure 2, as A, B, C, D, and E point.

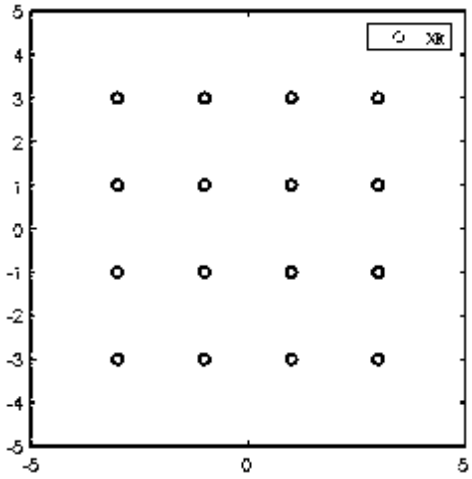


Figure 4. The constellation 16-QAM of the signal at A

The constellation 16-QAM of the signal at A is shown in Figure 4.

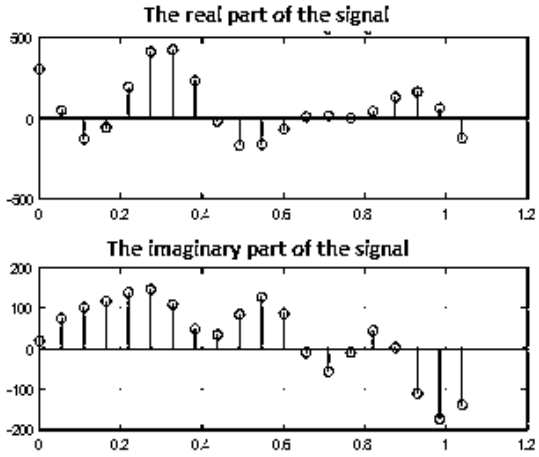


Figure 5. The signal x(t) at B

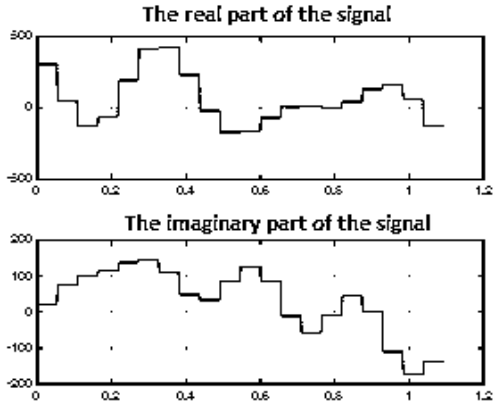


Figure 6. The signal y(t) at C

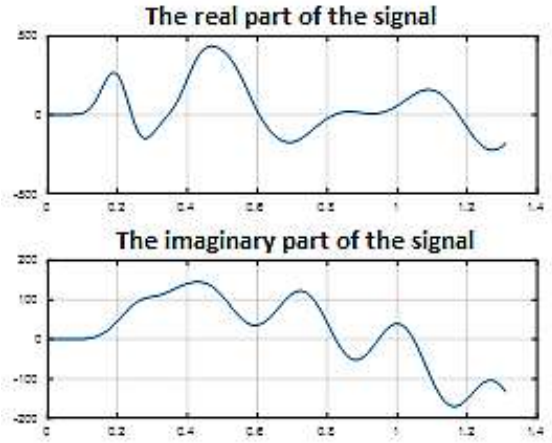


Figure 7. The signal u(t) at D

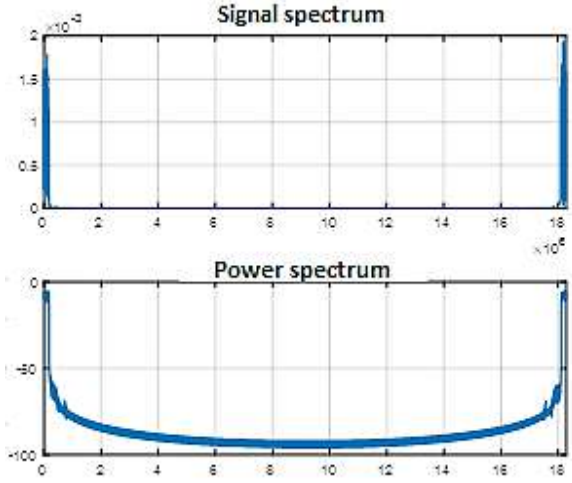


Figure 8. The signal spectrum and the power spectrum at D

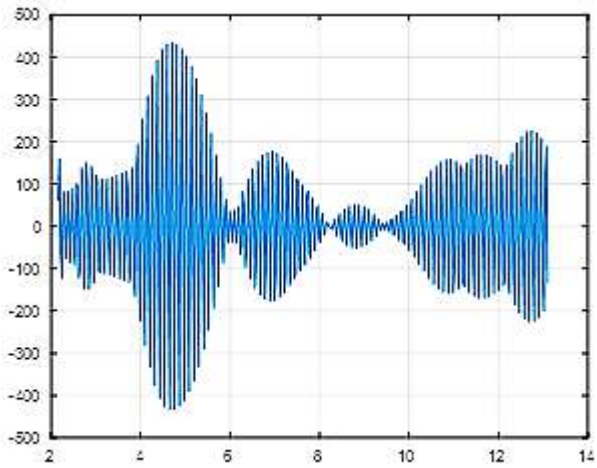


Figure 9. The OFDM signal s(t) at E

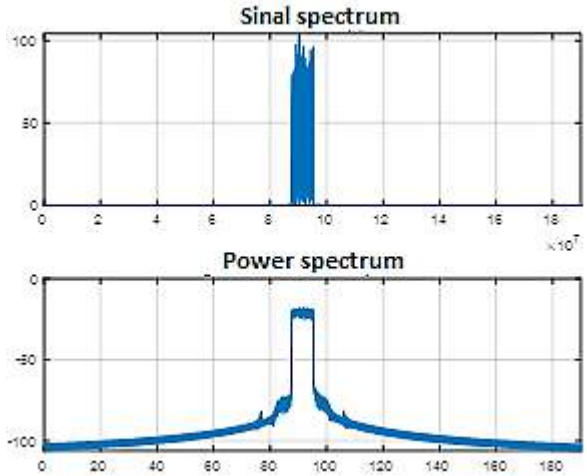


Figure 10. The signal spectrum and the power spectrum at E

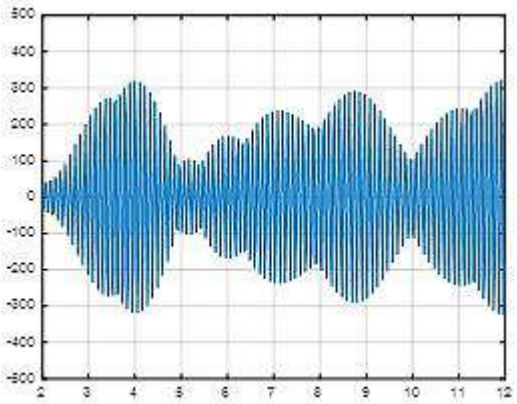


Figure 11. The simulation based on formula (3)

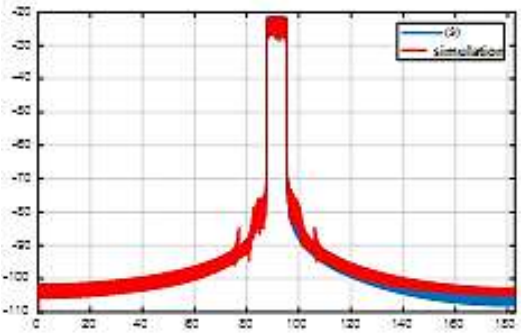


Figure 12. The signal spectrum based on formula (3) and based on simulation

The signal spectrum based on formula (3) and based on simulation is shown in Figure12. We can see that the result in in good agreement.

3.2 The simulation of OFDM signal reception

The simulation of OFDM signal generation is performed based on Figure 3, as F, G, H, I, and J point.

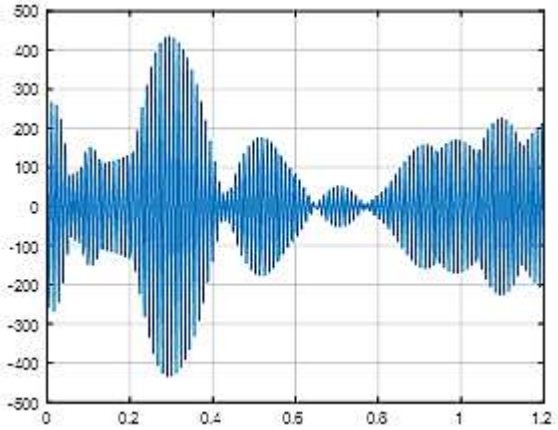


Figure 13. The signal s(t) at F

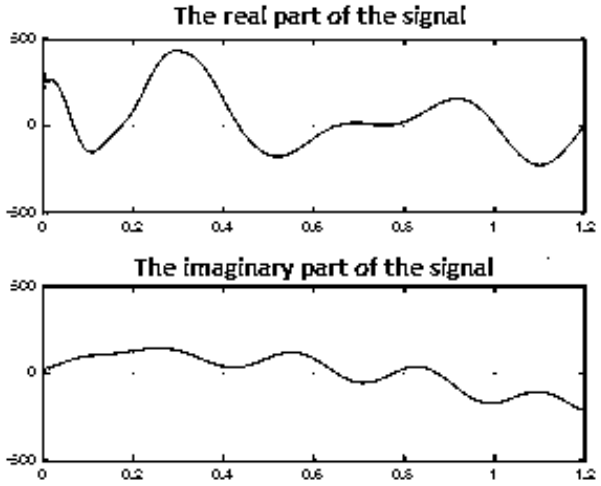


Figure 14. The signal at G

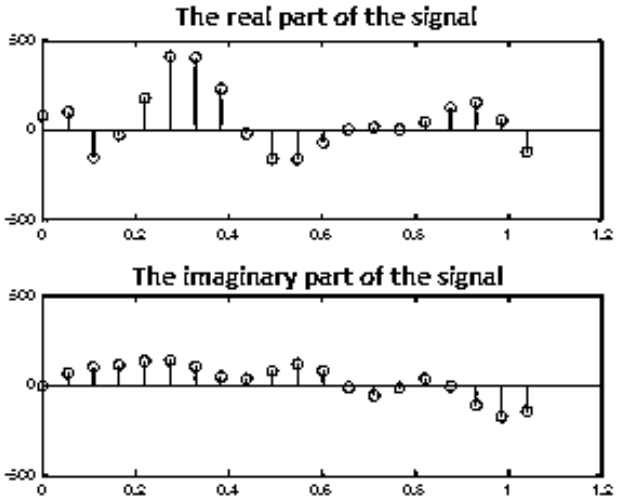


Figure 15. The signal at H

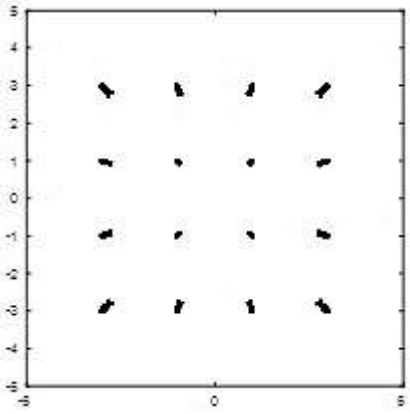


Figure 16. The constellation 16-QAM of the signal at I

The constellation 16-QAM of the signal at I is shown in Figure 16 which is still good agreement.

3.3 The simulation on AWGN channel

The influence of the AWGN channel on the 16-QAM digital modulation is also simulated.

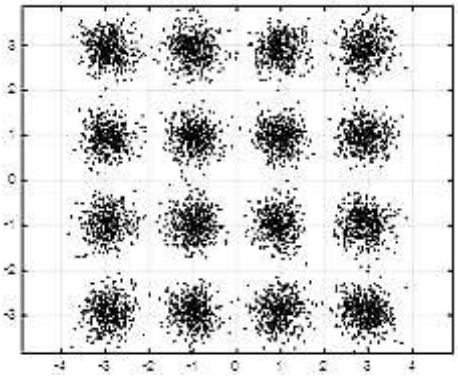


Figure 17. The constellation 16-QAM with AWGN noise

The constellation 16-QAM with AWGN noise is shown in Figure 17. It shows little phase distortion due to AWGN noise, but constellation diagram performance is still accepted.
The measured BER curve is shown in Figure18.

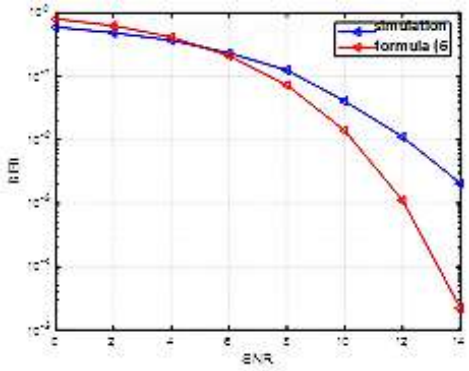


Figure 18. The BER curve

In AWGN channel simulation, the bit error rate (BER) difference between simulation and calculation is quite small. In general, we have the higher the SNR, the lower the bit error rate.

4 CONCLUSIONS

We have simulated modulation and demodulation OFDM signal in DVB-T2 for 8k mode. The application of 16-QAM modulation technique to increase the bit rate of transmission and use IFFT / FFT transform to show orthogonality between subcarriers before digitizing the signal to broadcast long distances. The application of 16-QAM digital modulation technique compared to 4-QAM will help increase the bit rate of transmission. The simulation results have analyzed that the received signal is almost the same as the original signal. Finally, the application of OFDM technology in terrestrial digital television (DVB-T2) has been highly effective.

REFERENCES

[1] Advanced Television Systems Committee, "ATSC Digital Television Standard," Document A/53, Sept. 1995.

[2] Digital video broadcasting (DVB); Framing structure, channel coding and modulation for terrestrial television, European Standard (EN) 300 744 V1.5.1, European Telecommunications Standards Institute (ETSI), Nov. 2004.

[3] Association of Radio Industries and Businesses, "Terrestrial Integrated Services Digital Broadcasting (ISDB-T) — Specifications of Channel Coding, Framing Structure, and Modulation," Sept. 1998.

[4] Nguyễn Quốc Nam, *Điều chế OFDM trong truyền hình số mặt đất DVB-T,* Đại học Quốc gia Hà Nội Trường Đại học Công nghệ, 2014.

[5] Tạ Nam Thắng, *Xây dựng mạng đơn tần trong truyền hình số mặt đất,* Đại học quốc gia Hà Nội, 2004.

[6] O. Omijeh and S. D. O. Agoye "Simulation and Evaluation of OFDM Based Digital Video Broadcasting over Different Wireless Communication Channels" International Journal of Scientific & Engineering Research Volume 6 Issue 10, pp. 738- 745 October-2015

[7] Dillip Dash, Arun Agarwa Kabita Agarwal, "Performance Analysis of OFDM Based DVB-T over Diverse Wireless Communication Channels", International Journal of Electronics and Communication Engineering. Volume 6, Number 1 (2013), pp. 131-141.

[8] Y. A. Lafta and P. Johnson "High performance OFDM systems for digital video broadcasting-terrestrial (DVB-T)" International Journal of Digital Information and Wireless Communications (IJDIWC) 2(1): 66-74 The Society of Digital Information and Wireless Communications, 2012.

[9] Ajay Siwach Sunil Malhotra,Ravi Shankar "OFDM based digital video broadcasting(dvb-t) technique analysis" International Journal of Engineering Trends and Technology (IJETT) - Volume4Issue4- April 2013.

[10] I. Eizmendi, M. Velez, D. G. Barquero, and J. Morgade, DVB-T2: The second generation of terrestrial digital video broadcasting system, IEEE Transactions on Broadcasting 60(2): 258-271, 2014.

[11] A. C. Honfoga, M. Dossou, and V. Moeyaert, Analysis of Channel uncertainty on OFDM/FBMC DVB-T2 simulations, [Online]. Available: https://doi.org/10.1007/978-3-031-06374-9_12, 2022.

[12] Honfoga, A.-C., Nguyen, T.T., Dossou, M., Moeyaert, V.: Application of FBMC to DVB-T2: a comparison vs classical OFDM transmissions. In: IEEE GlobalSIP Conference (2019). [Online]. Available: https://doi.org/10.1109/GlobalSIP45357.2019.8969550.

[13] Y. Yang, C. Lim, A. Nirmalathas, Investigation on Transport Schemes for Efficient High-Frequency Broadband OFDM Transmission in Fibre-Wireless Links, Journal of Lightwave Technology, VOL.32, NO. 2, 2014.

.

ICATSD2F.119

A NOVEL MULTILEVEL INVERTER USING SMALL CAPACITORS

NGO BAC BIEN, PHAN XUAN DUNG, TRINH NGOC DUC, NGO THI KIM LINH
School of Engineering, Eastern International University
bien.ngo@eiu.edu.vn, dung.phan@eiu.edu.vn, duc.trinh@eiu.edu.vn, linh.ngo@eiu.edu.vn

Abstract. This study proposes a novel switched-capacitor eleven-level inverter which can operate with small capacitors. The proposed topology controls the charged/discharged processes of the capacitors by the switching frequency. If the switching frequency is higher, the ripple voltage of capacitors is lower. In simple terms, when using the high switching frequency, the small capacitors can be used in this circuit. Therefore, the size of the inverter is reduced, and the reliability is improved. In addition, the paper presents the sine pulse-width modulation (sine PWM) method used to control the operation of the circuit. The simulation results of the proposed inverter are performed by PSIM software at the different switching frequencies (5 kHz and 20 kHz) along with the different capacitor values (100 µF and 400 µF) shown in the paper.

Keywords. switched-capacitor, multilevel inverter, control method, sine-PWM.

1 INTRODUCTION

The inverter takes an important role in converting the DC power supply to the AC power, which is applied in motor drives, renewable energy systems, uninterruptible power supplies, static var compensators, active filters, flexible AC transmission systems, and voltage compensators. The simplest form of the inverter has two different voltage values, and it is called a two-level inverter [1]. However, this topology has several drawbacks reducing the performance. The multilevel inverter is the advanced form with a staircase output waveform that eliminates the disadvantage of the simplest inverter [1] – [4]:

- Increase voltage levels.
- Improve output voltage quality.
- Extend power range.
- Decrease switching voltage rate change.
- Reduce common-mode voltage

Despite its advantages, the conventional multilevel inverter has more challenges such as the high numbers of semiconductors, multiple DC sources, the complexity of control methods, the size, and the cost of the inverter [5] – [7].

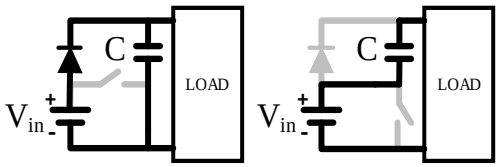


Figure 1. The charged/discharged characteristic of the capacitor.

To solve the issues of the conventional inverter, there are several topologies have been proposed. One of them is switched-capacitor multilevel inverters (SCMI) which optimize cost, size as well as the control method [8] – [11]. SCMI is the combination of a switched-capacitor (SC) structure and an H-bridge circuit, including a DC source, switches, diodes, and capacitors. In SC structure, the charged/discharged characteristic of the capacitor is used to create the staircase voltage waveform at output. In Figure , the capacitor charges when it is connected in parallel with the DC source and discharges when it is connected in series with the DC source. If the switches are turned on and turned off altenately, the charged and discharge processes of the capacitor in SC will occur consecutively. Therefore, the capacitor in SC can be self-balanced by switches without using the auxiliary circuits. In addition, SCMI features by multilevel

output waveform, and the output voltage value is higher than the input DC source.
In SCMI, the capacitor is one of the important components, but it has many challenges to design and develop topology. The ripple voltage waveform of the capacitor directly affects the output quality; the bigger of ripple voltage value, the higher the total distortion harmonics [8] – [11]. To solve this problem, the capacitor with high capacitance is selected (normally 2200 µF). However, it increases the size of capacitors as well as the overall size of the inverter. Furthermore, the reliability of the inverter also is affected. Another method to overcome this problem, it increases the charge/discharge frequency of the capacitors [12]. In simple terms, the charge/discharged time is shorter, the ripple voltage value is smaller.
This paper proposed a novel SCMI that increases the charged/discharged frequency to reduce the size of capacitors. Moreover, capacitor size depends on the switching frequency, the higher frequency, the smaller size of the capacitor. This paper presents the topology along with the control method of the proposed circuit. Finally, the simulation result performed by PSIM 9.0 software is shown to validate the operation of the proposed circuit.

2 PROPOSED MULTILEVEL INVERTER USING SMALL CAPACITORS

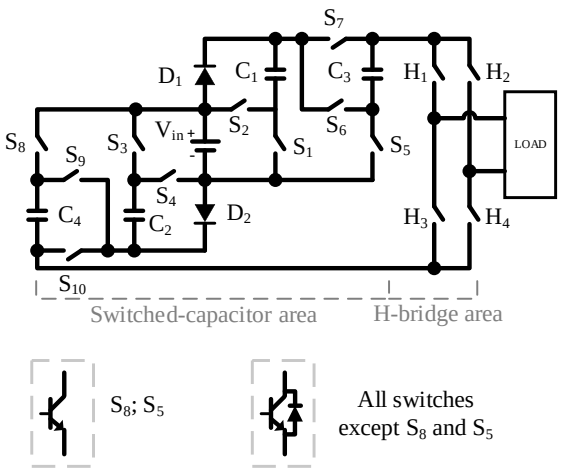


Figure 2. The proposed eleven-level inverter using small capacitors.

Figure shows the proposed eleven-level inverter using small capacitors. This topology includes an H-bridge area with 4 switches and a switched-capacitor area using 4 capacitors, 2 diodes, and 10 switches to create an eleven-level voltage waveform at output. In the switched-capacitor area, every capacitor comes with switches or diodes to control its operation such as C_1 is S_1 and S_2, C_2 is S_3 and S_4, C_3 is S_5, S_6 and S_7, C_4 is S_8, S_9 and S_{10}.

2.1 The operation principle

2.1.1 The charged/discharged process of capacitors

Table describes the charged/discharged process of capacitors:
- Capacitor C_1 ($V_{C1}=V_{in}$): When S_1 turn on and S_2 turn off, C_1 connects in parallel with V_{in}; and C_1 is charged from V_{in}. When S_1 turns off and S_2 turns on, C_1 connects in series with V_{in}; and C_1 is discharged to load.
- Capacitor C_2 ($V_{C2}=V_{in}$): When S_3 turn on and S_4 turn off, C_2 connects in parallel with V_{in}; and C_2 is charged from V_{in}. When S_3 turns off and S_4 turns on. C_2 connects in series with V_{in}, and C_2 is discharged to load.
- Capacitor C_3 ($V_{C3}=V_{in}+ V_{C1}$): When C_1 discharge, S_6 turns off and S_5, S_7 turn on, C_3 connects in parallel with V_{in} and C_1; and C_3 is charged from V_{in} and C_1. When S_6 turn on and S_5, S_7 turn off. C_3 connects in series with V_{in}, and C_3 is discharged to load.
- Capacitor C_4 ($V_{C4}=V_{in}+ V_{C2}$): When C_2 discharge, S_9 turns off and S_8, and S_{10} turn on, C_4 connects in parallel with V_{in} and C_2; and C_4 is charged from V_{in} and C_2. When S_9 turns on and S_8, S_{10} turn off. C_4 connects in series with V_{in}, and C_4 is discharged to load.

Table 1. The charged/discharged process of capacitors.

Capacitor	Status	Voltage value	Switches		Note
			ON	OFF	
C_1	Charge	V_{in}	S_1	S_2	
	Discharge		S_2	S_1	
C_2	Charge	V_{in}	S_3	S_4	
	Discharge		S_4	S_3	
C_3	Charge	$2V_{in}$	S_5, S_7	S_6	C_1 discharge
	Discharge		S_6	S_5, S_7	
C_4	Charge	$2V_{in}$	S_8, S_{10}	S_9	C_2 discharge
	Discharge		S_9	S_8, S_{10}	

2.1.2 The proposed topology operation

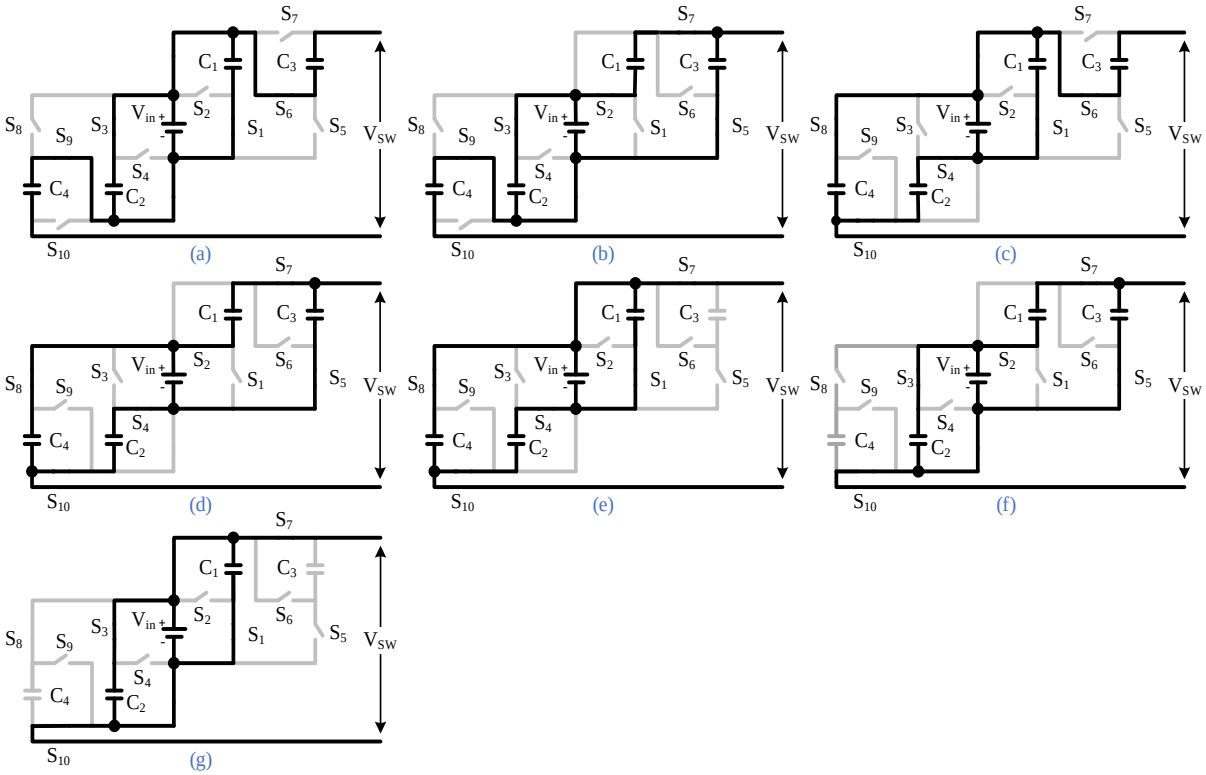


Figure 3. The operation stages of switched-capacitor area: (a) stage a: $V_{sw}=5V_{in}$; (b) stage b: $V_{sw}=4V_{in}$; (c) stage c: $V_{sw}=4V_{in}$; (d) stage d: $V_{sw}=3V_{in}$; (e) stage e: $V_{sw}=2V_{in}$; (f) stage f: $V_{sw}=2V_{in}$; (g) stage g: $V_{sw}=V_{in}$.

The proposed topology is separated into two areas: The H-bridge area and the switched-capacitor area. The H-bridge area includes 4 switches that are used to create the positive period, the negative period and zero-voltage at the output. When the circuit operates at the positive period, switches H_1 and H_4 turn on; H_2 and H_3 turn off. When the circuit is in the negative period, the state of the switches is the opposite. In zero-voltage status, H_1 and H_2 turn on, H_3 and H_4 turn off; or vice versa.

Figure shows seven operation stages of the switched-capacitor area with the output voltage (V_{SW}) changes from $5V_{in}$ to V_{in}. The seven operation stages include:

➢ **Stage a (Fig 3.a):**

- S_1, S_3 turn on; S_2, S_4 turn off. Capacitors C_1 and C_2 are charged from the input source $V_{C1}=V_{C2}=V_{in}$.
- S_6, S_9 turn on; S_5, S_7, S_8, S_{10} turn off. Capacitors C_3 and C_4 are discharged.

▪ The output voltage $V_{SW}=V_{C3}+V_{C4}+V_{in}=5V_{in}$.

➢ **Stage b (Fig 3.b):**

▪ S_2, S_3 turn on; S_1, S_4 turn off. Capacitor C_1 is discharged, whereas capacitor C_2 is charged from the input source $V_{C2}=V_{in}$.

▪ S_5, S_7, S_9 turn on; S_6, S_8, S_{10} turn off. Capacitor C_4 is discharged, capacitors C_3 is charged from the input source and C_1; $V_{C3}=V_{in}+V_{C1}=2V_{in}$.

▪ The output voltage $V_{SW}=V_{C1}+V_{C4}+V_{in}=4V_{in}$.

➢ **Stage c (Fig 3.c):**

▪ S_1, S_4 turn on; S_3, S_2 turn off. Capacitor C_2 is discharged, whereas capacitor C_1 is charged from the input source $V_{C1}=V_{in}$.

▪ S_5, S_7, S_9 turn off; S_6, S_8, S_{10} turn on. Capacitor C_3 is discharged, capacitors C_4 is charged from the input source and C_2; $V_{C4}=V_{in}+V_{C2}=2V_{in}$.

▪ The output voltage $V_{SW}=V_{C2}+V_{C3}+V_{in}=4V_{in}$.

➢ **Stage d (Fig 3.d):**

▪ S_2, S_4 turn on; S_1, S_3 turn off. Capacitors C_1 and C_2 are discharged.

▪ S_5, S_7, S_8, S_{10} turn on; S_6, S_9 turn off. Capacitors C_3 and C_4 are charged from the input source and C_1, C_2; $V_{C3}=V_{in}+V_{C1}=2V_{in}$, $V_{C4}=V_{in}+V_{C2}=2V_{in}$.

▪ The output voltage $V_{SW}=V_{C1}+V_{C2}+V_{in}=3V_{in}$.

➢ **Stage e (Fig 3.e):**

▪ S_1, S_4 turn on; S_2, S_3 turn off. Capacitor C_2 is discharged, whereas capacitor C_1 is charged from the input source $V_{C1}=V_{in}$. S_7, S_8, S_{10} turn on; S_6, S_9, S_5 turn off.

▪ Capacitor C_4 is charged from the input source and C_2; $V_{C4}= V_{C2} + V_{in} =2V_{in}$.

▪ The output voltage $V_{SW}= V_{C2}+V_{in}=2V_{in}$.

➢ **Stage f (Fig 3.f):**

▪ S_2, S_3 turn on; S_1, S_4 turn off. Capacitor C_1 is discharged, whereas capacitor C_2 is charged from the input source $V_{C2}=V_{in}$.

▪ S_5, S_7, S_{10} turn on; S_6, S_8, S_9 turn off. Capacitor C_3 is charged from the input source and C_1; $V_{C3}=V_{in}+V_{C1}=2V_{in}$.

▪ The output voltage $V_{SW}= V_{C1}+V_{in}=2V_{in}$.

➢ **Stage g (Fig 3.g):**

▪ S_1, S_3 turn on; S_2, S_4 turn off. Capacitors C_1 and C_2 are charged from the input source $V_{C1}=V_{C2}=V_{in}$.

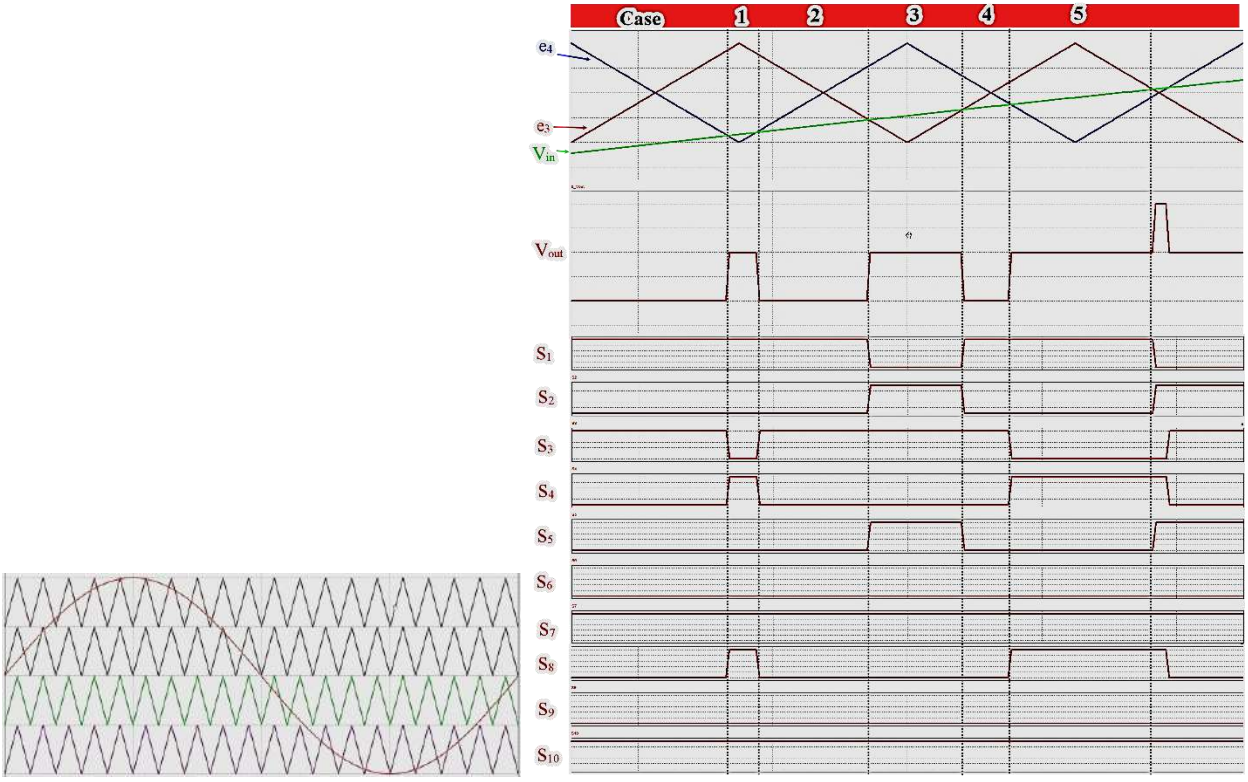
▪ The output voltage $V_{SW}= V_{in}$

Table 2. The operation states of the proposed topology.

Switching state	Switches (ON)	Switched-capacitor area	Capacitor status (C: Charge; D: Discharge)				V_{out}
			C_1	C_2	C_3	C_4	
1	S_1, S_3, S_6, S_9, H_1, H_4	Fig. 3.a	C	C	D	D	$5V_{in}$
2	S_2, S_3, S_5, S_7, S_9, H_1, H_4	Fig. 3.b	D	C	C	D	$4V_{in}$
3	S_1, S_4, S_6, S_8, S_{10}, H_1, H_4	Fig. 3.c	C	D	D	C	$4V_{in}$
4	S_2, S_4, S_5, S_7, S_8, S_{10}, H_1, H_4	Fig. 3.d	D	D	C	C	$3V_{in}$
5	S_1, S_4, S_7, S_8, S_{10}, H_1, H_4	Fig. 3.e	C	D	-	C	$2V_{in}$
6	S_2, S_3, S_5, S_7, S_{10}, H_1, H_4	Fig. 3.f	D	C	C	-	$2V_{in}$
7	S_1, S_3, S_7, S_{10}, H_1, H_4	Fig. 3.g	C	C	-	-	V_{in}
8	S_1, S_3, S_7, S_{10}, H_1, H_2	Fig. 3.g	C	C	-	-	0
9	S_1, S_3, S_7, S_{10}, H_2, H_3	Fig. 3.g	C	C	-	-	$-V_{in}$
10	S_2, S_3, S_5, S_7, S_{10}, H_2, H_3	Fig. 3.f	D	C	C	-	$-2V_{in}$
11	S_1, S_4, S_7, S_8, S_{10}, H_2, H_3	Fig. 3.e	C	D	-	C	$-2V_{in}$
12	S_2, S_4, S_5, S_7, S_8, S_{10}, H_2, H_3	Fig. 3.d	D	D	C	C	$-3V_{in}$
13	S_1, S_4, S_6, S_8, S_{10}, H_2, H_3	Fig. 3.c	C	D	D	C	$-4V_{in}$
14	S_2, S_3, S_5, S_7, S_9, H_2, H_3	Fig. 3.b	D	C	C	D	$-4V_{in}$
15	S_1, S_3, S_6, S_9, H_2, H_3	Fig. 3.a	C	C	D	D	$-5V_{in}$

Table describes all switching states of the proposed topology including the turned-on switches, the status of capacitors and the output voltage of the proposed topology (V_{out}).

2.2 Control method



(a) (b)

Figure 4. (a) Sine-PWM technique; (b) PWM technique for stage.

To control the operation of the inverters, PWM techniques are used such as space vector modulation [13], phase disposition (PD) PWM [14], phase opposite disposition PWM [15], and hybrid modulation [16]. In PWM techniques, the sine-PWM technique is the most popular. Figure .a shows the sine PWM technique that uses a sine waveform (reference waveform) and four triangle waveforms (carrier waveform) to create the control signals for switches. The triangle waveform is usually the same peak-to-peak voltage and the same phase, but the offset is different. If two triangle is close, the peak value of a waveform is the offset value of another waveform as Figure .a. Depending on the level number of the output voltage of the circuit, the amplitude of the reference waveform is modified.

$$A_s = \frac{n_{level}-1}{2} A_c \quad (1)$$

Where A_s: the amplitude of sine waveform.

n_{level}: the level number of output voltage.

A_c: the amplitude of triangle waveform.

The proposed topology uses the sine-PWM technique to control the operation of the circuit. Figure .b shows the PWM for the switching stages 5 and 6 in Table . Two triangle waveforms are the same peak – to - peak voltage and 180 degrees out of phase.

In Figure .b, cases is corresponding with the stage in Figure and Table :

- Case 1: switching stage 5, Figure 3.e ($2V_{in}$).
- Case 2: switching stage 7, Figure 3.g (V_{in}).
- Case 3: switching stage 6, Figure 3.f ($2V_{in}$).
- Case 4: switching stage 7, Figure 3.g (V_{in}).
- Case 5: switching stage 5, Figure 3.e ($2V_{in}$).

From Figure .b, switching stages 5 and 6 occur alternatively; capacitors C_1 and C_2 are charged and discharged consecutively. Therefore, the ripple voltage waveform of C_1 and C_2 can control by switching frequency.

3 SIMULATION RESULTS

To verify the operation of the proposed topology, the simulation result is shown in this section with the parameter in Table .

Table 3. The parameters used in the simulation.

Input voltage (V_{in})	24 V
Capacitor (C_1, C_2, C_3, C_4)	400 µF, 100 µF
Inductor (L)	14 mH
Resistor (R)	30 Ohm
Reference frequency (f_r)	50 Hz
Switching frequency (f_{sw})	5 kHz, 20 kHz

3.1 Simulation result for capacitor value: 400 µF

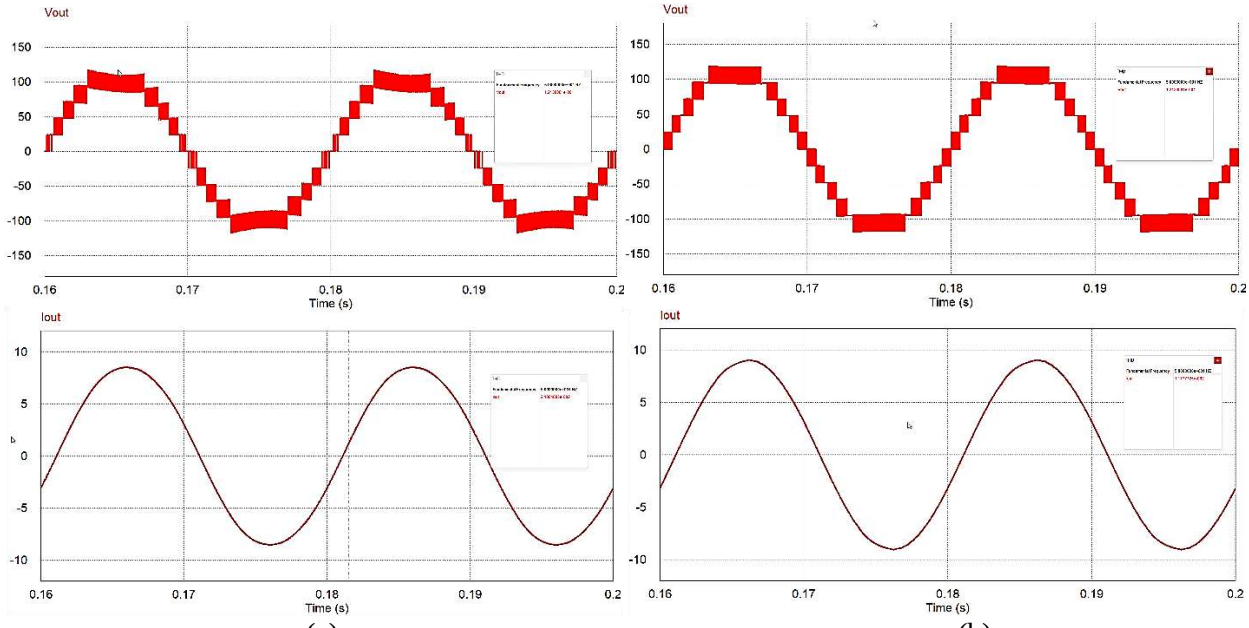


Figure 5. The output voltage and output current of the proposed inverter (a) f_{sw}=5 kHz; (b) f_{sw}=20 kHz.

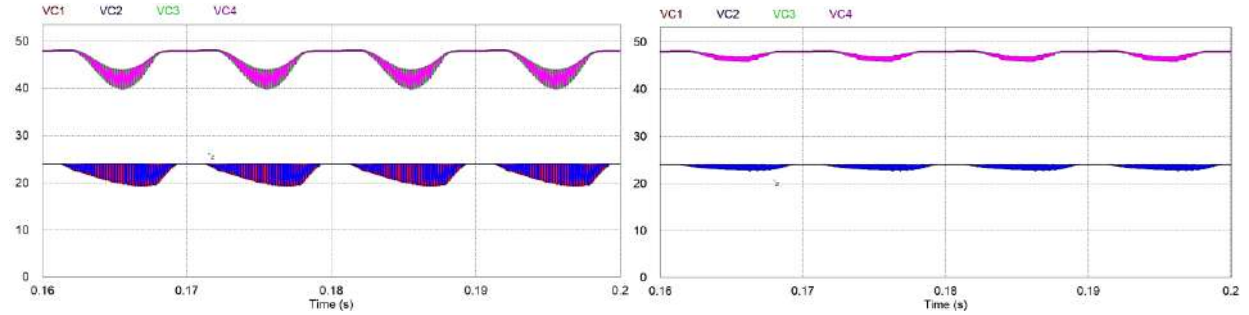


Figure 6. The voltage ripple of the capacitors (a) f_{sw}=5 kHz; (b) f_{sw}=20 kHz.

Figure and Figure show the simulation results of the proposed topology using the small capacitor 400 µF at the switching frequency of 5 kHz and 20 kHz. In Figure , although the THD value of the output voltage

is 12% for both cases, the waveform at 20 kHz is better. The reason for this case is that the voltage ripple on the capacitors is lower. Besides, the output current in both cases is a sine waveform, but THD at the 20 kHz switching frequency (1.18%) is better than the 5 kHz switching frequency (2.19%).
Figure presents the ripple voltage on the capacitors in both cases. At the 5 kHz switching frequency, the voltage ripple values of C_1 and C_2 are 4.5 V, and C_3 and C_4 are 8 V. At the 20 kHz switching frequency, the voltage ripple values of C_1 and C_2 are 1.3 V, and C_3 and C_4 are 1.9 V. Therefore, the ripple voltage at the 20 kHz switching frequency is improved.

3.2 Simulation result for capacitor value: 100 µF

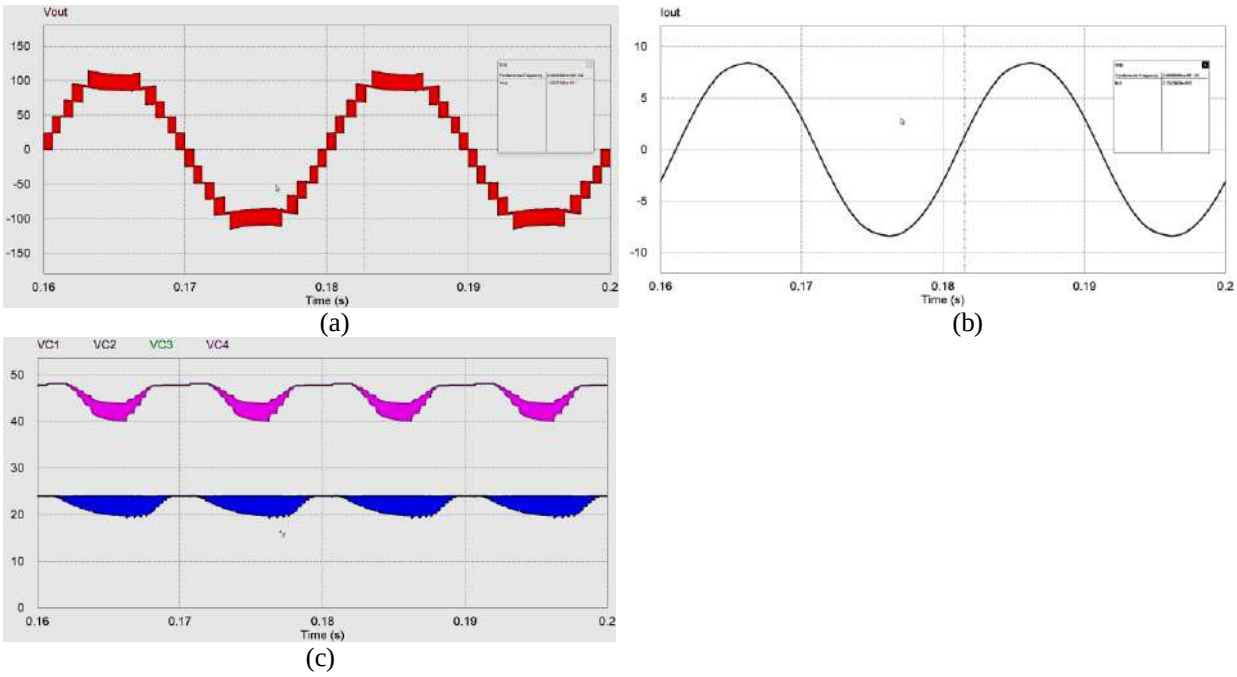


Figure 7. (a) The output voltage; (b) the output current; (c) The voltage ripple of the capacitors.

Figure 7 shows the simulation results of the proposed topology using the small capacitors of 100 µF at the switching frequency of 20 kHz. The output voltage is an eleven-level waveform with THD = 13%. The output current is the sine waveform with THD = 2.75%. Because using the small capacitors (100 µF), the ripple voltage of capacitors is higher (C_1 and C_2 are 4.7 V; C_3 and C_4 are 6.3 V). All the ripple voltage is summarized in Table .

Table 4. The ripple voltage in the different cases.

	5 kHz	20 kHz	
	400 µF	400µF	100 µF
C_1 and C_2	4.5 V	1.3 V	4.3 V
C_3 and C_4	8 V	1.9 V	6.3 V

4 CONCLUSIONS

The paper proposed a novel inverter using small capacitors which has several advantages such as reducing the inverter's size and improving its reliability. The proposed topology can reduce the ripple voltage of capacitors when increasing the switching frequency and it can run with the 100 µF capacitors at 20 kHz. The operation principle, as well as the control method, were shown in this paper. Moreover, the simulation results for the cases include f_{sw} = 5 kHz, C=400 µF; f_{sw} = 20 kHz, C = 100 µF; and f_{sw} = 20 kHz, C = 100 µF verified the operation of the proposed topology.

REFERENCES

[1] J. Rodriguez, J.-S. Lai, and F. Z. Peng, "Multilevel inverters: A survey of topologies, controls, and applications," IEEE Trans. Ind. Electron., vol. 49, no. 4, pp. 724–738, Aug. 2002.

[2] J. Rodriguez, S. Bernet, B. Wu, J. O. Pontt, and S. Kouro, "Multilevel voltage-source-converter topologies for industrial medium-voltage drives," IEEE Trans. Ind. Electron., vol. 54, no. 6, pp. 2930–2945, Dec. 2007.

[3] J. Rodriguez et al., "Multilevel Converters: An enabling technology for high-power applications," Proc. IEEE, vol. 97, no. 11, pp. 1786–1817, Nov. 2009.

[4] S. H. Hosseini, K. Varesi, J. F. Ardashir, A. A. Gandomi, and S. Saeidabadi, "An attempt to improve output voltage quality of developed multi-level inverter topology by increasing the number of levels," in 2015 9th International Conference on Electrical and Electronics Engineering (ELECO), 2015, pp. 665-669: IEEE.

[5] Nabae, A., Takahashi, I., Akagi, H.: 'A new neutral-point-clamped PWM inverter', IEEE Trans. Ind. Appl., 1981, IA-17, (5), pp. 518–523.

[6] Khazraei, M., Sepahvand, H., Corzine, K.A., et al.: 'Active capacitor voltage balancing in single-phase flying-capacitor multilevel power converters', IEEE Trans. Ind. Electron., 2012, 59, (2), pp. 769–778.

[7] Malinowski, M., Gopakumar, K., Rodriguez, J., et al.: 'A survey on cascaded multilevel inverters', IEEE Trans. Ind. Electron., 2010, 57, (7), pp. 2197– 2206.

[8] Y. Hinago and H. Koizumi, "A switched-capacitor inverter using series/parallel conversion with inductive load," IEEE Trans. Ind. Electron., vol. 59, no. 2, pp. 878–887, Feb. 2012.

[9] Y. Ye, K. W. E. Cheng, J. Liu, and K. Ding, "A step-up switched-capacitor multilevel inverter with self-voltage balancing," IEEE Trans. Ind. Electron., vol. 61, no. 12, pp. 6672–6680, Dec. 2014.

[10] Bac-Bien Ngo, Minh-Khai Nguyen, Jae-Hong Kim and Firuz Zare (2018) Single-phase multilevel inverter based on switched-capacitor structure, IET Power Electronics, vol. 11, no. 11, pp. 1858 – 1865.

[11] J. S. M. Ali and V. Krishnasamy, "Compact switched capacitor multilevel inverter (CSCMLI) with self voltage balancing and boosting ability," IEEE Trans. Power Electron., vol. 34, no. 5, pp. 4009–4013, May 2019.

[12] Y. Ye, S. Chen, X. Zhang, and Y. Yi, "Half-Bridge Modular Switched-Capacitor Multilevel Inverter with Hybrid Pulsewidth Modulation", IEEE Trans. Power Electron., vol. 35, no. 8, pp. 8237-8247, May 2020.

[13] Rodriguez, J., Lai, J.S., Peng, F.Z.: 'Multilevel inverters: a survey of topologies, controls, and applications, IEEE Trans. Ind Electron., 2002, 49, (4), pp. 724–738.

[14] Ma, M., He, X., Cao, W., et al.: 'Optimised phase disposition pulse-width modulation strategy for hybrid-clamped multilevel inverters using switching state sequences', IET Power Electron., 2015, 8, (7), pp. 1095–1103.

[15] Shi, X., Wang, Z., Tolbert, L., et al.: 'A comparison of phase disposition and phase shift PWM strategies for modular multilevel converters. Proc. IEEE Energy Conversion Congress and Exposition (ECCE'2013), Denver, Colorado, USA, 2013, pp. 4089–4096.

[16] Ren, L., Gong, C., He, K., et al.: 'Modified hybrid modulation scheme with even switch thermal distribution for H-bridge hybrid cascaded inverters', IET Power Electron., 2017, 10, (2), pp. 261–268.

ICATSD2F.120

BUILDING A QUESTION ANSWERING MODEL TO SUPPORT STUDENTS USING DEEP LEARNING

DO TAN HAI [1]*, DANG THI PHUC [1]*, TRAN CONG THINH[1], NGUYEN PHUC HUNG[2]

[1]*Faculty of Information Technology, Industrial University of Ho Chi Minh City*

[2]*Center of System Management, Industrial University of Ho Chi Minh City*

**dotanhaiit@gmail.com, phucdt@iuh.edu.vn, trancongthinh1999@gmail.com, nguyenphuchung@iuh.edu.vn*

Abstract. Currently, the demand for applying information technology to solve practical problems of social life is increasingly concerned. The advent of new technologies in the field of artificial intelligence and Deep Learning has created new possibilities supporting humans to solve problems that previous techniques could not resolve. Having determined these problems combined with the process of learning and doing research at the Industrial University of Ho Chi Minh City (IUH), in the paper, we will introduce the process of building an automatic answering system for students, which would use Deep Learning techniques to support automatic question answering related to regulation and admissions consulting. We trained the model based on a dataset of 20,000 pairs of questions and answers created following the IUH's as well as the Ministry of Education and Training's regulations. As reliable solutions for this problem, we propose two models: Sequence-to-Sequence combined with Attention and Transformer model. The training results show that the two models have achieved accuracy on the test set of 79% and 87%, respectively. Thanks to high accuracy and lightweight model architecture, the Transformer model was selected to deploy on the system. The system operations are based on the AWS platform which creates convenience to access, ensures question-answering functionality, updates the new sets of questions and answers, and improves the model over time.

Keywords. Deep learning, Question answering model, Sequence-to-Sequence-With-Attention, Transformer, Natural language processing.

1 INTRODUCTION

In the present status, the Industrial University of Ho Chi Minh City (IUH) has been deploying several websites on various platforms. These websites provide relatively complete information for students who are studying at the IUH. However, due to the complex organizational structure as well as the enormous scale of the university, the number of regulations is very large. As can be easily pointed out, this problem caused certain difficulties for learners while approaching the regulations of the university and the Ministry of Education and Training as well. In fact, a proportion of students had to contact their professors for support or look up on social media for frequently similar questions and problems.

In addition, through the process of studying, we noticed that the demand for admission counseling for new students wishing to register after finishing high school is sharply going up in number. In the current state, IUH also includes available communication channels for these students to exchange questions about the university's admissions; however, after taking a brief approach to these channels, we have found that the information exchange between students and the board only takes place commonly through asking and answering questions from students by professors or other students that are a part of groups or fanpages. Thus, this has resulted in different opinions on the matter of admissions counseling leading to the trouble that students who received counseling can make unwise decisions. Additionally, sending information via email to Admissions Counseling Department can positively provide more accurate results but it will take a lot of time for both students and professors in charge with simple and similar questions. In general, this causes an increase in workload but does not bring high efficiency.

In several previous pieces of research, there have been a couple of approaches to solving the question-answering problems. An outstanding example that could be observed is Bordes et al.'s method, which would

utilize a 'structured knowledge base' to answer simple questions [1]. This approach depends mainly on man-made rules or those from previous experiments, which do not bring high efficiency to processing big data. In contrast, the deep learning method brings has remarkably effective results in extracting highlighted features of the question and selecting flexible answers in a diversity of contexts [2]. Another illustration that should not be overlooked is that Yih et al.'s method, the application of the 'bag-of-word' to represent words, then through the CNN model to extract information [3]. In one other report, to improve the extraction of syntactic and semantic relationships of the sentences, Modak et al.'s have combined RNN with the Attention mechanism, which has brought several positive solutions for the NLP problems [4]. Another research by Shao et al.'s also offered the Transformer model, which has a lighter network architecture, to unfold the problem without negatively affecting its efficiency [5]. In fact, there have been several solutions to the mentioned problem; one of the most common ways can be observed is using chat-bots by developers of popular platforms such as Facebook, WordPress, Zalo, etc [6]. However, this solution is facing a challenge that the data required to respond to numerous different contexts and scenarios need to be in preparation, which is considered infeasible in practice. The system not applied with Deep Learning hardly responds to questions of relatively high difficulty and flexible context. With the desire to apply an application for the acquired knowledge of information technology, especially in the field of Deep Learning to solve the practical problem as well as contribute a part to the development of IUH, we have proposed to use Deep Learning techniques to solve the question-answering problem.

In the paper, we have proposed to use Deep Learning techniques to solve the question-answering problem using advanced models Sequence-to-Sequence based on RNN combined with Attention, Transformer model, and other techniques in natural language processing will be utilized to process data during implementation. Meanwhile, we also aim to build a website that not only allows users to make several questions about current regulations of the IUH as well as the Ministry of Education and Training but also automatically responds to the enrollments. In addition, the system also includes fundamental statistical functions, especially in supporting the automatic creation of datasets to help the model easily respond to new requirements and changes.

2 RESEARCH METHODS

2.1 Related algorithms

In this paper, we use Deep Learning technique including the application of recurrent neural network (RNN). In addition, to find the dependence of the components in the sentence, we apply the Attention mechanism to the Sequence-to-Sequence model based on RNN.

Recurrent Neural Network: For the given problem, the RNN model [7] is used to store information as well as the relationship among the elements in the input sequence. RNN is referred as recurrent networks because RNN performs the same task for all elements of a sequence whose output depends on previous calculations. In other words, RNN can store previously computed information. Theoretically, the information of an overlong text can be stored by RNN; however, RNN can only practically retain information from a few previous steps due to the vanishing gradient problem that appears when transmitting long distances in the network. To overcome this problem, LSTM (Long Short-Term Memory), GRU (Gated Recurrent Unit), and BiLSTM (Bidirectional LSTM) are commonly known as support models for calculating and storing important information about the input sequence. In this document, we propose to use the GRU network model to store and filter the information using update and reset ports. Moreover, to increase the efficiency of the processing of complex input sequences, we use the Attention mechanism to orient the attention to important elements of the input sequence.

Attention Mechanism: Attention means directing concerns to something and concentrating on that. The Attention mechanism [8] in Deep Learning is based on the mentioned concept of Attention to certain factors when processing data. The most significant benefit of this mechanism is that it enables the possibility to select vital information from the input string and store these weighted values for the decryption process. The Attention mechanism will be specifically described in the following problem.

Transformer Model: The transformer model [8] was introduced in 2017 after the scientific paper "Attention Is All You Need" had been published by Google. Like previous recurrent networks, the Transformer model is designed to process sequential data for natural language processing problems.

However, unlike the previous mechanism, the Transformer model allows processing inputs simultaneously; thereby greatly improving the model training performance as well as solving the problem of data loss when transmitting distant ranges in the network model. The application of the Transformer model to resolve the problem will be described in detail later in this paper.

2.2 Model for question-answering problem based on regulations

In the process of researching Deep Learning techniques for question-answering problems, we propose to use two processing models: Sequence to Sequence with Attention and the Transformer model to solve the problems we built-up.

2.2.1 Sequence-to-Sequence with Attention

Sequence-to-Sequence with Attention [9] is based on the architecture of RNN. However, there must have several inheritance and improvements to overcome the limitations of the fundamental Sequence-to-Sequence model when Attention is not combined; in which, important information when transmitting through different network states is lost. By applying the Attention mechanism, vital information of the input sequence can be retained through the method of storing these weighted values in each hidden state of the Encoder, then synthesizing and transferring those values to the Decoder along with Context Vector, which were encoded when passing through hidden states of RNN is the LSTM or GRU in the Encoder. The Decoder will rely on this information to predict words in the output string. To be more specific, this plays a significant role in choosing words that match the input context to help output match with input string content.

In this paper, we propose to use Bahdanau Attention for Sequence-to-Sequence model combined with Attention [10]. To calculate the score between all hidden states of the Encoder and the output of the previous step of the Decoder, we use the following formula:

$$score(h_t, \bar{h}_s) = v_a^T tanh(W_1 h_t + W_2 \bar{h}_s) \quad (1)$$

where:

h_s- All the hidden states of the Encoder; h_t- Previous hidden states of the Decoder; W- Weighting matrix for parameterizing the calculations, tanh- Activation function, v- Another single hidden layer network model. Below is Sequence-to-Sequence with Attention that we implemented.

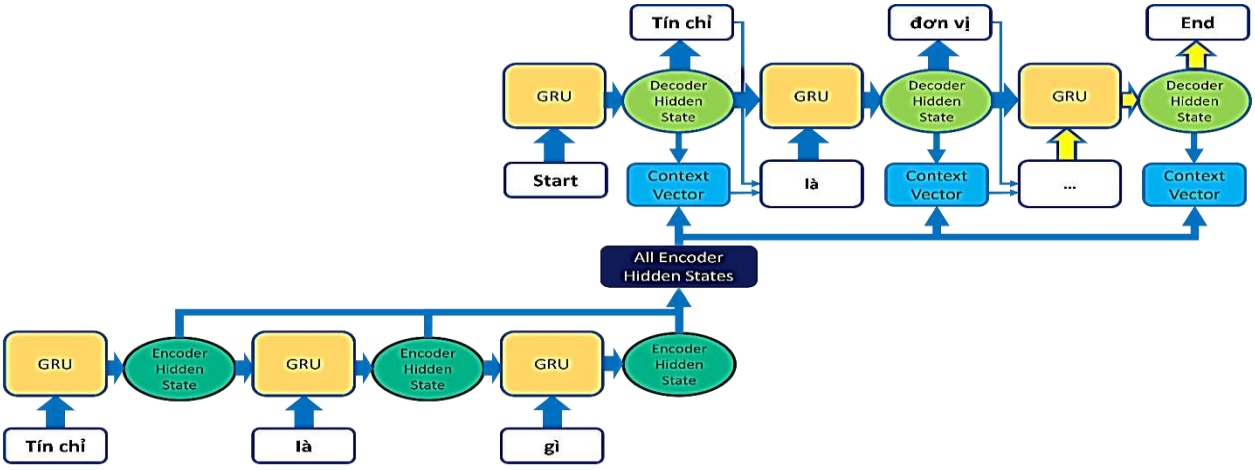


Figure 1. Sequence-to-Sequence with Attention for question-answering problem

2.2.2 Transformer Model

For the Transformer model [8,11], a solution mechanism is required as the input is processed concurrently while keeping the components order and relationship unchanged in the input sequence at the Encoder. To solve the above problem, the Transformer model not only uses Embedding but also applies the Positional Encoding method. For this method, the positions of the words are encoded by a vector whose size equals Word Embedding and they are added directly to Word Embedding. This method is defined as follows [8]:

$$p_{i,j} = \begin{cases} \sin\left(\frac{i}{10000^{\frac{j}{d_{emb\_dim}}}}\right) \; if \; j \; is \; even \\ \cos\left(\frac{i}{10000^{\frac{j-1}{d_{emb\_dim}}}}\right) \; if \; j \; is \; odd \end{cases} \quad (2)$$

where:

i – position of the word in the sentence; j – the dimension index in Positional Encoding; d_{emb_dim}- number of dimensions of Positional Encoding.

The Transformer model also applies the Attention mechanism specifically Self-Attention. By this, the Transformer model can understand the relationship of the components in a sentence. Self-Attention allows a model when encoding a word can use information about the words associated with it. Self-Attention can be calculated as the following formula [8]:

$$\mathrm{Z} = \mathrm{Attention}(\mathrm{Q}, \mathrm{K}, \mathrm{V}) = \mathrm{Softmax}\left(\frac{\mathrm{QK^T}}{\sqrt{\mathrm{dk}}}\right) * \mathrm{V} \quad (3)$$

Instead of calculating Self-Attention only once, a number of Self-Attention (Head) are created with different initialized weights to learn a diverse variety of aspects of the input sequence, which is called Multi-Head Attention. In the paper, we conduct experiments with 8 Heads in Encoder. Multi-Head Attention can be calculated as the following formula [8]:

$$\mathrm{Z} = \mathrm{Attention}(\mathrm{Q}, \mathrm{K}, \mathrm{V}) = \mathrm{Concat}(\mathrm{Z_0}, \mathrm{Z_1}, \ldots, \mathrm{Z_n}) * \mathrm{W_0} \quad (4)$$

In the Transformer model, to avoid the problem where information can be lost when traversing through different layers, we also apply Residuals Connection as well as ResNet networks. To be clearer, Residuals Connection and Normalization Layer are well-known in the architecture of the Transformer model. In general, both techniques benefit in making the training model faster in converging and preventing information loss. The Z vectors after being 'Normalized' will transfer to a Fully Connected Network called Feed Forward Neural Network to extract latent features once more time before performing Residuals Connection and switching to Decoder. In this paper, we conduct experiments with 4 the Encoder layers similar and sequence. Alternatively, we also built the Feed Forward Neural Network have input and output size at d_{model} = 512 as well as number of nodes in Hidden Layer is d_{ff} = 2048. The formula of Feed Forward Neural Network is calculated as follows:

$$\mathrm{FFN(x)} = \max(0, \mathrm{xW_1} + \mathrm{b_1})\mathrm{W_2} + \mathrm{b_2} \quad (5)$$

where: W_1, W_2 are weighting matrices.

Basically, the Decoder model is built the same as the Encoder; however, there will be an extra layer of Masked Multi-Head Attention. The next layer in the Decoder is Multi-Head Attention which receives both outputs from the Encoder and Masked Multi-Head Attention Layer; after being "Residual & Normalized" in the Decoder, these outputs are the Q, K, and V matrices. After having the set of those matrices, the Attention calculation will be performed similarly to the Encoder. In this paper, we conduct experiments with 4 the Decoder layers similar and sequence. After going through 4 layers of the Decoder, data will be pushed into Linear Layer and projected into a dimension with the length of vocab_size, then we calculate Softmax to get log_probs. After getting the values from log_probs, the task is to get the maximum-valued element on log_probs and return the received value to the word corresponding to the index in the dictionary. The Transformer model for the question-answering problem about the regulations is described in Figure 2:

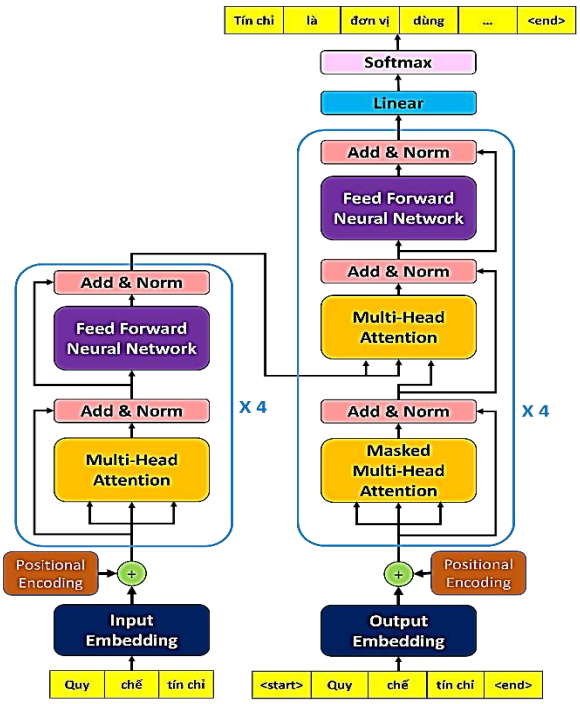


Figure 2. Transformer model for question-answering problem

2.3 Model Evaluation

Among the models that we implement, Sequence-to-Sequence with Attention help extracts and store important information from the input sequence. Otherwise, the Transformer model applies a new architecture compared to the traditional RNN, which provides the ability to concurrently process components of the input sequence; thereby significantly increasing the efficiency of the training model in both speed and accuracy.

To evaluate accuracy of models, we perform an evaluation based on BLEU Scorewith the following formula[12]:

$$BLEU = \underbrace{min\left(1, exp\left(1 - \frac{reference - length}{output - length}\right)\right)}_{\text{Brevity penalty}} \underbrace{\left(\prod_{i=1}^{4} precision_i\right)^{1/4}}_{\text{N-gram overlap}}$$

$$\text{precision}_i = \frac{\sum_{snt \in Cand-Corpus} \sum_{i \in snt} \min\left(m^i_{cand}, m^i_{ref}\right)}{\omega^i_t = \sum_{snt' \epsilon Cand-Corpus} \sum i' \epsilon snt, m^{i'}_{cand}} \quad (6)$$

where:

m^i_{cand} - the count of i-gram in the candidate translation matching the reference translation; m^i_{ref} - the count of i-gram in the reference translation; m^i_t- the total number of i-grams in the candidate translation

The formula consists of two parts:

+ *Brevity Penalty* penalizes generated translations that are too short compared to the closest reference length with an exponential decay. The brevity penalty compensates for the fact that the BLEU score has no recall term.

+ *N-Gram Overlap* counts the number of unigrams (1-gram), bigrams (2-grams), trigrams (3-grams), and four-grams (4-grams), matching their n-gram counterpart in the reference translations. This term acts as a precision metric. Unigrams account for the adequacy while longer n-grams account for the fluency of the translation. To avoid over-counting, the n-gram counts are clipped to the maximal n-gram count occurring in the reference m^n_{ref}.

BLEU scores are explained in Table 1.

Table 1. Interpretation of BLEU scores

BLEU Score	Description
< 10	Nearly has no reference to the gist
10 - 19	Unfocused on the gist
20 - 29	The gist is relatively clear, but numerous significant grammatical errors
30 - 40	Understandable quality translations
40 - 50	High quality translations
50 - 60	Very high quality, adequate, and fluent translations
> 60	The performance is almost as good as humans

3 RESULTS AND DISCUSSIONS

3.1 Dataset

Our dataset includes 20,000 data about the regulations needed to be universal to students at the IUH. In the process of creating the dataset, we decided to manually create a complete set including 13,650 questions-answer pairs about regulations of IUH and the Ministry of Education and Training. The rest include 6,350 data related to regulations and activities of the Information Technology Faculty at the IUH, we had an outstanding move in combining two methods of manual generation and an automatic data generator which was developed by our team.

Table 2. List of regulations used

Regulation Code	Name
MC1.1	Regulations of the National High School Exam
MC1.2	Announcement about admission to university and college
MC1.3	Announcement about quality assurance threshold entrance admission
MC1.4	Regulations on admission to universities and colleges
MC1.5	Announcement on requiring enrollment scores
MC1.6	Planning on separation of majors
MC1.7	Regulations on credit training of IUH
MC1.8	Regulations on assessment of training scores for regular university students
MC1.9	Regulations for the professors in charge of classes
MC1.10	Announcement of graduation examination plan

For each regulation, we filtered out the answer that can provide the correct information. Corresponding to each answer, there will be approximately 50 different questions. Table 3 briefly describes the dataset with several questions linked to one answer.

For the training process, we proceed to divide the dataset into 3 parts: 12,800 data are used for the training of the model, 3,200 data for checking the model validation throughout the training process, and the rest 4,000 data are used for the testing of the model after training.

Table 3. Illustration example for asking and answering pairs in the dataset

Question	Answer
Các sinh viên trúng tuyển cần nộp đơn nhập học ở đâu để làm các thủ tục liên quan đến quy trình nhập học? (*Where do admitted students have to submit their applications to fulfill the admission process's procedures?*)	Sinh viên trúng tuyển nộp đơn nhập học về phòng công tác chính trị và công tác sinh viên (*Admission students submit their applications to the Office of Political Affairs and Student Affairs*).
Em vừa nhận được thông báo trúng tuyển, em còn băn khoăn không biết sinh viên trúng tuyển nộp đơn nhập học ở đâu? (*I have just*	Sinh viên trúng tuyển nộp đơn nhập học về phòng công tác chính trị và công tác sinh viên (*Admission students submit their*

received the announcement of admission, so I wonder where to submit my application as an admitted student?) Vui lòng cho em biết, khi đến đăng ký nhập học tại trường, em nên làm việc gì trước tiên và liên hệ với phòng/ban nào trước để nhận được hỗ trợ? (*Please guide me through the registration process, what should I prioritize in this process, and also which department/office should I contact?*)	*applications to the Office of Political Affairs and Student Affairs*). Sinh viên trúng tuyển nộp đơn nhập học về phòng công tác chính trị và công tác sinh viên (*Admission students submit their applications to the Office of Political Affairs and Student Affairs*).
Mình là học sinh đến ở một tỉnh cách xa đến Thành phố Hồ Chí Minh để đăng ký nhập học vào trường. Mình vừa vượt qua kỳ thi đầu vào của trường, do đó, mình cần phải hoàn thành thủ tục nhập học. Bạn có thể cung cấp cho mình biết mình cần làm những gì được không? *(I am a student from a distant province who goes to Ho Chi Minh City to enroll in our university. I have already passed our university's entrance exam; therefore, I need to complete my admission procedures. Can you provide me with what I need to do?*)	Sinh viên trúng tuyển nộp đơn nhập học về phòng công tác chính trị và công tác sinh viên (*Admission students submit their applications to the Office of Political Affairs and Student Affairs*).
Theo cơ cấu tổ chức hiện nay của trường Đại học Công nghiệp Thành phố Hồ Chí Minh, văn phòng nào chịu trách nhiệm tiếp nhận đơn nhập học sau khi có kết quả trúng tuyển của học sinh? (*According to the current organizational structure at the Industrial University of Ho Chi Minh City, which office is responsible for receiving applications after having admitted students' results?*)	Sinh viên trúng tuyển nộp đơn nhập học về phòng công tác chính trị và công tác sinh viên (*Admission students submit their applications to the Office of Political Affairs and Student Affairs*).

3.2 Implementation Details

For the Sequence-to-Sequence with Attention, the input sequence is pushed into the Encoder after Embedding with a dimension of 300. Next step, these data will transmit through several GRU layers with 512 hidden states to extract the information of the sequence. At each hidden state, the calculation with Bahdanau Attention will be performed before transmitting to the next layers. Finally, data is pushed into the Decoder to decode the output sequence from the Encoder.

For the Transformer model, the input string after Embedding with a dimension of 256 will perform Positional Encoding with the dimension of 512. After completing this step, data is transmitted through 4 Encoder layers with 8 Multi-Head Attention. After that, data will perform few steps as section described above to get the output at the Encoder. Finally, the output of the Encoder is pushed into the Decoder to perform decoding, which also is described above.

Implement model: After researching through different platforms, we decided to use Google Colab Pro to take complete advantage of GPU hardware.

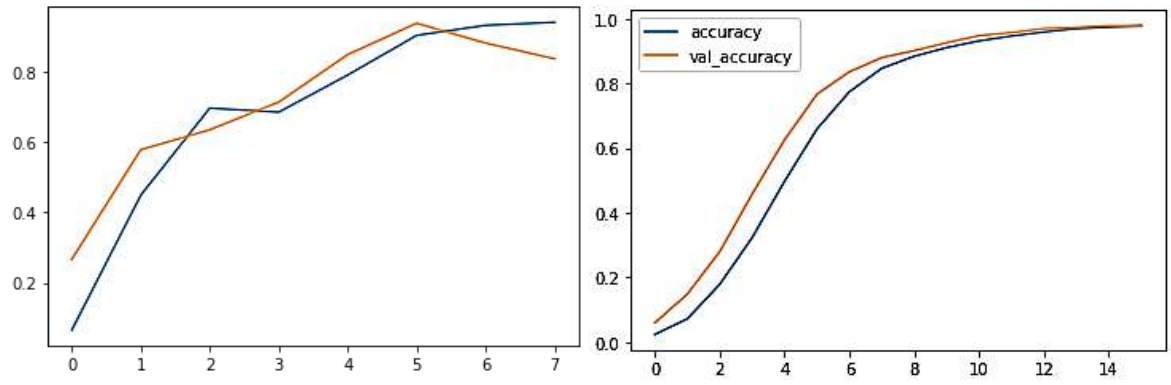
For the Sequence-to-Sequence with Attention implemented on PyTorch library, we do it with the model hyperparameters proceeded as follows: the number of GRU units in Encoder and Decode is 512, the drop rate is 0.25, epoch is 8, and the learning rate is adjusted during model training.

For the Transformer model, we decided to choose the number of layers for both Encoder and Decoder is 4, not only to help reduce in size and complexity of the model but to maintain the accuracy as well. In this model, we choose the model hyperparameters as follows: $d_{model} = 256$, $d_{ff} = 512$, the drop rate is 0.05, batch size is 64, epoch is 16, and the learning rate is adjusted during model training.

3.3 Results and Analysis

3.3.1 Evaluation of model accuracy

Overall, the training results of Sequence-to-Sequence with Attention and the Transformer model are described in Figure 3. Accuracy results on training dataset and testing dataset are described in Table 4. The results indicate that the Sequence-to-Sequence-with-Attention (Figure 3a) is not really stable and accurate. The accuracy is not high. To be more specific, there is a difference about the accurate between the training set and the testing set, which leads to the problem called "overfitting". Otherwise, the Transformer model (Figure 3b) has better results in significant features: stability, convergence, and high accuracy.



a) b)

Figure 3. Graph of model's accuracy on the training set and validation set; a) Sequence-to-Sequence with Attention, b) Transformer Model

Table 4. Model accuracy results based on data set

Model	BLEU Score on training dataset	BLEU Score on testing dataset
Sequence-to-Sequence with Attention (PyTorch)	0.88	0.79
Transformer	0.99	0.87

In which, the Transformer model has a lightweight network architecture; as a result, in and Table 5, the required training time is reduced to take up an amount of 57 minutes compared to 127 minutes of Sequence-to-Sequence with Attention model. Therefore, the Transformer model architecture is deployed to the system, the operation will have greater efficiency than applying Sequence-to-Sequence with Attention model architecture.

Table 5. Statistics of model execution time

Model	Epochs	Training time (mins)
Sequence-to-Sequence with Attention (PyTorch)	8	127
Transformer	16	57

3.4 Deploying the model on application

3.4.1 System architecture

After the process of carefully observing, testing, and evaluating the experimental results between two models, the Transformer model is chosen to deploy on the system. During implementation process, we utilized the support from libraries and frameworks such as: Flask, Vue.js, Docker, ... to build FrontEnd and BackEnd of the application. The implementation model is shown in Figure 4.

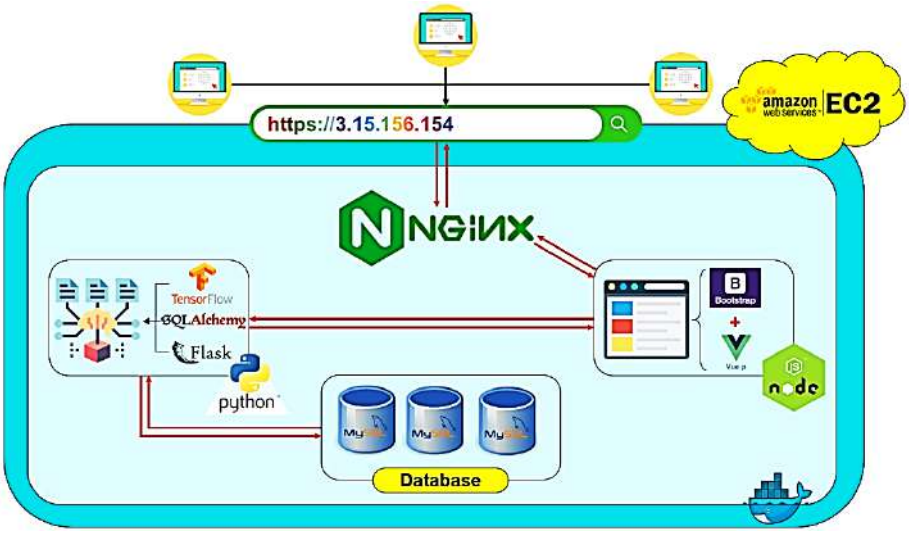


Figure 4. General model of application

3.4.2 Practical application

Based on the general model, we implemented the system's functions according to our theoretical model, then tested its operation on the Amazon Web Services platform. The experimental results of the practical application we achieved can be detailedly listed beneath.

The system includes the following functions: log-in, automatic question-answering, evaluating the answers, registering an account, viewing statistics of the application, updating personal information, changing password, managing reviews, deleting reviews, exporting reviews to CSV file, and generating data from template files.

In which, the automatic question-answering function allows students who are studying at IUH or wishing for admission to the university can ask questions about IUH regulations as well as the Ministry of Education and Training.

Figure 5. The display of the application while using the asking function

Figure 6. The automatic answering function displayed in the application

In addition, we have integrated the function of evaluating answers after the user asks a question to improve the quality of answers. With this function, users can rate their satisfaction on a scale from 1 to 5. Meanwhile, users can also enter their comments as a contribution to improve the quality of answers as well as complete the system.

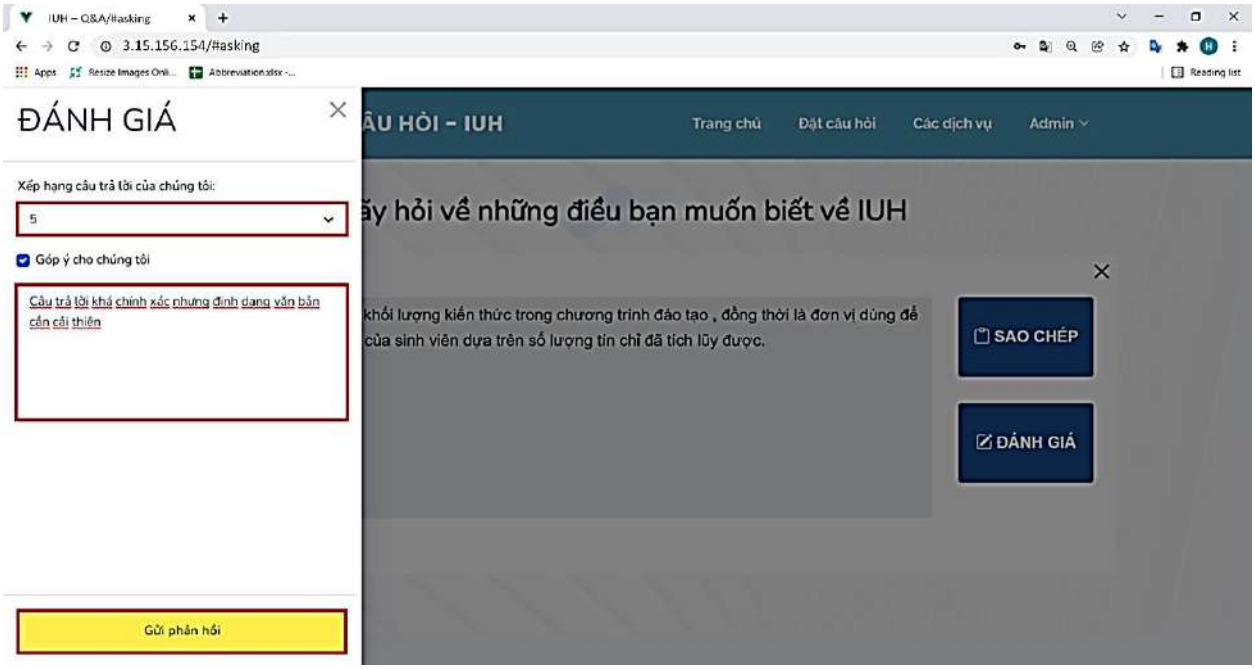


Figure 7. The display of the application during the rating answer

The statistics function will present statistics about the number of utilizations, charts showing the distribution of reviews, and application ratings, which benefit both users and administrators in apprehending the overall website.

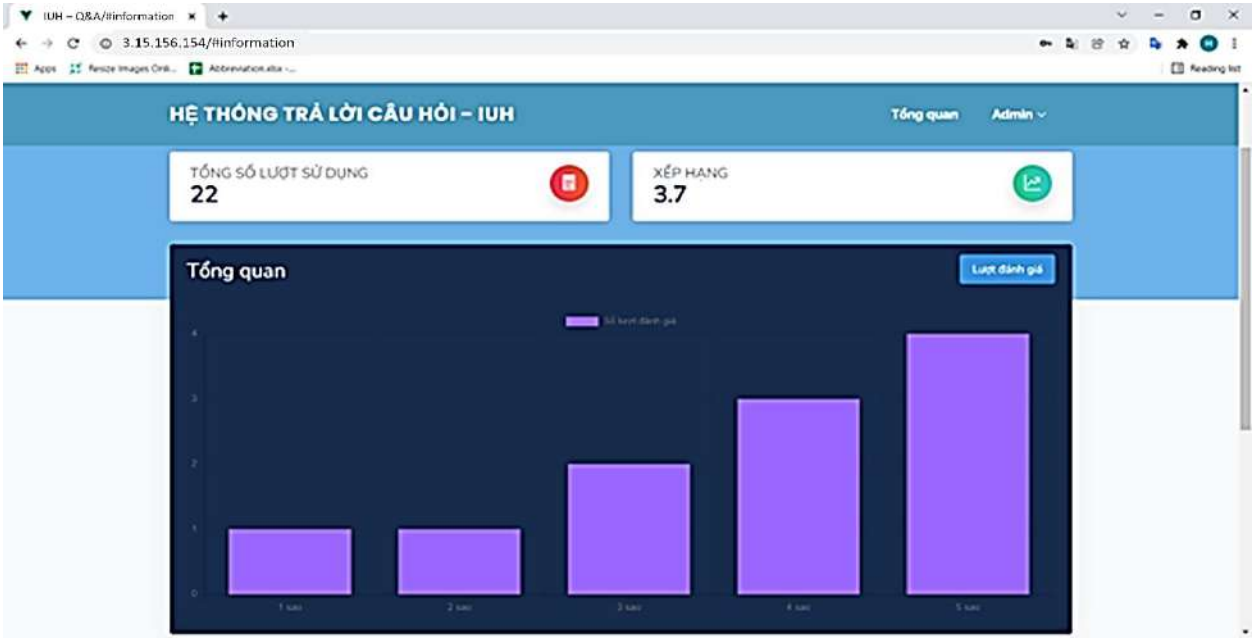


Figure 8. Statistical charts and information provided in the application

Moreover, another important function of the system is generating data to update the model based on the new contributions or new questions. These inputs by users require additional answers for the purpose of re-training the model or updating new data for the model toward the new questions which not have answers that exist in the dataset. This functionality is designed to respond to constantly changing requirements of regulations that always require the model to be updated quickly to meet the demand for the new data.

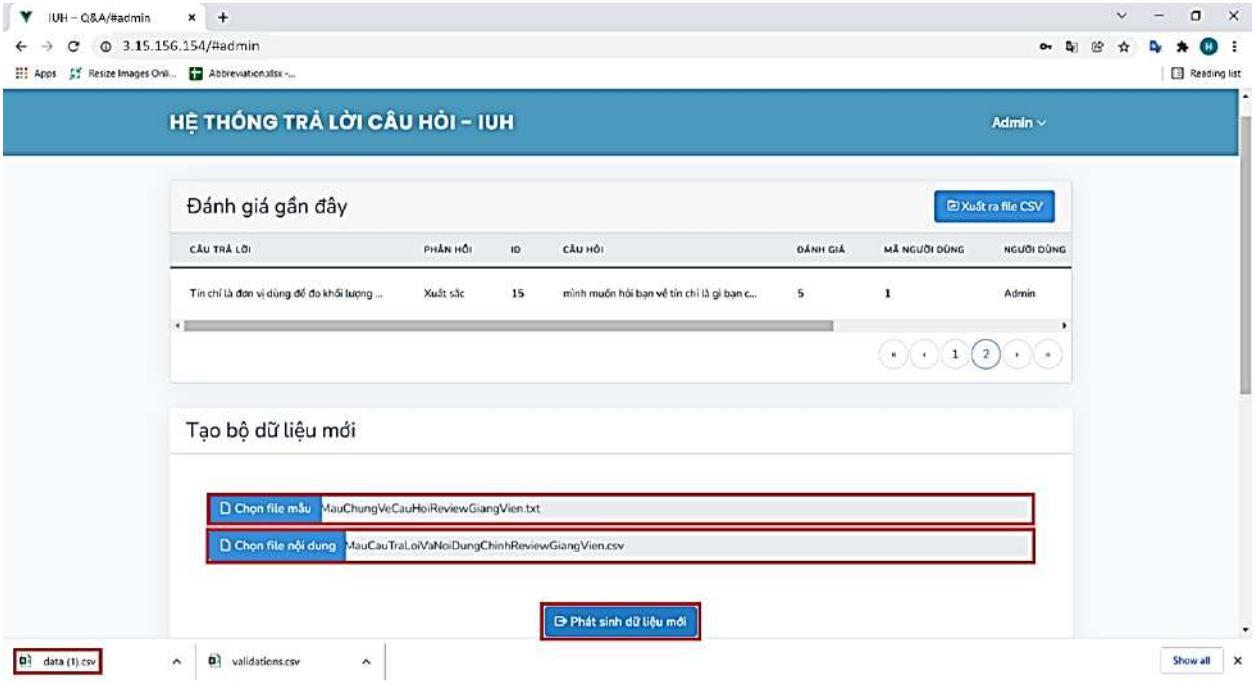


Figure 9. The display of the application while using the generating a dataset function

4 CONCLUSIONS

In this paper, we built an automatic system to answer questions for students at the Industrial University of Ho Chi Minh City using the Deep Learning technique. The experimental results collected during several different tests have shown that the Transformer model outperformed the Sequence-to-Sequence with

Attention both in accuracy and stability. Besides, since Sequence-to-Sequence with Attention is based on RNN architecture, the training process takes up a relatively long amount of time. Also, because this network architecture is quite heavy, it requires a lot of GPU hardware for the training process. In contrast, the Transformer model applies a novel architecture that greatly shortens the training time as well as reduces the hardware required during the training process. Therefore, the Transformer model is selected to deploy on the question-answering system. However, any Deep Learning problem facing a persistent challenge is that the requirement of having large datasets goes along with the constant expansion in complexity and quantity of data. This problem promotes us to catch the changing requirements by upgrading and completing the automatic data generation tool; thereby, we can accomplish the purpose of getting the application to fit into different actual situations of the IUH. In fact, that is also one of our further research directions in the question-answering problem. Finally, we will also research new techniques and architectures that can be applied to this problem as our main purpose is continuous to improve the system quality.

REFERENCES

[1] A. Bordes, N. Usunier, S. Chopra and J. Weston, "Large-scale Simple Question Answering with Memory Networks," in *CoRR abs/1506.02075*, 2015.

[2] S. Yashvardhan and G. Sahil, "Deep Learning Approaches for Question Answering System," in *Proceedings of International Conference on Computational Intelligence and Data Science*, Vol. 132, 2018, pp. 785-794.

[3] W. Yih, X. He and C. Meek, "Semantic Parsing for Single-Relation Question Answering," in *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics* (Volume 2: Short Papers), Baltimore, Maryland, 2014, pp. 643–648.

[4] S. Modak, S. Chaudhury, A. Rawat and S. Deb, "Improving performance of Recurrent Neural Networks for Question-Answering with Attention-based Context Reduction," in *2021 IEEE Mysore Sub Section International Conference (MysuruCon)*, 2021, pp. 723-728, doi: 10.1109/MysuruCon52639.2021.9641626.

[5] T. Shao, Y. Guo, H. Chen and Z. Hao, "Transformer-Based Neural Network for Answer Selection in Question Answering," *IEEE Access*, vol. 7, pp. 26146-26156, 2019, doi: 10.1109/ACCESS.2019.2900753.

[6] A. Trivedi, V. Gor and Z. Thakkar, "Chatbot generation and integration: A review," International in *Journal of Advance Research Ideas and Innovations in Technology*, vol. 5, no. 2, pp. 1308-1311, 2019.

[7] Y. Yu, X. Si, C. Hu, and J. Zhang, "A review of recurrent neural networks: Lstm cells and network architectures," in *Neural Comput*, pp. 1235–1270, 2019. https://doi.org/10.1162/neco_a_01199

[8] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, et al., "Attention is all you need," *Advances in neural information processing systems*, pp. 5998-6008, 2017.

[9] T. Luong, H. Pham, and C. D. Manning, "Effective Approaches to Attention based Neural Machine Translation," in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, Lisbon, Portugal: Association for Computational Linguistics, pp. 1412–1421, 2015.

[10] D. Bahdanau, K. Cho, and Y. Bengio, "Neural Machine Translation by Jointly Learning to Align and Translate," in *ICLR*, 2014. https://arxiv.org/abs/1409.0473

[11] S. Khan, M. Naseer, M. Hayat, S. W. Zamir, F. S. Khan and M. Shah, "Transformers in vision: A survey," in *ACM Computing Surveys (CSUR)*, 2021.

[12] K. Papineni et al., "Bleu: a method for automatic evaluation of machine translation," in *Proceedings of the 40th annual meeting on association for computational linguistics*, 2002, pp. 311–318.

ICATSD2F.121

AN EFFICIENTLY METHOD DETERMINATION THE REACHABLE SET OF GEOSPATIAL DATA IN NETWORK SPACE

TRANG T.D. NGUYEN[1,2], LOAN T.T. NGUYEN[3,4,*], L.N. DUY[1]

[1]*Faculty of Information Technology, Industrial University of Ho Chi Minh City, Ho Chi Minh City, Vietnam.*

[2]*Faculty of Information Technology, Nha Trang University, Nha Trang, Vietnam.*

[3]*School of Computer Science and Engineering, International University, Ho Chi Minh City, Vietnam.*

[4]*Vietnam National University, Ho Chi Minh City, Vietnam.*

*21142301.trang@student.iuh.edu.vn, *nttloan@hcmiu.edu.vn, lenhatduy@iuh.edu.vn*

Abstract. Now, data is not only increasing in volume and velocity, but also in variety. In addition to structured data, unstructured and semi-structured data is increasingly common, including geospatial data. Clustering is an important technique in geospatial analysis. The main algorithm of the geospatial clustering algorithm finds the neighborhood set of a point. The problem of finding a reachable set is related to the problem of finding a neighborhood. This paper proposes a solution to find the reachable set according to the MapReduce model.

Keywords. MapReduce, Neighborhoods, Network Spatial Analysis, Reachability, Spatial Clustering.

1 INTRODUCTION

Today, the advent and increasing popularity of devices that can collect data anytime, anywhere has contributed to accelerating the growth of data. This rapid development of data is one of the important characteristics that led to the formation of the concept of "Big Data". Living in a world that is not only "flat" but also needs to be "fast" but this excess of data and lack of information has posed many challenges for us. How can we quickly extract information from that huge data block to promptly meet human needs? This paper will present MapReduce model - one of the models for big data - for the problem of finding the reachable set for from a point on the graph. The article implements the solutions of this model for a computer according to non-parallel, parallel, and improved solutions. Experimental results show that the parallel solution, and especially the improved parallel solution, has greatly increased the processing speed. Some of the main contributions are presented following:

- Experimenting the algorithm to determine Reachability according to MapReduce model on a computer.
- Based on the experimental results, propose solutions to improve the algorithm.
- Experiment and compare the original algorithm with the improved algorithm, parallel and non-parallel.
- Propose some more solutions to further increase processing speed in the discussion.

The rest of this paper is structured as follows. Section 2 discusses related work. The technical details for the reachability are presented in Section 3. The experimental results and evaluations of algorithms will be presented in Section 4. Some discussion is in Section 5 and finally the conclusion.

2 RELATED WORKS

The problem of finding the reachable set is related to the problem of finding a neighborhood. This paper proposes a solution to find the reachable set according to the MapReduce model. The reachability set of mode u is defined as the union of all neighborhoods of any radius for a node u. Therefore, the reachable set for u is determined by determining its neighborhoods with the smallest radius d where $N(u, d) = N(u, d + 1)$.

Neighborhood construction models are important in demonstrating interrelations between the data points. Therefore, determining the neighborhood of a data point is a problem that always challenges the scientific community to constantly find effective improvement solutions.

So far, there are many methods to find the neighborhood like K-Nearest Neighbor (k-NN) [1], kNN-IS [2], K-nearest neighbor based structural twin support vector machine (KNN-STSVM) [3]… The result of most of the proposed algorithms depends on the input parameter, the value of k. One of the problems with these approaches is fixing the value of k for the entire data set. Obviously this is not efficient because spatial data regions with different characteristics will need different values of k, so García-Pedrajas et al proposed a solution to set the local value of k [4].

The work [5] determines the neighborhood according to the input parameter epsilon radius, the result depends on the epsilon value. To overcome this limitation, [6] have proposed the solution of forming a neighborhood using Apollonius circle [7].

Here are some related definitions:

Definition 1. **Directed graph** *[8]***.** Directed graph has a set of nodes and a set of arcs. The arc from u to v is a pair of nodes: $u \rightarrow v$.

Definition 2. **Undirected graph** *[8]***.** Undirected graphs can be represented by directed graphs. $Edge(u, v)$: use two arcs $u \rightarrow v$ and $v \rightarrow u$.

Definition 3. **Path** *[8]*. A path in a directed graph is a sequence of nodes v_0, v_1,. . . , v_k such that there exist arcs $v_i \rightarrow v_{i+1}$ for all $i = 0, 1, \ldots, k - 1$.

A path with $k + 1$ nodes has a length of k, which is the number of arcs the path passes through.

A path with a node has a length of 0.

Definition 4. **The neighborhood of radius d** *[8]*. The neighborhood of radius d of node v is the set of nodes u whose path with the greatest length from v to u is d.

Notation: $N(v, d)$.

Example:

$N(v, 0)$ is $\{v\}$,

$N(v, 1)$ is v along with the set of nodes with an arc from v.

$N(V, d)$: the set of nodes u has a path of length ≤ d from at least one node in V.

Consider an undirected graph:

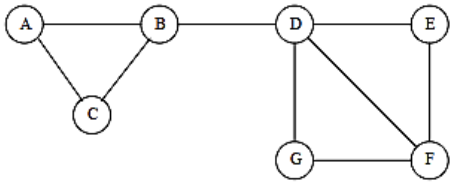


Figure 1. Undirected graph [8]

Convert to a directed graph by treating each edge as a pair of arcs, each in a direction.

$Edge(A, B)$ is represented by two arcs $A \rightarrow B$ and $B \rightarrow A$.

Consider the neighborhoods of node A.

$N(A, 0) = \{A\}$.

$N(A, 1) = \{A, B, C\}$, because there are only arcs from A to B and C.

$N(A, 2) = \{A, B, C, D\}$.

$N(A, 3) = \{A, B, C, D, E, F, G\}$.

Neighborhoods for larger radius are all the same as $N(A, 3)$.

Definition 5. **The Diameter of a Graph** *[8]***.** The diameter of a directed graph is the smallest integer d such that any two nodes u and v have a path of length not exceeding d from u to v.

Example: The graph in Figure 1 has a diameter of 3.

Definition 6. ***T*****ransitive closure** *[8]*. A transitive closure of a graph is the set of pairs of nodes (u, v) such that there is a path from u to v of length 1 or more.

Notation: $Path(u, v)$ related to the concept of reachability.

We say that node u reaches (reaches) node v if $u = v$ or $Path(u, v)$ is true.

The problem of transitive closure is to find all pairs of nodes u and v in a graph for which $Path(u, v)$ is true.
Calculates $N(v, \infty)$, the set of nodes accessible from node v by computing the sequence of neighborhoods, which is essentially breadth-first search of the graph from v.

3 REACHABILITY

3.1 The reachability problem [8]

For a node u in the graph, find all $v \neq u$ such that $Path(u, v)$ is true. Given a node u in the graph, find all $v \neq u$ such that there is a path from u to v.

Definition: $N(u, \infty) = \cup_{i \geq 0} N(u, i)$.

Therefore, the reachability problem is the union of all neighborhoods of any radius for a node u.

The reachable set for u is determined by determining its neighborhoods with the smallest radius d where $N(u, d) = N(u, d + 1)$.

3.2 Reachability Via MapReduce [8]

Idea:
In relational terms, suppose there is a relation Arc (X,Y) containing pairs (x, y) such that there is arc $x \rightarrow y$.
Iterative computation on the relation Reach(X) that is the set of nodes reachable from node v. After the ith loop, Reach(X) will contain all nodes in $N(v, i)$.
First: Reach(X) contains only node v.
We assume it contains all nodes at $N(v, i)$ after i loops with MapReduce. To determine $N(v, i + 1)$ we need to join of Reach(X) with $Arc(X, Y)$ relation, then project onto Y and union the result with the old value of Reach.
In SQL terms:

```
SELECT    DISTINCT Arc.Y
FROM      Reach, Arc
WHERE     Arc.X = Reach.X
```

Algorithm consists of three main steps with as follows:

– $Reach(X) = \{v\}$
– Assume $Reach(X)$ contains all nodes at $N(v, i)$ after the i[th] iteration.
– To determine $N(v, i + 1)$ in the i+1[th] iteration, we need:
+ Step 1: Join Reach relation and Arc relation using Algorithm Join (natural join to get nodes u from arc $v \rightarrow u$ with v contained in Reach).
+ Step 2: Projection to get u.
+ Step 3: Union with the old value of Reach.
+ Iterate the above steps until no new nodes are added to Reach.

Algorithm 1. Join via MapRedue:

– ***Key: X.***
– ***Value: List of nodes reachable from X.***
– ***Map:***
+ For each $tuple(v)$ of $Reach(X)$, create a key-value pair (v, v).
+ For each $tuple(v, u)$ of $Arc(X, Y)$, generate a key-value pair (v, u).
– ***Reduce:***
+ Each key v is associated with a list of u values where the key value v is in the relation Reach and there is the tuple (v, u) in the relation Arc.

Algorithm 2. Determine Reachability.

Input **Two relations: Reach, Arc.**
Output **List of nodes reachable from X.**
Method Initialization: $Reach(X) = \{v\}$.
Assume $Reach(X)$ contains all nodes at $N(v, i)$ after i rounds.
To generate $N(v, i + 1)$, we need:

– Step 1: Join (natural union to get nodes u from arc $v \rightarrow u$) Reach with Arc relations. In this step there are 2 phases: Map phase and Reduce phase.
  + Map:
    - Map for Reach relation.
    - Map for Arc relation.
  + Reduce: Join Reach relation and Arc relation using Algorithm Join.
– Step 2: Projection (projection to get u).
– Step 3: Combine the result with the old value of Reach:
– Reach = Old Reach union the Result found in this loop.

3.3 Illustration

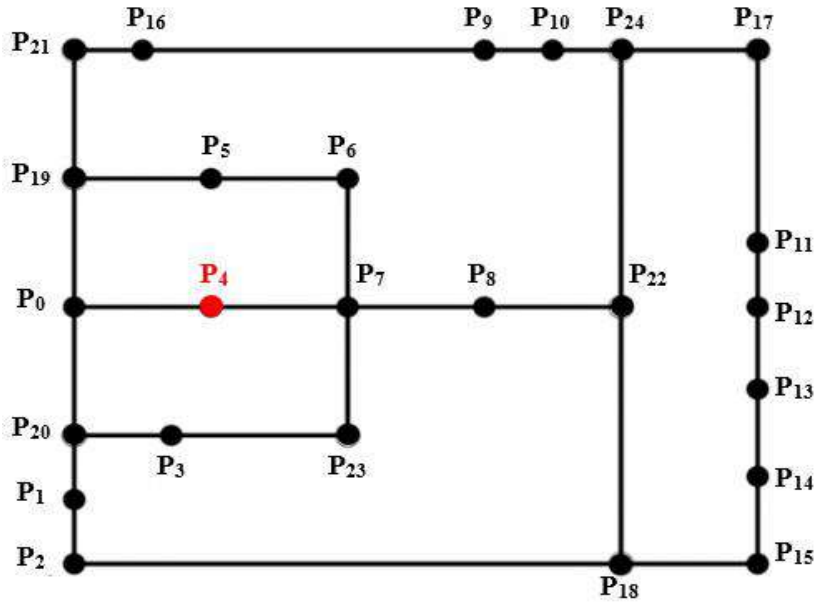


Figure 2. Illustration data

The map and reduce phases are illustrated as follows:

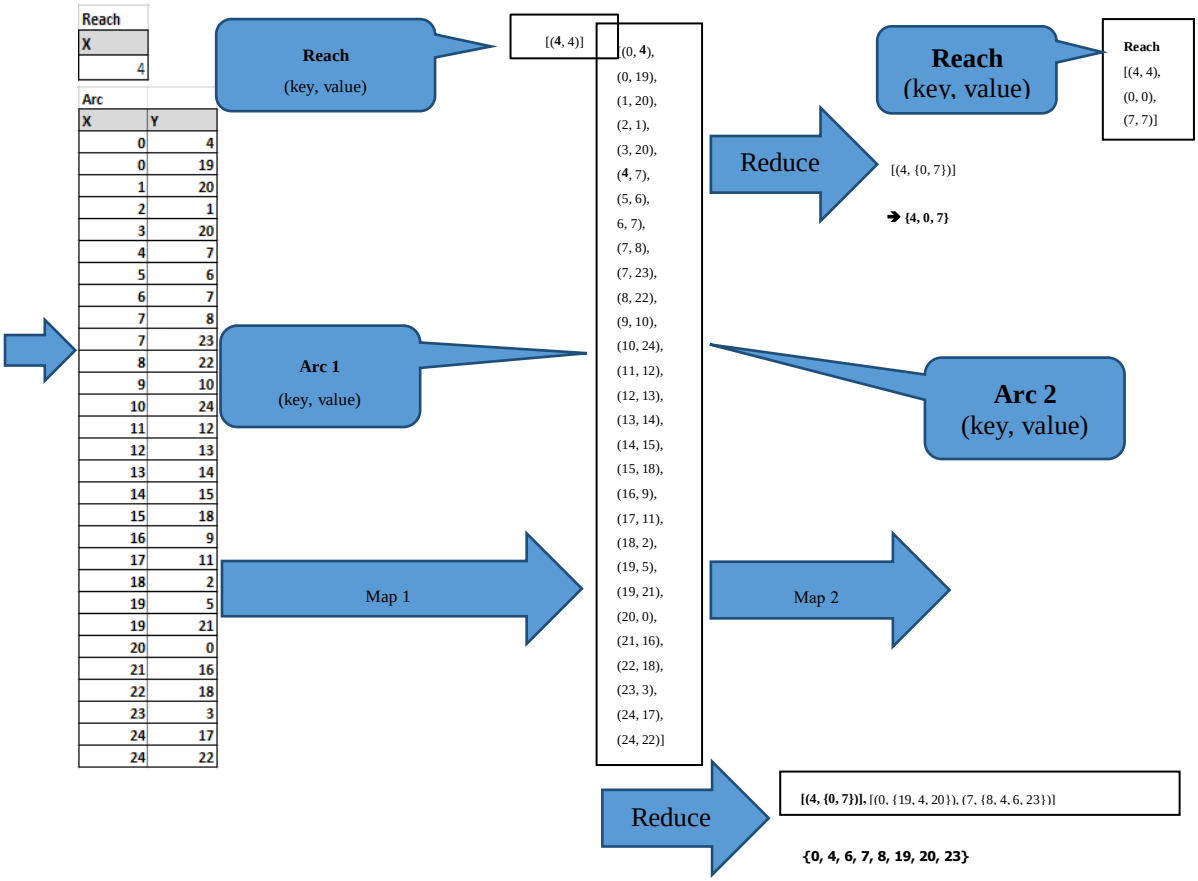


Figure 3. Simulate the Map and Reduce phases.

Comment:

Map: The Arc relation does not change during the execution of the algorithm, so it is only performed once, so it is out of the loop. This reduces the execution time of the algorithm.

Reduce: In each iteration, join only the points not considered in the previous iterations. That is, only include the results for the next iteration of the points that are not considered. The next iteration only takes Map and Reduce on this set, so it will reduce the algorithm processing time. For example, we see that we only need to take the values $\{6, 8, 19, 20, 23\}$ instead of $\{0, 4, 6, 7, 8, 19, 20, 23\}$.

3.4 Suggestions for improvement

Storing two arcs for each edge for undirected graphs wastes storage space as well as wastes the search operation in the Arc relation to join in the algorithm. So, to reduce the storage space and reduce the number of algorithm iterations, we suggest storing in the following way:

+In undirected graphs, we only store one arc for each edge. When processing, if it is an undirected graph, we consider two dimensions. That is, consider both arcs (v, u) and (u, v) in the Arc relation.

+The graph is directed, we store all the arcs and when processing we consider only one dimension.

In summary, along with the comments in the illustration above, we have made the following improvements:

– Storage solution: Store once for each edge in the undirected graph and handle two dimensions in the algorithm as mentioned above.

– Map phase: Perform Map for Arc set only once instead of doing it in a loop → reducing time.

– Reduce phase: Only perform on the latest and unreviewed Reach result set (not in Reach) → reduce the considered points.

```
SELECT    DISTINCT Arc.Y
FROM  Reach, Arc
WHERE     Arc.X = Reach.X
```

```
SELECT    DISTINCT    Arc.Y
FROM  NewReach,       Arc
WHERE     Arc.X = NewReach.X
```

NewReach = NewReach \ Reach

Following is the pseudocode of the improved algorithm.

Algorithm 2. Determine Reachability.
Input: Two relations: Reach, Arc.
Output: List of nodes reachable from X.

Table 1. Algorithm Determine Reachability

Algorithm 2. Determine Reachability	
(1)	Reach(X) = {v}
(2)	ArcMap ← funcMap(Arc) *# Map for Arc relation*
(3)	Change = True *# FALSE: stop*
(4)	Result = set() # contains the result
(5)	while (Change == True):
(6)	r ← []
(7)	ReachMap ← funcMap(df_ReachX, r) *# Map for Reach relation into r*
(8)	ReachCurrent ← apply_async_with_funcReduce (ReachMap, ArcMap) *# Reduce # contains the reachable points at the current loop*
(9)	ReachCurrentKq ← Filter ReachCurrent only gets list of new points not in Result
(10)	Result = Result.union(ReachCurrentKq)
(11)	If |Result| not change then Change = False

4 EXPERIMENTS

The experimental results and evaluations of algorithms will be presented here.

4.1 Environmental experiment

The algorithms are tested on the computer Intel(R) Core (TM) i5-8250U @1.60GHz (4 cores), memory: 12GB DDR4 SDRAM.

CPU Intel(R) Core (TM) i5-8250U @ 1.60GHz (5 cores).
Memory: 12.0 GB.
Operating System: Windows 64-bit.
Programming language: Python.

4.2 Datasets

The algorithm is run experimentally on the illustrated dataset and OSM geospatial data (http://download.geofabrik.de/asia/vietnam.html).
Illustrative data includes 25 points from P_0 to P_{24} as Figure 2.
This report uses two layers of point (POIs) and line data from DaNang city, Vietnam. Points are locations such as schools, fast food stores, supermarkets, etc. Lines are sections of the national road network connecting the nodes. This data source is good enough and comparable to authoritative data to some extent [9].
In addition, the algorithm is also tested on the dataset of social networks Facebook (anonymous) downloaded from: https://snap.stanford.edu/data/index.html [10]. Nodes are users, edges are related users. The following table is a description about these three datasets.
1) Illustrated data.
2) Spatial data of provinces in Vietnam is downloaded from OpenStreetMap (OSM).
3) Social network data from Facebook downloaded from SNAP.

Table 2**.** Data description

Name	Nodes	Edges
Illustrative data	25	29
Da Nang	3284	8340
ego_Facebook	4,039	88,234

4.3 Data preprocessing

The data downloaded is preprocessed before executing the algorithm. The network-constrained spatial data is preprocessed and represented as Arc relations as described above, where tuples are arcs of the graph.

4.4 Results

Table 3. Description of data and experimental parameters

Name	Nodes	Edges	block_size	block_numbers
Illustrative data	25	29	5	6
Da Nang	3284	8340	834	10
ego_Facebook1	4,039	88,234	1000	100

Table 4. The results of comparing comparison of time of algorithms

Name	Non-Parallel (1)	Parallel (2)	Parallel has improved MapReach search (3)	(2) versus (1)	(3) versus (1)
Illustrative data	0.031917334	0.048868895	0.030916214	-53%	3%
Da Nang	606.5766108	523.9522090	7.413079500	14%	99%
ego_Facebook1	342.9537766	134.5248480	40.86810422	61%	88%

Table 5. The results of comparing comparison of memory of algorithms

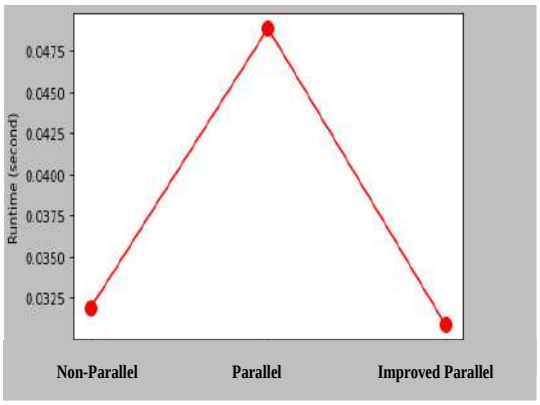
Name	Non-Parallel (1)	Parallel (2)	Parallel has improved (3)	(2) versus (1)	(3) versus (1)
Illustrative data	111,104,000	111,079,424	111,255,552	0%	0%
Da Nang	116,805,632	158,560,256	113,799,168	-36%	3%
ego_Facebook1	138,313,728	144,162,816	131,375,104	-4%	5%

Experimental results show that the parallel algorithm and the improved algorithm reduce the execution time. However, for very small data the converse is more time consuming because it takes time to switch the context as the dataset illustrates with only 25 points the parallel algorithm takes longer with 53%. While on a larger data set like Da Nang downloaded from OSM, it saves 14% with parallel solutions and up to 99% time savings with the improved algorithm.

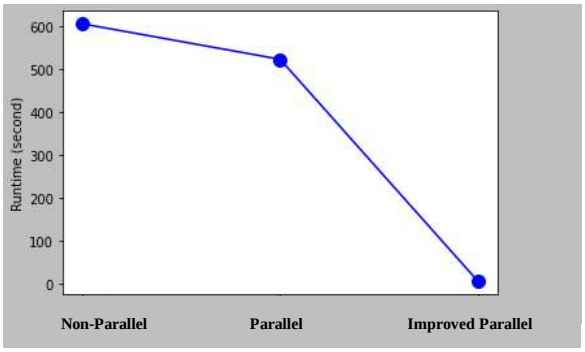
With a larger dataset of Facebook friends downloaded from the SNAP site (https://snap.stanford.edu/), time savings are up to 61% and 88% for the parallel algorithm and improved algorithm. Thus, the improved algorithm greatly reduces the execution time.

Memory comparison results in Table 5 shows that the parallel algorithm consumes more memory The improved parallel algorithm saves memory but is not significant.

The following charts show the comparison of execution time between algorithms: non-parallel, parallel and improved algorithm on three data sets: illustrative data, Da Nang and ego_Facebook1.



c) Illustrated data



b) Da Nang (OSM)

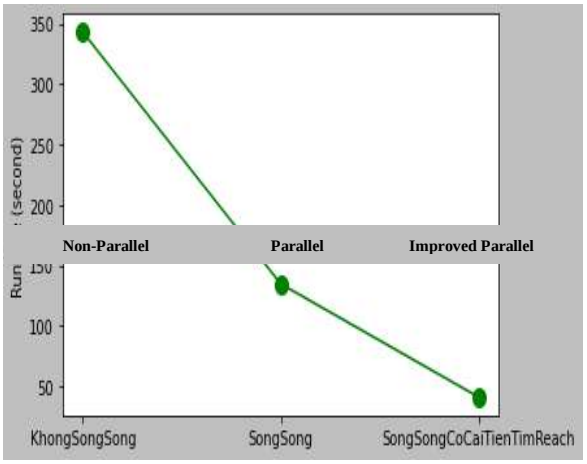


Figure 4. Compare the processing time of algorithms

5 DISCUSSIONS

Storing according to the proposed solution in the above improvement is only effective for undirected graphs, for directed graphs that still have to store two arcs. This proposed storage solution does not improve performance.

However, although it is a directed graph, the number of 2-dimensional edges is more than 1-dimensional edges, for example in a traffic network, the number of two-way roads is much larger than the number of one-way roads.

Therefore, for a directed graph where the two-dimensional edges take up the majority: To reduce the storage space, thus reducing the searching space in the algorithm, we can use this undirected storage solution (save

only one arc per edge) and add a column to indicate 2 or 1 dimension. If the edge is two-dimensional, store the value as 2, otherwise it is 1. Obviously, adding columns reduces the search time more than adding rows because adding rows increases the number of loops.
In addition, to further speed up the processing time, the following improvements can be made:
In the process of determining Reachibility, if arc $4-7$ have been considered, do not consider arc $7-4$ anymore to avoid useless repetition because going back to the node that has already passed certainly does not generate new nodes. The problem of finding Reachibility is to find the set of nodes that can be reached from a node, that is, the next consideration must generate a new node. That means not going back to the node that went through, by add the Flag column with value 0 to the Arc relation, if considered, it will turn up to 1. When considering an edge, it must check that Flag is 0 to consider.
For example: Considering $(4,7)$ then not considering arcs up to 4 like $(7,4),(8,4)$... by turning $(4,7) = 1$. That is, when considering an arc to find a vertex to be able to access it, it is necessary to check whether the end vertex of that arc is considered or not, if not, then it will be considered.
Add flag value $Flag = 0$ for vertices in Reach relation, if vertex is considered, turn this value on to 1. Only continue to consider for vertices whose Flag value is 0.

6 CONCLUSION AND FUTURE WORKS

This article implements the MapReduce solution for the problem of finding the reachable set of a point on a single computer. The concept of the reachability set is related to the notions of a neighborhood.
Future research direction is to experiment MapReduce solution on many computers for the problem of finding the reachable set of a point. From there, the MapReduce solution is given to the problem of finding a neighborhood of a point for a weighted graph in network space. From there, the MapReduce solution can be applied to the problem of geospatial data clustering in network space, improving the algorithm presented in [11]. In addition, Spark model [12] will be considered for research and experiment.

REREFENCES

[1] Y. Qin, Z.L. Yu, C.-D. Wang, Z. Gu, Y. Li, A Novel clustering method based on hybrid K-nearest-neighbor graph, Pattern Recognit. 74 (2018) 1–14. https://doi.org/https://doi.org/10.1016/j.patcog.2017.09.008.

[2] J. Maillo, S. Ramírez, I. Triguero, F. Herrera, kNN-IS: An Iterative Spark-based design of the k-Nearest Neighbors classifier for big data, Knowledge-Based Syst. 117 (2017) 3–15. https://doi.org/https://doi.org/10.1016/j.knosys.2016.06.012.

[3] X. Pan, Y. Luo, Y. Xu, K-nearest neighbor based structural twin support vector machine, Knowledge-Based Syst. 88 (2015) 34–44. https://doi.org/https://doi.org/10.1016/j.knosys.2015.08.009.

[4] N. García-Pedrajas, J.A. del Castillo, G. Cerruela-García, A Proposal for Local k Values for k -Nearest Neighbor Rule, IEEE Trans. Neural Networks Learn. Syst. 28 (2017) 470–475. https://doi.org/10.1109/TNNLS.2015.2506821.

[5] W. Pedrycz, The design of cognitive maps: A study in synergy of granular computing and evolutionary optimization, Expert Syst. Appl. 37 (2010) 7288–7294. https://doi.org/https://doi.org/10.1016/j.eswa.2010.03.006.

[6] S. Pourbahrami, L.M. Khanli, S. Azimpour, A novel and efficient data point neighborhood construction algorithm based on Apollonius circle, Expert Syst. Appl. 115 (2019) 57–67. https://doi.org/https://doi.org/10.1016/j.eswa.2018.07.066.

[7] M.B. Partensky, The Circle of Apollonius and Its Applications in Introductory Physics, Phys. Teach. 46 (2008) 104–108. https://doi.org/10.1119/1.2834533.

[8] J.U. Jure Leskovec, Anand Rajaraman, Mining of Massive Datasets, (n.d.). http://www.mmds.org/#ver21.

[9] M. Haklay, How Good is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets, Environ. Plan. B Plan. Des. 37 (2010) 682–703. https://doi.org/10.1068/b35097.

[10] Stanford Large Network Dataset Collection, (n.d.). https://snap.stanford.edu/data/index.html.

[11] T.T.D. Nguyen, L.T.T. Nguyen, A. Nguyen, U. Yun, B. Vo, A method for efficient clustering of spatial data in network space, J. Intell. Fuzzy Syst. 40 (2021) 11653–11670. https://doi.org/10.3233/JIFS-202806.

[12] Apache Spark, (n.d.). https://spark.apache.org.

ICATSD2F.122

COOPERATIVE NOMA-ENABLED CELLULAR INTERNET-OF-THINGS SYSTEMS: PERFORMANCE ANALYSIS

TIEN-TUNG NGUYEN, ANH VINH NGUYEN, HUNG LE VAN, TAN LOC NGUYEN

Industrial University of Ho Chi Minh City

nguyentientung@iuh.edu.vn, nguyenanhvinh@iuh.edu.vn, levanhung@iuh.edu.vn, nguyentanloc@iuh.edu.vn

Abstract. In this paper, we study a cooperative non-orthogonal multiple access (NOMA)-enabled cellular Internet-of-Things (C-IoT) systems, where an IoT master node assists an information source to communicate with a cell-edge user while forwarding data of far IoT user to a fusion center. In this system, NOMA is deployed at two stages: (1) the information source and the far IoT user transmit their data to IoT master node in the uplink phase and (2) IoT master node operating as a relay to decode and forward data the egde user along with fusion center in the downlink. To evaluate the performance of the proposed system, exact closed-form expressions for the outage probability of the users are derived. Our results show that under interference from the near IoT user, the proposed system may not gain outage performance at users assisted by relaying transmission but improve efficient spectrum usage when compared with orthogonal multiple access (OMA)-based system.

Keywords. Cellular Internet-of-Things (C-IoT), non-orthogonal multiple access (NOMA), outage performance.

1 INTRODUCTION

Recently, with ever-increasing number of Internet-of-Things (IoT) devices and the convergence of emerging wireless communication technologies, a novel integrated platform where communication of IoT devices can be enabled in the existing cellular networks, has begun to emerge [1], [2]. However, a major problem of the rapid development is spectrum shortage due to a dramatically increasing in spectrum usage. To address this challenge, non-orthogonal multiple access (NOMA) technique has been anticipated as an auspicious candidate. In principle, NOMA-based systems achieve high spectral efficiency thanks to simultaneously serve multiple users by the same resource block such as time, frequency, and spreading code [3]. Along the years, due to the demand of improving the spectral efficiency and expanding the transmission coverage in wireless communication systems, cooperative NOMA strategy has been widely proposed in the literature. In [4], the outage probability (OP) and the ergodic capacity (EC) of a relay-aided NOMA network are investigated. The overall OP of a two-way relaying network was analyzed and optimized by jointing power and time allocation in [5]. The authors in [6] studied the performance of a virtual full-duplex (FD) relaying scheme in terms of OP and EC. Considering the joint effects of in-phase and quadrature-phase imbalance (IQI) and imperfect successive interference cancellation (SIC) on OP and ergodic sum rate evaluated in a cooperative NOMA system [7]. In [8], the OP for different relay selection schemes of a NOMA-based cooperative network using multiple relays was derived. Considering an IoT system setup, the outage performance of a cooperative NOMA-based network, where the IoT master node performs as an energy harvesting relay assisting the source to communicate with the destination and simultaneously conveying data to its IoT user, was investigated in [9]. The closed-form and asymptotic expressions for the average block-error-rate, throughput, and the goodput of the coordinated direct and relay transmission (CDRT) over Rayleigh fading channels were derived in [10]. The cooperative NOMA with CDRT was studied in [11], where the proposed system provided higher ergodic sum capacity than the OMA counterpart. In [12], a cooperative underlay cognitive NOMA-based CDRT was proposed, where the primary and secondary receivers in the considered system achieved higher outage performance than those of the cognitive OMA-based CDRT. Considering the uplink NOMA transmission, the power allocation

strategy was proposed [13], where uplink NOMA transmission achieved double sum data rate as compared to OMA at high SNR regions. In [14], the uplink NOMA with two-user scenario was studied, where the ergodic sum capacity was analyzed under imperfect SIC. The authors in [15] investigated the uplink and downlink NOMA transmissions in cellular wireless networks, where the NOMA transmission provided impressive throughput gains over OMA one in both uplink and downlink. In [13], an uplink power back-off control scheme was proposed to improve the system OP and achievable sum data rate in 5G systems. However, the SIC error processing incorporating with uplink and downlink cooperative NOMA transmission still has not properly investigated in these works.

Differently from previous works[1], in this paper, we propose a novel cooperative NOMA network, where an IoT system is combined with a cellular system, which will be termed as scheme NOMA-enabled cellular IoT system (NECTS). The proposed network enhances the spectrum utilization by deploying NOMA for both uplink and down link phases. Based on this system setup, exact closed-form for the users' OP and approximated expression for the ergodic sum capacity of NECTS are derived and compared with conventional orthogonal multiple access (OMA)-based system. Simulation and analytical results demonstrate that the proposed scheme achieves superior ergodic sum capacity (ESC) with slight outage probability loss, compared with the other benchmark strategy.

The main contributions of this paper are briefly summarized as follows.

- First, we propose cooperative NOMA-enabled cellular IoT with down link and uplink considered. In the system, two IoT devices, i.e., two cameras, send their data to a fusion center, however, the far camera needs to assistance of one relay serving for communication between one cellular source and one cellular destination.
- Second, we derive the closed-form expressions for the OP of the users.

The remaining paper is organized as follows. Section II describes the system model. Section III explores the performance analysis. Section IV addresses the numerical results and discussion. Finally, Section V provides a conclusion for this paper.

2 SYSTEM MODEL

We consider a cooperative incorporated down/uplink NOMA-enabled cellular IoT, as illustrated as in Figure 1 where one information source (denoted by $\mathcal{S}$) communicates with one cell-edge user $\mathcal{U}$ in downlink. Two IoT devices (e.g., near camera $\mathcal{D}_1$ and far camera $\mathcal{D}_2$) send their collecting data to one fusion center $\mathcal{F}$ in uplink through the assistance of one master IoT node acting as DF relay (denoted by $\mathcal{R}$).

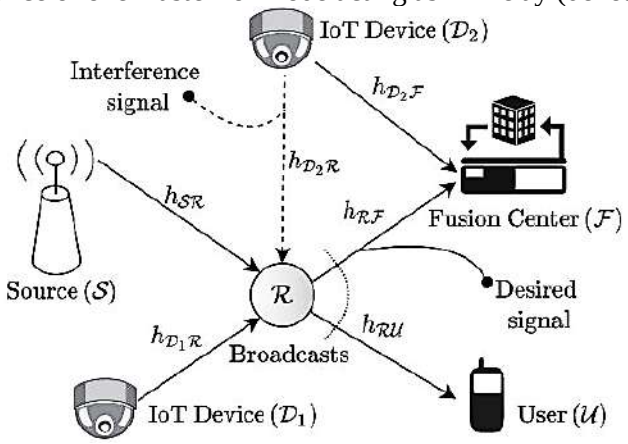


Figure 1. An illustrate of cooperative NOMA-based cellular IoT systems.

The user $\mathcal{D}_2$ directly transmits its data to the fusion center $\mathcal{F}$. The $\mathcal{S}$ and $\mathcal{D}_1$ communicate with their corresponding to destinations via assistance of $\mathcal{R}$ without direct links due to the deep shadowing [6], [10]. The whole data collection at $\mathcal{F}$ and information transmission of $\mathcal{S}$ process is carried out in two consecutive time slots. Let denote $h_{nm} \sim \mathcal{CN}(0, \Omega_{nm})$ as the channel coefficient of the link between the n transmitter and

[1] The studies [4]–[8] focus on assessing performance of cellular-based systems, while works [9], [10] adopting NOMA to simultaneously serve information node and IoT user node in the downlink phase.

m receiver with $n \in \{\mathcal{S}, \mathcal{R}, \mathcal{D}_1, \mathcal{D}_2\}$, $m \in \{\mathcal{R}, \mathcal{F}, \mathcal{U}\}$. In which, $\Omega_{n,m}$ is represented by $\Omega_{n,m} = \frac{\mathcal{L}}{(d_{n,m}/d_0)^{\theta}}$ [16], where $d_{n,m}$ denote the distance between the n transmitter and m receiver, d_0 denotes the reference distance, $\mathcal{L}$ is the average signal power attenuation at d_0 and θ is the path-loss exponent.

2.1 First Communication Stage

In the first timeslot, $\mathcal{D}_2$ directly sends data to $\mathcal{F}$ while $\mathcal{S}$ and $\mathcal{D}_1$ simultaneously transmit their information to $\mathcal{R}$. The received signals at $\mathcal{F}$ while $\mathcal{R}$ are, respectively, expressed as follows:

$$y_{\mathcal{F}} = \sqrt{P_{\mathcal{D}_2}} h_{\mathcal{D}_2\mathcal{F}} x_{\mathcal{D}_2} + n_{\mathcal{F}}, \tag{1}$$

$$y_{\mathcal{R}} = \sqrt{P_{\mathcal{S}}} h_{\mathcal{S}\mathcal{R}} x_{\mathcal{U}} + \sqrt{P_{\mathcal{D}_1}} h_{\mathcal{D}_1\mathcal{R}} x_{\mathcal{D}_1} + \sqrt{P_{\mathcal{D}_2}} h_{\mathcal{D}_2\mathcal{R}} x_{\mathcal{D}_2} + n_{\mathcal{R}}, \tag{2}$$

where $x_{\mathcal{D}_1}$, $x_{\mathcal{D}_2}$ and $x_{\mathcal{U}}$ are the decoded signals of $\mathcal{D}_1$, $x_{\mathcal{D}_2}$ and $\mathcal{U}$ with $\mathbb{E}\left\{\left|x_{\mathcal{D}_1}\right|^2\right\} = \mathbb{E}\left\{\left|x_{\mathcal{D}_2}\right|^2\right\} = \mathbb{E}\{|x_{\mathcal{U}}|^2\} = 1$. And $n_{\mathcal{F}}$ and $n_{\mathcal{R}}$ denote the additive white Gaussian noise (AWGN) at $\mathcal{F}$ and $\mathcal{R}$, respectively, with zero-mean and the corresponding variance $N_{\mathcal{F}}$ and $N_{\mathcal{R}}$. We examine two case studies at $\mathcal{R}$ when the SIC is employed.

The signal of S come R before the signal of D_1 coming. Based on this assumption, the signal-to-interference-plus-noise ratios (SINRs) for decoding x_U and decoding x_{D_1} at R are, respectively, written as:

$$\psi_{\mathcal{R},C_1^1}^{x_{\mathcal{U}}} = \frac{\rho_{\mathcal{S}}|h_{\mathcal{S}\mathcal{R}}|^2}{\rho_{\mathcal{D}_1}\left|h_{\mathcal{D}_1\mathcal{R}}\right|^2 + \rho_{\mathcal{D}_2}\left|h_{\mathcal{D}_2\mathcal{R}}\right|^2 + 1}, \tag{3}$$

$$\psi_{\mathcal{R},C_1^1}^{x_{\mathcal{D}_1}} = \frac{\rho_{\mathcal{D}_1}\left|h_{\mathcal{D}_1\mathcal{R}}\right|^2}{\varepsilon_1\rho_{\mathcal{S}}|h_{\mathcal{S}\mathcal{R}}|^2 + \rho_{\mathcal{D}_2}\left|h_{\mathcal{D}_2\mathcal{R}}\right|^2 + 1}, \tag{4}$$

where $\rho_{\mathcal{S}} \triangleq P_{\mathcal{S}}/N_{\mathcal{R}}, \rho_{\mathcal{D}_1} \triangleq P_{\mathcal{D}_1}/N_{\mathcal{R}}, \rho_{\mathcal{D}_2} \triangleq P_{\mathcal{D}_2}/N_{\mathcal{R}}$ and ε_1 presents the level of residual interference (RSI) caused by imperfect SIC at $\mathcal{R}$.

2.2 Second Communication Stage

To enhance the spectral efficiency, NOMA protocol is adopted at $\mathcal{R}$. Let α_1 and α_2 are the power allocation ratio for the receiver destinations with $a_1 > a_2$ and $\alpha_1 + \alpha_2 = 1$. The power allocation policy assigned for $\mathcal{F}$ and $\mathcal{U}$ is examined two situations based on the quality-of-service (QoS) requirement of receiver destination. We assume that the QoS of $\mathcal{F}$ is higher than $\mathcal{U}$, therefore, more power allocation is assigned for the message of $\mathcal{D}_1$ and less for $\mathcal{U}$. The observation signal at $\mathcal{F}$ and $\mathcal{U}$ are, respectively, given by:

$$y_k = \left(\sqrt{P_{\mathcal{R}}\alpha_1} x_{\mathcal{D}_1} + \sqrt{P_{\mathcal{R}}\alpha_2} x_{\mathcal{U}}\right) h_{\mathcal{R}k} + n_k, k \in \{\mathcal{F}, \mathcal{U}\}, \tag{5}$$

where $n_k \sim \mathcal{CN}(0, N_k)$ presents the AWGN terms at n_k receiver and $P_{\mathcal{R}}$ is the transmit power of $\mathcal{R}$. The $\mathcal{U}$ uses SIC to first decode and remove $x_{\mathcal{D}_1}$ from the received signals before detecting its own message. Meanwhile, from the received signals, $\mathcal{F}$ directly decodes the message of $\mathcal{D}_1$. Thus, the received SINR and SNR of $\mathcal{U}$ for decoding the message of $\mathcal{D}_1$ and $\mathcal{U}$ are, respectively, given by:

$$\psi_{\mathcal{U},C_2^1}^{x_{\mathcal{D}_1}} = \frac{\rho_{\mathcal{R}}\alpha_1|h_{\mathcal{R}\mathcal{U}}|^2}{\rho_{\mathcal{R}}\alpha_2|h_{\mathcal{R}\mathcal{U}}|^2 + 1}, \tag{5}$$

$$\psi_{\mathcal{U},C_2^1}^{x_{\mathcal{U}}} = \frac{\rho_{\mathcal{R}}\alpha_2|h_{\mathcal{R}\mathcal{U}}|^2}{\varepsilon_2\rho_{\mathcal{R}}\alpha_1|h_{\mathcal{R}\mathcal{U}}|^2 + 1}, \tag{7}$$

where $\rho_{\mathcal{R}} \triangleq \frac{P_{\mathcal{R}}}{N_{\mathcal{U}}}$ and ε_2 is RSI caused by SIC process at $\mathcal{U}$. The received SINR at $\mathcal{F}$ for decoding $\mathcal{D}_1$'s message can be expressed as:

$$\psi_{\mathcal{F},C_2^1}^{x_{\mathcal{D}_1}} = \frac{\rho_{\mathcal{R}}\alpha_1|h_{\mathcal{R}\mathcal{F}}|^2}{\rho_{\mathcal{R}}\alpha_2|h_{\mathcal{R}\mathcal{F}}|^2 + 1}. \tag{8}$$

3 PERFORMANCE ANALYSIS

Based on the analysis of first and second communication stage, we have two cases corresponding each communication stage as depicted in Table 1. Therefore, four schemes have been evaluated in this section

accordingly. To characterize the performance of one wireless communication networks, the outage probability (OP) and ergodic sum capacity (ESC) are two importance metrics, which should be investigated; thus, we evaluate the of four schemes in this section, accordingly. The OP of the receiver is defined as the outage event which their achievable rate is lower than a given target data rate. Denote by $R_{\mathcal{D}_1}$, $R_{\mathcal{D}_2}$ and $R_{\mathcal{U}}$ the constant rate of $\mathcal{D}_1$, $\mathcal{D}_2$ and $\mathcal{U}$, respectively. Let $\gamma_{\mathcal{D}_1} = 2^{2R_{\mathcal{D}_1}} - 1$, $\gamma_{\mathcal{D}_2} = 2^{2R_{\mathcal{D}_2}} - 1$ and $\gamma_{\mathcal{U}} = 2^{2R_{\mathcal{U}}} - 1$ be the SINR threshold for decoding signal of $\mathcal{D}_1$, $\mathcal{D}_2$ and $\mathcal{U}$, respectively.

Under the DF protocol, the achievable rate of the considered system setup can be calculated based solely on the weakest link. Thus, the end-to-end (e2e) achievable rate of $\mathcal{D}_1$ is acquired from (4), (6) and (8) as:

$$C_{\mathcal{D}_1}^{C_1^1} = \frac{1}{2} log\left(1 + min\{\psi_{\mathcal{R},C_1^1}^{x_{\mathcal{D}_1}}, \psi_{\mathcal{U},C_2^1}^{x_{\mathcal{D}_1}}, \psi_{\mathcal{F},C_2^1}^{x_{\mathcal{D}_1}}\}\right). \quad (9)$$

The achievable rate of $\mathcal{U}$ is given by:

$$C_{\mathcal{U}}^{C_1^1} = \frac{1}{2} log\left(1 + min\{\psi_{\mathcal{R},C_1^1}^{x_{\mathcal{U}}}, \psi_{\mathcal{U},C_2^1}^{x_{\mathcal{U}}}\}\right). \quad (10)$$

From the definitions, the exact closed-form analytical expressions for the OP of $\mathcal{D}_1$ and $\mathcal{U}$ can be attained in the following theorems.

Theorem 1. *Exact analytical closed-form expression for OP of $\mathcal{D}_1$ can be expressed as*

$$\mathcal{P}_{\text{out}}^{\mathcal{D}_1} = 1 - \frac{exp(-\varsigma_1\gamma_{\mathcal{D}_1})}{(\varepsilon_1\gamma_{\mathcal{D}_1}/\mathcal{A}_1+1)(\mathcal{A}_2\gamma_{\mathcal{D}_1}/\mathcal{A}_1+1)} exp\left(-\frac{\gamma_{\mathcal{D}_1}}{(\alpha_1-\alpha_2\gamma_{\mathcal{D}_1})}\left(\frac{1}{\rho_{\mathcal{R}}\Omega_{\mathcal{R}U}} + \frac{1}{\rho_{\mathcal{R}}\Omega_{\mathcal{R}F}}\right)\right). \quad (11)$$

Theorem 2. *Exact analytical closed-form expression for OP of $\mathcal{U}$ can be expressed as*

$$\mathcal{P}_{\text{out}}^{\mathcal{U}} = 1 - \frac{exp(-\varsigma_2\gamma_{\mathcal{U}})}{(\mathcal{A}_1\gamma_{\mathcal{U}}+1)(\mathcal{A}_2\gamma_{\mathcal{U}}+1)} exp\left(-\frac{\gamma_{\mathcal{U}}}{(\alpha_2-\varepsilon_2\alpha_1\gamma_{\mathcal{U}})\rho_{\mathcal{R}}\Omega_{\mathcal{R}U}}\right). \quad (12)$$

Proof. See Appendix A. □

4 NUMERICAL RESULTS AND DISCUSSION

In this section, we present analytical results for evaluation of the performance of the proposed system. The key parameters used in this paper are provided in Table 1. We can see that the results of analysis and simulation are coincident in all Figures.

Table 1. Simulation parameters.

Parameter	Value
Bandwidth	1MHz
Antenna noise power density, $N_{\mathcal{R}} = N_{\mathcal{F}}$	-100 dBm/Hz
Antenna noise power density, $N_{\mathcal{U}}$	-90 dBm/Hz
Fixed target data rate, $\mathcal{U}$	0.5 Bit/s/Hz
Fixed target data rate of $\mathcal{D}_1$ and $\mathcal{D}_2$	0.25 Bit/s/Hz
Fixed transmit power of $\mathcal{D}_1$	30 dB
Fixed transmit power of $\mathcal{D}_2$	20 dB
Distance of $\mathcal{S} \to \mathcal{R}$	30 meters
Distance of $\mathcal{D}_1 \to \mathcal{R}$	50 meters
Distance of $\mathcal{D}_2 \to \mathcal{F}$	60 meters
Distance of $\mathcal{R} \to \mathcal{F}$	40 meters
Distance of $\mathcal{R} \to \mathcal{U}$	30 meters
Path-loss exponent, θ	3
Path-loss at reference distance, $\mathcal{L}$ (at d_0 = 1 meter)	-30dB
Power allocation factors, α_1	0.8
Power allocation factors, α_2	0.2
Sample numbers	10^6

The OPs at the receivers changing to the transmit power of the source for perfect and imperfect SIC are illustrated as Figure 2. We can see that under imperfect SIC, the performances of the the system

significantly decreases. Due to without depending on P_S, the OP of $\mathcal{D}_2$ is constant. As can be seen that the OP of $\mathcal{U}$ gradually decreases as P_S increases. Moreover, the OP of $\mathcal{U}$ obtains much better than that of OMA-based system. For $\mathcal{D}_1$, the OP decreases to a saturation level as P_S increases a sufficient value while suffering severely outage at high transmit power of source node.

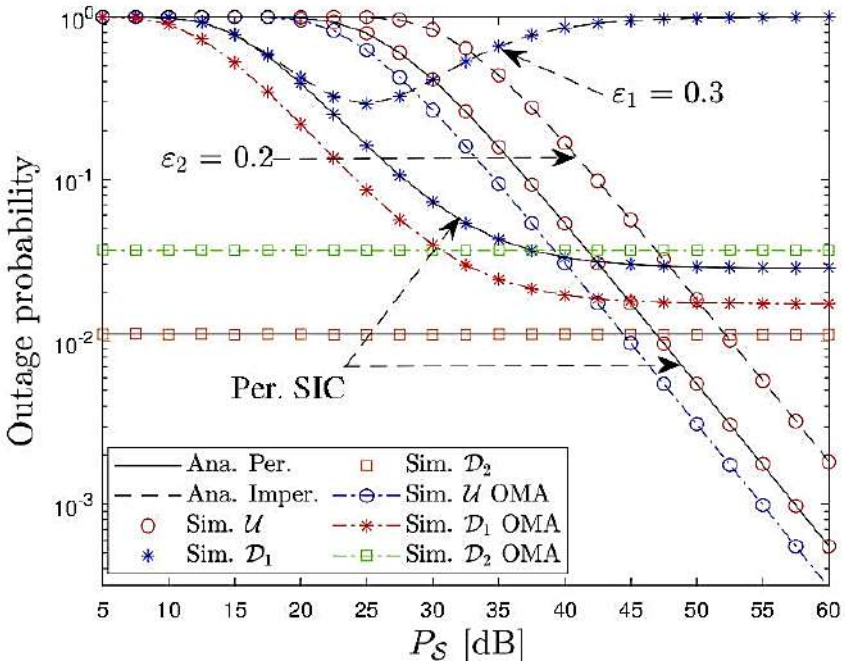


Figure 2. Outage probability of cell-user and IoT devices versus transmit power of source node under pSIC process.

5 CONCLUSIONS

In this paper, a cooperative NOMA-enabled C-IoT system, where an IoT master node assists an information source to communicate with a cell-edge user while forwarding data of far IoT user to a fusion center, is studied under perfect and imperfect SIC. In the uplink phase, the information source and the far IoT user transmit their data to IoT master node, in the downlink phase, IoT master node plays a role of relay to decode and forward data the edge user along with fusion center. Our results show that under interference from the near IoT user, the proposed system may not gain outage performance at users assisted by relaying transmission but improve efficient spectrum usage when compared with orthogonal multiple access (OMA)-based system.

APPENDIX

Appendix A

Proof of Theorem 1

From (9), the OP of $\mathcal{D}_1$ can be expressed as

$$\begin{aligned} \mathrm{P}_{\mathrm{out}}^{\mathrm{D}_1} &= \Pr(\min\{\psi_{\mathrm{R},C_1^1}^{x_{\mathrm{D}_1}}, \psi_{\mathrm{U},C_2^1}^{x_{\mathrm{D}_1}}, \psi_{\mathrm{F},C_2^1}^{x_{\mathrm{D}_1}}\} < \gamma_{\mathrm{D}_1}) \\ &= 1 - \underbrace{\Pr(\psi_{\mathrm{R},C_1^1}^{x_{\mathrm{D}_1}} \geq \gamma_{\mathrm{D}_1})}_{I_1} \underbrace{\Pr(\psi_{\mathrm{U},C_2^1}^{x_{\mathrm{D}_1}} \geq \gamma_{\mathrm{D}_1})}_{I_2} \underbrace{\Pr(\psi_{\mathrm{F},C_2^1}^{x_{\mathrm{D}_1}} \geq \gamma_{\mathrm{D}_1})}_{I_3}, \end{aligned} \tag{13}$$

where the expression of I_1, I_2 and I_3 can be calculated, respectively, as follows:

$$I_1 = \int_{x=0}^{\infty} \int_{y=0}^{\infty} [1 - F_{|h_{\mathcal{RD}}|^2}(\hat{\psi}_1)] f_{|h_{\mathcal{SR}}|^2}(x) f_{|h_{\mathcal{RDD}}|^2}(y) dx dy, \tag{14}$$

where $\hat{\psi}_1 = \frac{\rho_S \varepsilon_1 \gamma_{\mathcal{D}_1} x}{\rho_{\mathcal{D}_1}} + \frac{\rho_{\mathcal{D}_2} \gamma_{\mathcal{D}_1} y}{\rho_{\mathcal{D}_1}} + \frac{\gamma_{\mathcal{D}_1}}{\rho_{\mathcal{D}_1}}$. By substituting $F_{|h_{\mathcal{RD}}|^2}(\cdot), f_{|h_{\mathcal{SR}}|^2}(\cdot)$ and $f_{|h_{\mathcal{RDD}}|^2}(\cdot)$ into (20), where the CDF and PDF are given in (19), we can attain the formulate of I_1 as

$$I_1 = \frac{exp(\varsigma_1 \gamma_{\mathcal{D}_1})}{(\varepsilon_1 \gamma_{\mathcal{D}_1} / \mathcal{A}_1 + 1)(\mathcal{A}_2 \gamma_{\mathcal{D}_1} / \mathcal{A}_1 + 1)}, \tag{15}$$

where $\mathcal{A}_1 = \frac{\rho_{\mathcal{D}_1}\Omega_{\mathcal{R}D}}{\rho_S\Omega_{S\mathcal{R}}}$, $\mathcal{A}_2 = \frac{\rho_{\mathcal{D}_2}\Omega_{\mathcal{R}DD}}{\rho_S\Omega_{S\mathcal{R}}}$ and $\varsigma_1 = 1/(\rho_{\mathcal{D}_1}\Omega_{\mathcal{R}D})$. Meanwhile, the formulate of I_2 and I_3 can be easily obtained as

$$I_2 = exp(-\frac{\gamma_{\mathcal{D}_1}}{(\alpha_1-\alpha_2\gamma_{\mathcal{D}_1})\rho_{\mathcal{R}}\Omega_{\mathcal{R}U}}), \qquad I_3 = exp(-\frac{\gamma_{\mathcal{D}_1}}{(\alpha_1-\alpha_2\gamma_{\mathcal{D}_1})\rho_{\mathcal{R}}\Omega_{\mathcal{R}F}}). \tag{16}$$

Plugging I_1, I_2 and I_3 into (19), we attained the desired result as in (11). The proof of **Theorem 1** is completed.

Proof of Theorem 2

From (10), the OP of $\mathcal{U}$ can be expressed as

$$\begin{aligned} \mathrm{P}_{\mathrm{out}}^{\mathrm{U}} &= \Pr\left(\min\left\{\psi_{\mathrm{R},C_1^1}^{x_{\mathrm{U}}}, \psi_{\mathrm{U},C_2^1}^{x_{\mathrm{U}}}\right\} < \gamma_{\mathrm{U}}\right) \\ &= 1 - \underbrace{\Pr\left(\psi_{\mathrm{R},C_1^1}^{x_{\mathrm{U}}} \geq \gamma_{\mathrm{U}}\right)}_{I_4}\underbrace{\Pr\left(\psi_{\mathrm{U},C_2^1}^{x_{\mathrm{U}}} \geq \gamma_{\mathrm{U}}\right)}_{I_5}. \end{aligned} \tag{17}$$

In which, the formulate of I_4 can be computed as

$$I_4 = \int_{x=0}^{\infty}\int_{y=0}^{\infty}[1 - F_{|h_{S\mathcal{R}}|^2}(\hat{\psi}_2)]f_{|h_{\mathcal{R}D}|^2}(x)f_{|h_{\mathcal{R}DD}|^2}(y)dxdy, \tag{18}$$

where $\hat{\psi}_2 = \frac{\rho_{\mathcal{D}_1}\gamma_{\mathcal{U}}x}{\rho_S} + \frac{\rho_{\mathcal{D}_2}\gamma_{\mathcal{U}}y}{\rho_S} + \frac{\gamma_{\mathcal{U}}}{\rho_S}$. Likewise, in (20), we attained the closed-form expression for I_4 as follows:

$$I_4 = \frac{exp(\varsigma_2\gamma_{\mathcal{U}})}{(\mathcal{A}_1\gamma_{\mathcal{U}}+1)(\mathcal{A}_2\gamma_{\mathcal{U}}+1)}, \varsigma_2 = \frac{1}{\rho_S\Omega_{S\mathcal{R}}}. \tag{19}$$

The formulate of I_5 can be easily obtained as

$$I_5 = exp(-\frac{\gamma_{\mathcal{U}}}{(\alpha_2-\varepsilon_2\alpha_1\gamma_{\mathcal{U}})\rho_{\mathcal{R}}\Omega_{\mathcal{R}U}}). \tag{20}$$

Substituting I_4 and I_5 (23), we obtained the desired result as in (12). The proof of **Theorem 2** is completed.

REFERENCES

[1] J. Ding, M. Nemati, C. Ranaweera, and J. Choi, "IoT connectivity technologies and applications: A survey," *IEEE Access*, vol. 8, pp. 67 646–67 673, 2020.

[2] S. K. Sharma and X. Wang, "Toward massive machine type communications in ultra-dense cellular IoT networks: Current issues and machine learning-assisted solutions," *IEEE Comm. Surv. Tutor.*, vol. 22, no. 1, pp. 426–471, 2020.

[3] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (NOMA)," in *Proc. IEEE Pers. Ind. Mobile Radio Commun.* (PIMRC), Sep. 2013, pp. 611–615.

[4] A. Tregancini, E. E. B. Olivo, D. P. M. Osorio, C. H. M. de Lima, and H. Alves, "Performance analysis of full-duplex relay-aided NOMA systems using partial relay selection," *IEEE Tran. Veh. Technol.*, vol. 69, no. 1, pp. 622–635, 2020.

[5] J. Bae and Y. Han, "Joint power and time allocation for two-way cooperative NOMA," *IEEE Tran. Veh. Technol.*, vol. 68, no. 12, pp. 12 443–12 447, 2019.

[6] Q. Y. Liau, C. Y. Leow, and Z. Ding, "Amplify-and-forward virtual full-duplex relaying-based cooperative NOMA," *IEEE Wireless Comm. Lett.*, vol. 7, no. 3, pp. 464–467, 2018.

[7] X. Li, M. Liu, C. Deng, P. T. Mathiopoulos, Z. Ding, and Y. Liu, "Full-duplex cooperative NOMA relaying systems with i/q imbalance and imperfect sic," *IEEE Wireless Comm. Lett.*, vol. 9, no. 1, pp. 17–20, 2020.

[8] Y. Li, Y. Li, X. Chu, Y. Ye, and H. Zhang, "Performance analysis of relay selection in cooperative NOMA networks," *IEEE Commun. Lett.*, vol. 23, no. 4, pp. 760–763, 2019.

[9] A. Rauniyar, P. E. Engelstad, and O. N. Østerbø, "RF energy harvesting, and information transmission based on NOMA for wireless powered IoT relay systems," *Sensors*, vol. 18, no. 10, p. 3254, 2018.

[10] T. -H. Vu, T. -V. Nguyen, T. -T. Nguyen, V. N. Q. Bao and S. Kim, "Short-Packet Communications in NOMA-CDRT IoT Networks with Cochannel Interference and Imperfect SIC," in *IEEE Transactions on Vehicular Technology*, vol. 71, no. 5, pp. 5552-5557, May 2022, doi: 10.1109/TVT.2022.3148988.

[11] T.-T. Nguyen, T.-V. Nguyen, T.-H. Vu, D. B. da Costa, and C. D. Ho, "IoT-based coordinated direct and relay transmission with non-orthogonal multiple access," *IEEE Wireless Commun. Lett.*, Nov. 2020.

[12] T. T. Nguyen, T. H. Vu, T. V. Nguyen, D. B. da Costa, and C. D. Ho, "Underlay cognitive NOMA-based coordinated direct and relay transmission," *IEEE Wireless Commun. Lett*., pp. 1–5, Dec. 2020.

[13] N. Zhang, J. Wang, G. Kang, and Y. Liu, "Uplink nonorthogonal multiple access in 5G systems," *IEEE Commun. Lett.*, vol. 20, no. 3, pp. 458–461, 2016.

[14] M. F. Kader and S. Y. Shin, "Coordinated direct and relay transmission using uplink NOMA," *IEEE Wireless Comm. Lett.*, vol. 7, no. 3, pp. 400–403, 2017.

[15] M. S. Ali, H. Tabassum, and E. Hossain, "Dynamic user clustering and power allocation for uplink and downlink nonorthogonal multiple access (NOMA) systems," *IEEE Access*, vol. 4, pp. 6325–6343, 2016.

[16] T. Nguyen, V. Nguyen, J. Lee, and Y. Kim, "Sum Rate Maximization for Multi-User Wireless Powered IoT Network with Non-Linear Energy Harvester: Time and Power Allocation," *IEEE Access*, vol. 7, pp. 149 698–149 710, 2019.

[17] A. Jeffrey and D. Zwillinger, Table of integrals, series, and products. Elsevier, 2007.

ICATSD2F.123

PLANT CLASSIFICATION IN SOUTHEAST ASIA USING HIGH-RESOLUTION NETWORK

DANG NGAN HA[1], TRAN HONG NGOC[1], HIEU TRUNG HUYNH[2]

[1]*Vietnamese-German University*

[2]*Industrial University of Ho Chi Minh City*

14328@student.vgu.edu.vn, hthieu@iuh.edu.vn, ngoc.th@vgu.edu.vn

Abstract. Intelligent technologies and applications are becoming more common in our modern environment, yet we seldom notice them. Machine learning has greatly reduced human difficulties and efforts in everyday tasks. It has improved the computer vision problems of recognition and classifying by advanced algorithms and improvements in computation. The information on the plant species is important for not only agriculture but also medicine. This paper presents an approach for plant classification in Southeast Asia by using a high-resolution network. The experiments are evaluated on a public dataset including 26 different species in Southeast Asia. The high-resolution network can obtain high accuracy.

Keywords. Plant classification, high-resolution network, deep learning.

1 INTRODUCTION

The conservation of natural resources depends heavily on plants and understanding species is crucial for preserving biodiversity. Identification of a plant's species offers important details about a plant's classification and features. Manual interpretation may be imprecise since it depends on the viewer's visual perception. Plant identification using traditional keys is difficult, time-consuming, and frustrating for non-experts due to the usage of specialized botanical words. For newcomers who are interested in learning about species, this poses a difficult obstacle to overcome. Digital leaf pictures may easily be sampled and taken in this way, and the textural elements make it easy to identify a particular pattern. Venation and leaf shape are the most crucial characteristics for identifying different plant species. The identification of plants based on the description of leaf shape and venation, which is the essential idea in the identification process, is done using techniques like image processing, pattern recognition, and so forth as information technology advances quickly. It is challenging to keep track of changing leaf properties throughout time.

Automating the process of identifying species is becoming more and more popular today. The notion of automatic species identification has become a reality because of the accessibility and rapid spread of essential technology including digital cameras and mobile devices, remote database access, new computer vision techniques, and pattern recognition algorithms. The use of image-based techniques is seen to be a viable strategy for identifying species. A user can use the built-in camera of a mobile device to take a picture of a plant in the wild and then analyze it using a recognition program loaded on the device to determine the species, or at the very least to get a list of potential species if a specific match is unattainable. Additionally, non-professionals can participate in this process by employing a computer-aided plant identification system. As a result, it is not surprising that several research investigations are focused on automating the process of identifying plant species. During the last decade, there were a large number of noticeable research on plants performed with machine learning algorithm, as well as many datasets made publicly available for research purposes.

One of those studies includes computer vision classification of leaves from Swedish trees [1] which has been captured as part of a joined leaf classification project between the Linkoping university and the Swedish museum of natural history. They contrast several descriptors that will explain the various aspects of the leaves. To create a system that could categorize the various tree classes, they also consider other classification models and integrate them with the descriptors. A leaf recognition algorithm for plant classification using a probabilistic neural network uses a dataset with leaves sampled on the campus of

Nanjing University and the Sun Yat-Sen arboretum, Nanking, China [2]. Most of them are common plants of the Yangtze Delta, China. Z. A. Nilsback ME et al. [3] [4] were conducted on the Oxford Flower 17 and 102 datasets to investigate to what extent combinations of features can improve classification performance on a large dataset of similar classes. Sonali Agrawala et al. [5] recognized a plant by examining the leaf image to assist novice users and botanists in determining which plants are edible or dangerous to people. An approach based on the deep convolutional neural network was proposed by Mads Dyrmann et al. [6]. The experiments were evaluated on 22 species.

Although the research on automatic plant taxonomy has yielded fruitful results, it is still challenging. We must notice that the models are not fully well developed, the datasets used in these studies are still heavily regional, and species oriented. To contribute to the plant classification research field, we have implemented deep high-resolution representation learning for plant classification. The experiments are evaluated on the dataset of Bali26 with 26 different plant species in Southeast Asia [7].

2 THE MODEL FOR PLANT CLASSIFICATION

Deep convolutional neural networks have attained state-of-the-art results in many applications in computer vision, several network models have been developed recently for image classification including VGGNet [8], GoogleNet [9], AlexNet [10], and ResNet [11]. These models process images with a low-resolution representation. Some applications require high-resolution representations. The models such as DeconvNet [12], SegNet [13], or U-Net [14] deal with medium-resolution representations. Recently, J Wang et al. [7] have proposed a method that can maintain the high-resolution representation, this model may offer good results in some applications including semantic segmentation, object recognition, and classification.

The high-to-low convolution streams are parallelly connected by the network. It maintains high-resolution representations throughout the entire process and repeatedly fuses the representations from multi-resolution streams to produce reliable high-resolution representations with high position sensitivity.

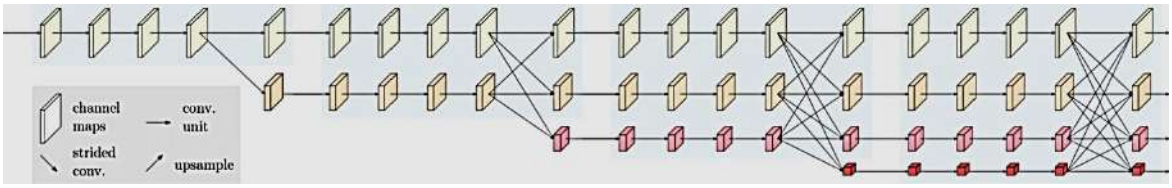


Figure 1. An example of a high-resolution network [7].

An image is fed into a stem, which comprises two stride-2 3x3 convolutions decreasing the resolution to $\frac{1}{4}$, and then the main body that generates the representation with the same resolution. The main body, demonstrated in Figure 1, is divided into multiple stages and consists of several components including parallel multi-resolution convolutions, repeated multi-resolution fusions, and a representation head. It begins with a high-resolution convolution stream as the first stage, progressively adding high-to-low resolution streams one by one, producing additional stages, then linking the multiresolution streams in parallel. Consequently, the resolutions for the parallel streams of a later stage comprise the resolutions from the previous stage, plus an additional lower one.

The fusion module's objective is to communicate data across representations with different levels of resolution. There are multiple repetitions (e.g., every 4 residual units). Let's look at an example of fusing 3-resolution representations, which is illustrated in Figure 2. The input consists of three representations. Each output representation is the sum of the transformed representations of the three inputs. The fusion across stages (from stage three to stage four) has an extra output. The choice of the transform function is dependent on the input resolution index (x) and the output resolution index (r).

If $x < r$, the function down-samples the input representations through stride-2 3x3 convolutions. For example, one stride-2 3x3 convolution for 2x down-sampling and two stride-2 3x3 convolutions for 4x downsampling.

If $x > r$, the function up-samples the input representations through bilinear up-sampling followed by a 1x1 convolution for aligning the number of channels.

If $x = r$, there is the resolution of the input representation is the same as the output representation.

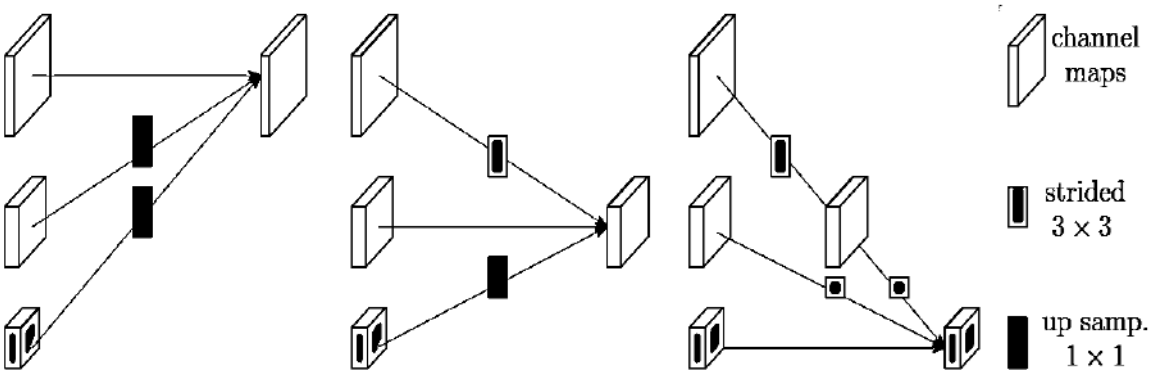


Figure 2. Illustrating how the fusion module aggregates the information for high, medium, and low resolutions [7].

In our implementation, the first stage consists of four residual units, each produced by a bottleneck with a width of 64, and is followed by one 3x3 convolution that changes the width of feature maps to *C*. The second, third, and fourth phases each have one, three, or four modularized blocks. Each branch of the modularized block's multiresolution parallel convolution has four residual units. For each resolution, each unit has two 3x3 convolutions, followed by batch normalization and the nonlinear activation ReLU. Convolution widths (numbers of channels) for the four resolutions are *C*, 2*C*, 4*C*, and 8*C*, respectively.

3 EXPERIMENTAL RESULTS

An image dataset called Bali26 [15, 16] is devoted to the study of human-plant interactions, or ethnobotany. In comparison to comparable samples in ImageNet, the images display a deliriously lush flora with higher information richness (entropy). Similar to the IRIS dataset [17], which contains 3 classes of 50 instances each, where each class refers to a type of iris plant, the data samples from Bali26 were gathered by a team of data collectors in the wild, in a constrained location, and over a brief period.

Figure 3. Typical images from the Bali26 dataset.

The Bali26 dataset contains images of 26 categories of ethnobotanically significant trees, plants, fruits, and seeds. Each category is represented by between 1000 and 2500 images. All images in this collection were collected on the Island of Bali in the vicinity of Ubud, specifically in the villages of Penglipuran, Kerta, Jatiluwih, Buahan, Sekaan, and Bayung Gede between the months of late February and May 2020 [18]. The information on categories is given in Table 1. The images (1080 x 1920 in .jpg format) were created from high-definition mobile phone video shot by data collectors in the field. Some typical images from the dataset are given in Figure 3.

Table 1. Information of plant categories in the Bali26 dataset

No.	Common plant name	Scientific plant name	Ethnobotany
1	Elephant foot yam (Suweg)	Amorphophallus paeoniifolius	Boiled tuberous roots eaten as a staple food
2	Bamboo	Gigantochloa apus	Cooked young shoots added to vegetable soups
3	Banana	Musa x paradisiaca	Cooked core stems added to vegetable soups
4	Cacao	Theobroma cacao	Decoction of seeds powder for inducing relaxation
5	Coffee arabica	Coffea canephora	Decoction of seeds powder for inducing relaxation
6	Dragon fruit	Hylocereus costaricensis	Fruits used in religious offerings
7	Durian	Durio zibethinus	Fruits used in religious offerings
8	Frangipani	Plumeria alba	Juice of flowers is applied to the skin for smallpox; Flowers used in religious offerings
9	Guava	Psidium guajava	Ripe fruit eaten fresh and used in religious offerings
10	Indonesian cinnamon	Cinnamomum burmanii	Barks used as spices; Decoction of leaves and barks used for heartburn, fever, cough, sore throat, hypertension, and to stimulate the appetite
11	Jackfruit	Artocarpus heterophyllus	Cooked young fruit and seeds added to vegetable soups; Decoction of leaves for diarrhea
12	Lychee	Litchi chinensis	Ripe fruit eaten fresh and used in religious offerings
13	Mango	Mangifera indica	Decoction of leaves for diabetes and hypertension
14	Mangosteen	Garcinia mangostana	Ripe fruit eaten fresh and used in religious offerings
15	Nilam	Pogostemon cablin	Decoction of leaves for headache, dysentery, rheumatic, and as a diuretic
16	Papaya	Carica papaya	Cooked fruit and young leaves added to vegetable soups
17	Passiflora	Passiflora edulis	Ripe fruit eaten fresh
18	Sawo	Manilkara zapota	Ripe fruit eaten fresh and used in religious offerings
19	Snake fruit	Salacca zalacca	Ripe fruit eaten fresh and used in religious offerings
20	Star fruit	Averrhoa carambola	Cooked fruit and leaves added to vegetable soups; Juice of leaves for heartburn
21	Sugar palm	Arenga pinnata	Boiled inner stems eaten as a staple food; Decoction of roots for urolithiasis; Edible boiled fruit; Leaves used in religious offerings
22	Taro	Colocasia esculenta	Boiled tuberous roots eaten as a staple food; Cooked leaves added to vegetable soups
23	Vanilla	Vanilla planifolia	Fruit powder used as vanilla flavoring
24	Water guava	Syzygium aqueum	Ripe fruit eaten fresh
25	White pepper	Piper nigrum	Seeds powder used as a spice
26	Zodiac	Evodia sauveolens	Leaves used for mosquito repellent

All experiments were run on the machine with a GPU that has the CPU of Quad-core with 2.9GHz and

RAM of 16GB. The algorithms were implemented in Python 3.7 environment with Pytorch on Ubuntu 18.04 operating system. The Adam optimizer Adam was used.
Based on the pre-trained weight of the ImageNet dataset, the proposed model is trained on the Bali26 dataset. The results are highly impressive since their accuracy rate is relatively high according to the announcement of the competition. The accuracy is 100% for the testing dataset, which is one of the top teams in the competition and belongs to first place. The remaining accuracies from other teams were ranging from 83.2% to 99.9%. The recognition results from the testing set for 26 categories are given in Table 2.

Table 2. The recognition results from the testing set

No.	Category	# Images	No.	Category	# Images
1	Water guava	205	14	Lychee	124
2	Jack fruit	297	15	Banana	191
3	Dragon fruit	234	16	Sawo	136
4	Papaya	183	17	Snake fruit	130
5	Taro	177	18	Guava	215
6	Vanilla	217	19	Coffee arabica	149
7	Star fruit	189	20	Frangipani	174
8	Zodiac	148	21	Cacao	187
9	White pepper	157	22	Sugar palm	140
10	Cinnamon	136	23	Suweg	132
11	Passiflora	151	24	Bamboo	214
12	Nilam	138	25	Durian	271
13	Mango	149	26	Mangosteen	180

The high-resolution network model costs more memory as the resolution of the images in the dataset is higher. In fact, the model is focused to improve the performance of the human pose estimation, semantic segmentation, and object detection since those tasks require higher accuracy. Therefore, it is undoubtedly that in my case, the model gives outstanding results on the dataset regarding the image classification task. The datasets provided are relatively small range in terms of classes. Most of the features in the datasets are leaves, fruits, or whole trees. The process would be more challenging if the images focus on more different parts of the trees. Moreover, the features of the images in the test dataset are not too different from the trained dataset so it may result in higher accuracy of the output of the model.

4 CONCLUSIONS

The information on the plant species is important for not only agriculture but also medicine. In this paper, we present an approach for plant classification in Southeast Asia by using a high-resolution network. The experiments are evaluated on a public dataset including 26 different species in Southeast Asia. The high-resolution network can obtain high accuracy. By training the model more to serve the purpose, we believe that our trained model can contribute to the variety of applications regarding the classification of plants, especially in the pharmaceutical industry. Plant classification can also be used in detecting the diseases in plants to help with agricultural problems, such as plant disease diagnosis.

REFERENCES

[1] S. O, "Computer vision classification of leaves from swedish trees," no. Department of Electrical Engineering, Computer Vision, Linköping University, 2001.

[2] Wu S, Bao F, Xu E, Wang YX, Chang YF, Xiang QL, "A leaf recognition algorithm for plant classification using probabilistic neural network," *IEEE,* no. IEEE international symposium on signal processing and information technology, 2007.

[3] Z. A. Nilsback ME, "A visual vocabulary for flower classification," *IEEE,* vol. 2, no. IEEE computer society conference on computer vision and pattern recognition, 2006.

[4] Z. A. Nilsback ME, "Automated flower classification over a large number of classes," 2008.

[5] A. S. J. M. A. K. Sonali Agrawala, "Plant Identification using Leaf Image Analysis," no. 3rd International Conference on Internet of Things and Connected Technologies, 2018.

[6] M. Dyrmann, H. Karstoft and H. S. Midtiby, "Plant species classification using deep convolutional neural network," *Biosystems Engineering*, vol. 151, pp. 72-80, 2016.

[7] Jingdong Wang, Ke Sun, Tianheng Cheng, Borui Jiang, Chaorui Deng, Yang Zhao, Dong Liu, Yadong Mu, "Deep High-Resolution Representation Learning," no. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, MARCH 2020.

[8] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in *ICLR*, 2015.

[9] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. E. Reed and D. Anguelov, "Going deeper with convolutions," in *CVPR*, 2015.

[10] A. Krizhevsky, I. Sutskever and a. G. E. Hinton, "Imagenet classification with deep convolutional neural networks.," in *NIPS*, 2012.

[11] K. He, X. Zhang, S. Ren and a. J. Sun, "Deep residual learning for image recognition," in *CVPR*, 2016.

[12] H. Noh, S. Hong and a. B. Han, "Learning deconvolution network for semantic segmentation," in *ICCV*, 2015.

[13] V. Badrinarayanan, A. Kendall and a. R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, p. 2481–2495, 2017.

[14] O. Ronneberger, P. Fischer and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *MICCAI*, 2015.

[15] M. B. a. W. Sujarwo, "Return to Bali," *Second International Conference on Transdisciplinary AI*, 2020.

[16] M. B. a. W. Sujarwo, "Machine Learning in Ethnobotany," *IEEE International Conference on Systems, Man, and Cybernetics* , 2020 .

[17] R. Fisher, "Iris Data Set," *NASA*, 1988.

[18] R. J. W. S. a. V. C. M. Böhlen, "From images in the wild to video-informed image classification," *IEEE International Conference on Machine Learning and Applications*, 2021 .

[19] Smitha S. Karanth, Suvijith S, "Plant Identification Methodologies using Machine Learning Algorithms," 2019.

ICATSD2F.124

AN IMPROVEMENT DESIGN OF MULTI-FUNCTION CONTROLLER FOR HIGH-TECH SHRIMP FARM

BUI THU CAO

Faculty of Electronics Technology, Industrial University of Ho Chi Minh City, Vietnam

buithucao@iuh.edu.vn

Abstract. Aquaculture is a potential sector in the coastal provinces of Vietnam. Shrimp farming is one of the potential fields with high economic value. Shrimp farming technology is developing strongly by applying high technology with modern equipment and machines such as plastic tarpaulin ponds, water fans, aeration systems, siphon pumps, shrimp feeding machines and environment measuring devices. However, the application of high technology has not been fully exploited yet. Most hi-tech shrimp farms today still supervise and control actuators by hand. Shrimp farmers have not paid more attention to the use of automatic controllers for shrimp farming. This paper evaluates the current controllers for shrimp farming. Therefrom, we propose an improvement design of a multifunction controller for shrimp farm with some preeminent features such as flexible device configuration, overload protection and motor damage detection and avoiding waste power consumption on the device. Moreover, the multifunction controller is integrated with IoT getaway to transmit shrimp farm parameters to the cloud and users can monitor and control automatically devices and system of shrimp farm via smartphone app.

Keywords. High-tech shrimp farm, Smart controller, IoT gateway.

1 INTRODUCTION

With terrain length of coastline over two thousand kilometers, aquaculture is a strong field of Vietnam. Shrimp farming is one of the fastest growing industries with high economic value. Shrimp farming techniques are shifting from extensive farming with low-density to super-intensive farming with high-density of shrimp stocks. To do this, it is imperative that farmers must apply high-tech farming with modern equipment and machinery such as plastic tarpaulin ponds, water fans, aeration systems, siphon pumps, shrimp feeding machines and water environment measuring devices, etc.

1.1 Model of a high-tech shrimp pond

Figure 1. Photo of typical high-tech shrimp pond with an area about 500 m^2.

The system and equipment model for a high-tech shrimp pond is arranged as shown in Figure 1. The layout

diagram of the actuator system is as shown in Figure 2. They include two water fans, one aerator system, one fresh water supply pump and one waste siphon pump and one shrimp-feeding machine. The water fans have function of circulating water in the pond. The aerator system has function of creating dissolved oxygen in the water to provide shrimp respiration. The water supply pump has function of pumping fresh water from external settling pond to change the water periodically. The siphon pump adsorbs out the leftovers and shrimp waste. The monitoring and controlling of the actuators are mostly done in manually. There is rarely high-tech shrimp farm (HTSF) in Vietnam that use automation controller for actuators.

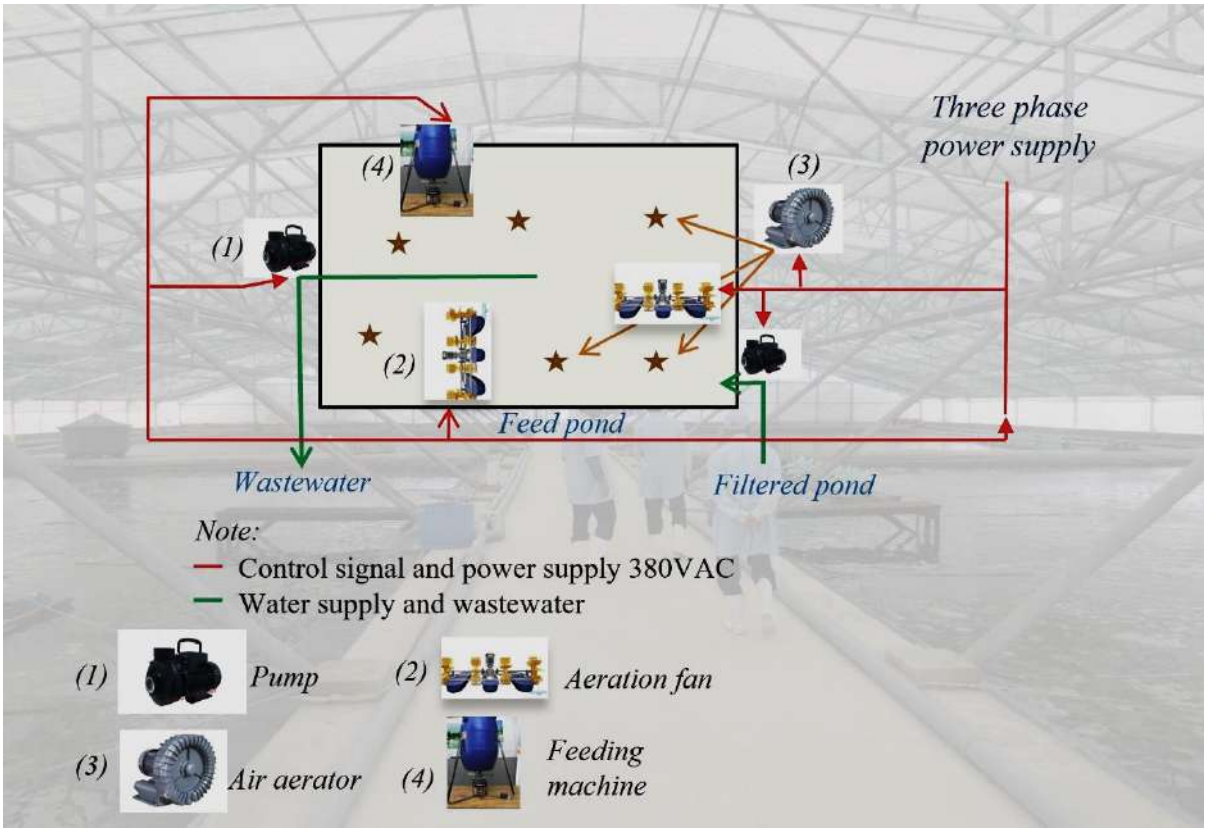


Figure 2. Popular model for equipment and control system for a high-tech shrimp pond.

1.2 State of using automation controller for HTSF

Automation controllers for HTSF are classified into two types of models:

- Large farm model with many ponds. This model uses professional automation equipment and controllers such as DCS (Distributed Control System), PLC (Programmable Logic Controller) and magnetic starter as shown in [1] and [2]. The controller is connected to actuators, water environmental measurement devices and database servers. An application software on computers or smartphone apps are used for supervising and controlling the actuators. This system requires huge investment costs and requires technical expertise in automation control for the operation.
- Small and medium farm model. This model uses an integrated controller which is a microcontroller integrated IoT gateway and electromagnetic relay to control actuators as shown in [3], [4], [5], [6], [7] and [8]. IoT gateway connects to an IoT flatform for storing database in cloud, monitoring, and controlling the actuators via smartphone app. This type of model offers low cost and simple operation without requiring automation technician.

Figure 3. Smart Shrimp Farm Management System of HYDRONEO, Thailand.

Figure 4. Aquaculture Controller of SENECT, Germany.

The integrated controller model is now available from a number of suppliers in the international market, such as Smart Shrimp Farm Management System of HYDRONEO [9] as shown in Figure 3 and Aquaculture Controller of SENECT [10] as shown in Figure 4. The common feature of the controllers of HYDRONEO and SENECT is to allow monitoring of environmental indicators and control actuators such as pumps, fans, and aerators via smartphone app. In Vietnam, there is a typical controller FARMEXT of TEPBAC company [11]. This controller can switch on/off actuators such as pumps, fans, and aerators via smartphone app, as shown in Figure 5 and Figure 6.

Figure 5. Cover photo of controller FARMEXT.

Figure 6. Inside photo of controller FARMEXT.

Through field survey, we found some restriction points of the current integrated controllers for shrimp farms as following:

1) The current controllers on the market are hard configured for output control ports with certain voltage and power parameters. Because of the hard configuration, user must not install actuators with larger or smaller values of the parameters. For example: If the user installs actuators with a larger power than the configured power, the controller will not be able to switch on the actuators at start-up time. If the user

installs actuators with a less power than the configured power, the controller will not effectively protect the actuators from overloading. Therefore, this is a difficulty for shrimp farmers to change the actuators. For example, the aerator they bought with the old technology has a power of 5HP. With the modern technology, the aerator only needs power of 2HP to ensure the oxygen supply to the pond. In this case, the shrimp farmer must contact the manufacturer and require reconfiguring the device for the controller.

2) The actuators operate in brackish water. Therefore, the power connections and the coils inside the motor will rust over time. This causes two problems. Firstly, if the connections are rusted, they will make increasing the impedance of the connections and reducing the current through the motor. As a result, motor is weaker. Secondly, if there is a rust inside the motor, it will cause an electric shock, reducing the coils impedance and increasing the consume current. The result is that motor takes more power.

3) Most current integrated controllers use magnetic starters to switch actuators. This will cause several impairments. The magnetic starters use the switching mechanical contact method so that they will cause ignition and contact corrosion with large power equipment. As a result, the life of the device will reduce, and the controller can be noised by ignition and create errors.

To solve the restrictions of current controllers, we propose a hardware design of a multifunction controller for high-tech shrimp farm.

2 DESIGNS OF THE MULTIFUNCTION CONTROLLER

2.1 Idea of designing the multifunction controller (MFC)

Figure 7. General structure of monitor and control system for HTSF.

General structure of the monitor and control system for shrimp farms is synthesized from the previous studies, as shown in Figure 7. The system is divided into 5 layers. The sensor layer contains sensors which able to measure water environmental parameters such as dissolved oxygen, pH, turbidity, salinity, etc.

The environmental parameters received from the sensor layers are transmitted to the IoT gateway layer. The actuator layer read control commands which are sent from the IoT gateway layer to control the actuators. The IoT gateway layer have ability to transmit environmental sensor data to the database on cloud server layer and receive control commands through changed states of data fields on the cloud server layer. The

application layer has function of setting parameters and controlling the device through changing the data state of the actuators on the smartphone app. The data states changed by users are transmitted to database on cloud.

Based on the general structure of the monitor and control system for HTSF as shown in Figure 7, we propose additional design features for the MFC as follows:

- Use solid-state relay (SSR) instead of magnetic starter. By using SSR, the switching of high-power actuators will not cause sparks at the switching contact points. This will increase the lifespan and especially the stability of the controller. While the cost of SSR is cheaper than the magnetic starter with the same power.
- Design a flexible configuration. The controller using SSR will allow user to reconfigure the device parameters such as voltage and power without the need to change the control hardware.
- Mount the phase current sensors at the input source of motors to monitor the operating status as well as promptly detect equipment damage.
- Design a controller for a common high-tech shrimp pond, with an area of 500m^2, as shown in Figure 8. The actuators include: two three-phase water fans, one supply pump and one three-phase aeration, one waste siphon pump and one automatic shrimp feeding machine.
- Ability to expand the control link with other controllers through Lora or Zigbee interfaces.
- Use a solar panel system to supply a DC power for the controller. This will increase stabilization and backup battery for the controller. When the supply power is lost, the controller still works and sends alarm to user.

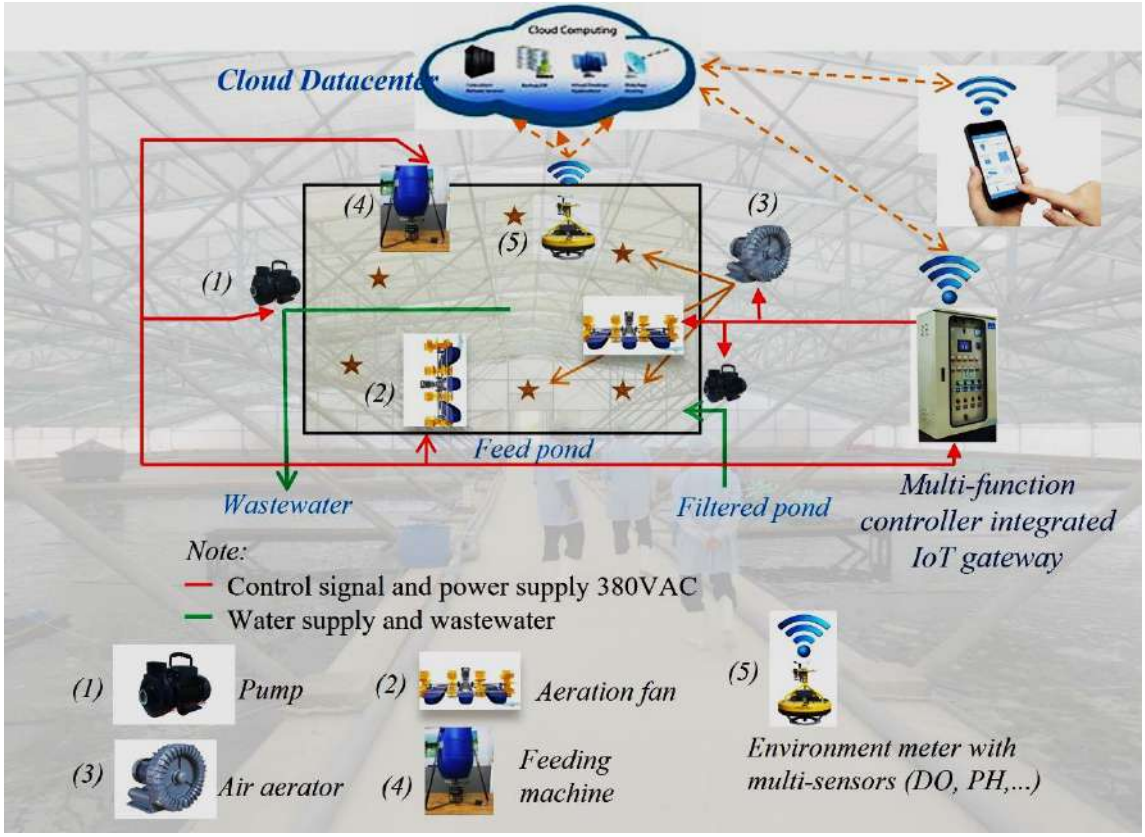


Figure 8. A proposed model of devices and control system for high-tech shrimp pond.

2.2 Structure design of the MFC

The structure design is shown in Figure 9, including the following functional modules:

- **Link-MCUs module:** This module is designed as an option. If it is linked to an external MFC, it will use the Lora or Zigbee transceiver modules as shown in Appendix 1. The link is used for the expansion control for the shrimp farm. In this case, the MFC will function as an intermediary IoT gateway node to connect an external MFC to database on the cloud. If the option is Local-MCU, we will use a microcontroller module

to control the MFC. Arduino mega 2560 module is selected for Local-MCU because the system design of the MFC is required twelve analog inputs for current sensors and 15 DI/O for 9 buttons, 4 temperature sensors and 2 sub-devices.

- **IoT gateway module:** NODEMCU ESP8266 is seletecd, as shown in Appendix 1. This module is designed as a microcontroller with built-in Wi-Fi. This module has high stability and low cost.
- **Modem 3G/4G:** The selected module is Alcatel MW40 Link Zone 4G LTE. This module has low cost and Wi-Fi connection can work for a long time without resetting the device.
- **Driver module:** This module is structured as Figure 10, with the following parts:

○ ***The sensor signal acquisition circuits.*** It includes acquisition circuits for phase current sensor signals, a data acquisition circuit for temperature sensing signal of the SSRs and a state sensing circuit of buttons on the keyboard.

○ ***The activator circuits for sub-device.*** It includes activator circuits for SSRs and activator circuits for sub-devices.

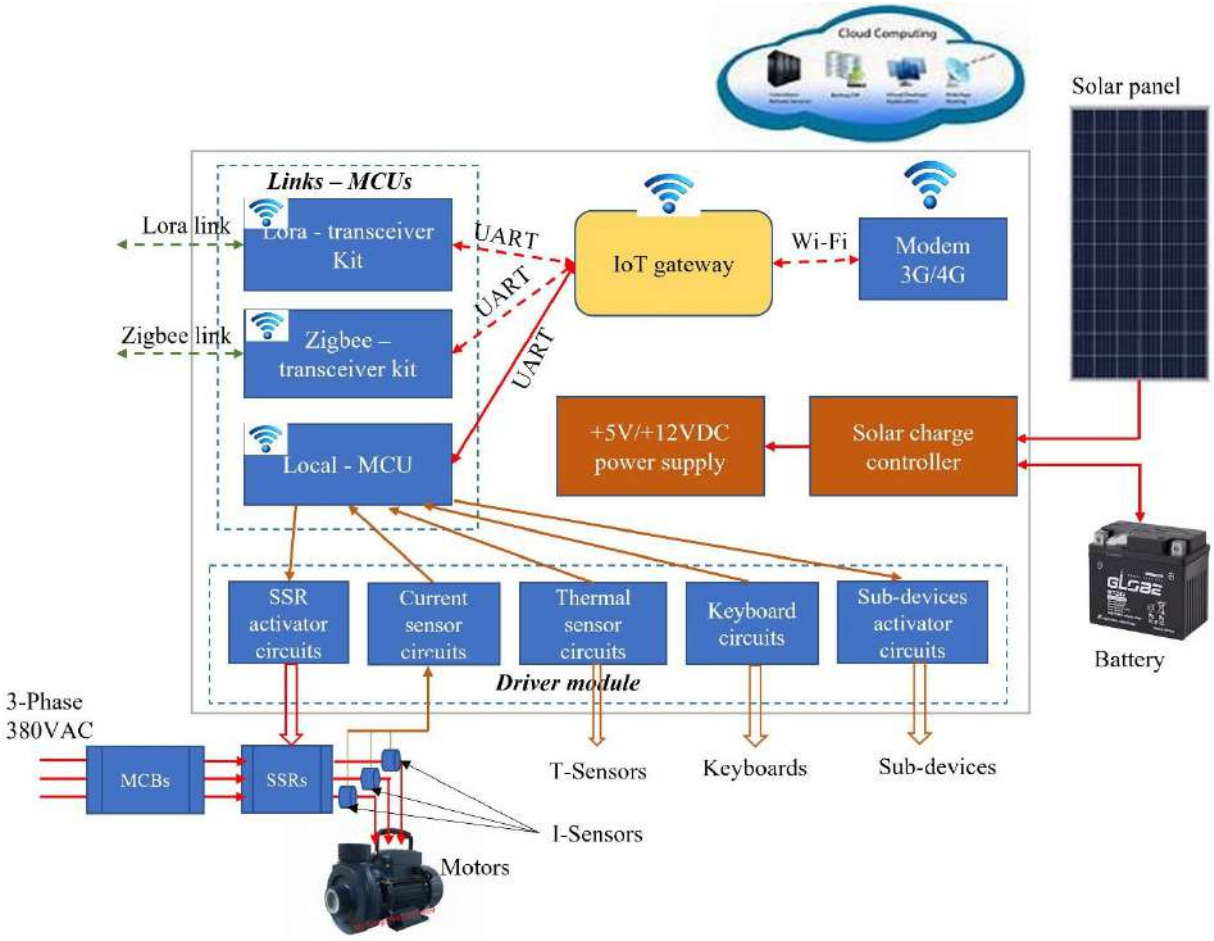


Figure 9. Structure design of the MFC.

- **SSR module:** The SSRs used are the type that can be switched on and off by a DC electrical signal. These SSRs are used to replace magnetic starters for actuator motors. The power of shrimp farm equipment is in statistic range of 2HP to 9HP. Therefore, the maximum starting current of the controller is referred in [12] and calculated as following:

$$I = 10 \times \frac{P}{U \times 3(phase)} = 10 \times \frac{9x750}{380 \times 3} \approx 60\text{A}.$$

So, we use HSR-3D704Z, as shown in Appendix 1. It is a three-phase SSR operating at 380VAC/70A.

- **I-Sensor module:** This module includes 12 output phase current sensors for 4 SSRs. The use of phase current sensors is to manage circuit breaker and motor overload protection. Besides, the use of phase current sensor will help monitor equipment damage. Since the maximum start current of the controller is 60A, we

choose a 100A AC current sensor, as shown in Appendix 1.
- **T-Sensor module:** This module includes 4 temperature sensors to monitor, warn and handle overheating protection for SSRs. The temperature sensor is selected, as shown in Appendix 1.
- **Sub-device module:** This module is a 8-channel optocoupler relay, as shown in Appendix 1. It is used to switch power 220VAC for four heat sink fans of the SSRs, a waste siphon pump and a shrimp feeding machine with a lower power of 0.25HP.
- **MCBs (Miniature Circuit Breaker)**: To facilitate for installation process as well as increase the safety of the system, the MCBs module is connected including a total MCB and four three-phase MCBs for the SSRs. Because the maximum starting current of the controller is 60A, we choose a MCB type B which a nominal current value is 16A, as shown in Appendix 1.
- **Power supply module:** The power module is voltage regulator from solar battery to +5VDC to supply for all modules in the controller. Based on the current consumption statistics of modules in Table 1, we see that the total current consumption is 765 mA. Therefore, power module XL4015 is selected as shown in Appendix 1. This module is compact, low voltage drop and capable of drawing high current up to 5A. This power module meets enough power for the controller.

Table 1. Statistics of current consumption of the modules in the controller.

No	Name of module	Current (A)
1	Lora-transceiver kit	20m
2	Zigbee-transceiver kit	30m
3	Arduino mega 2560	27m
4	NODEMCU ESP 8266	126m
5	Modem Wifi 4G Alcatel MW40	200m
6	4 HSR-3D404Z	56m
7	4 DS18B20	6m
8	4 Sensor YHDC STC013	200m
8	Relay for sub-devies	100m
	TOTAL CURRENT	**765m**

- **Solar charging module:** The controller is connected to the cloud to online monitor and control the actuators for 24/7, so its power must use solar cells. In case of power line failure, the controller still works and sends an alert to the smartphone. The solar charging module is selected a type which has parameters 130W and 10Ah to be able to charge for the battery 7A. This battery is capable of maintaining on the power source for more than 7 hours in case the power line is lost.

2.3 Layout design of the driver module

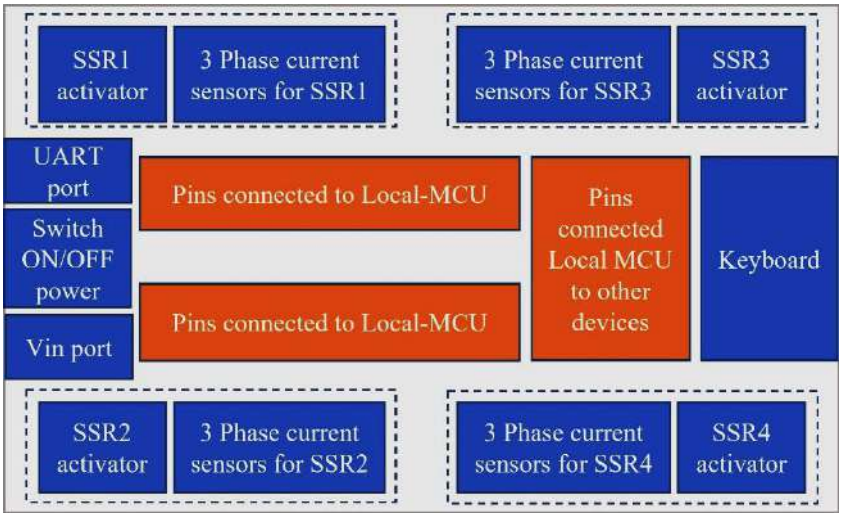


Figure 10. Layout part of driver module.

The architecture of the driver module is shown in Figure 10. It includes: four driver circuits for SSRs, four receiver circuit assemblies for three-phase current sensors, keyboard module, UART port for connecting IoT gateway to Local-MCU, Switch ON/OFF power port, Vin port for +5V/+12V power supply input, 2 PIN connection ports with Local-MCU module and a PIN connection port with other devices. The other devices include Sub-device module and T-sensor module.

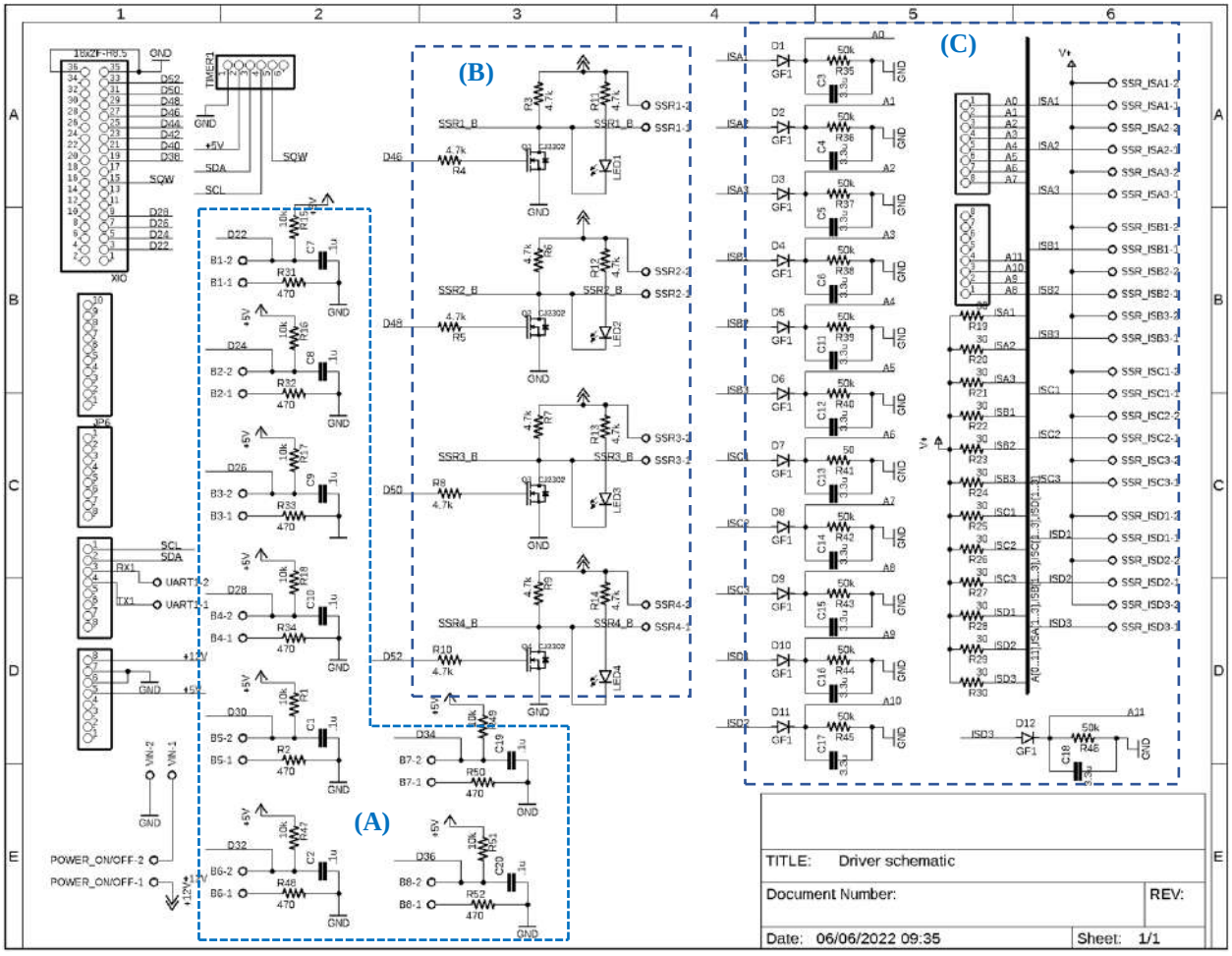


Figure 11. Schematic of the driver module

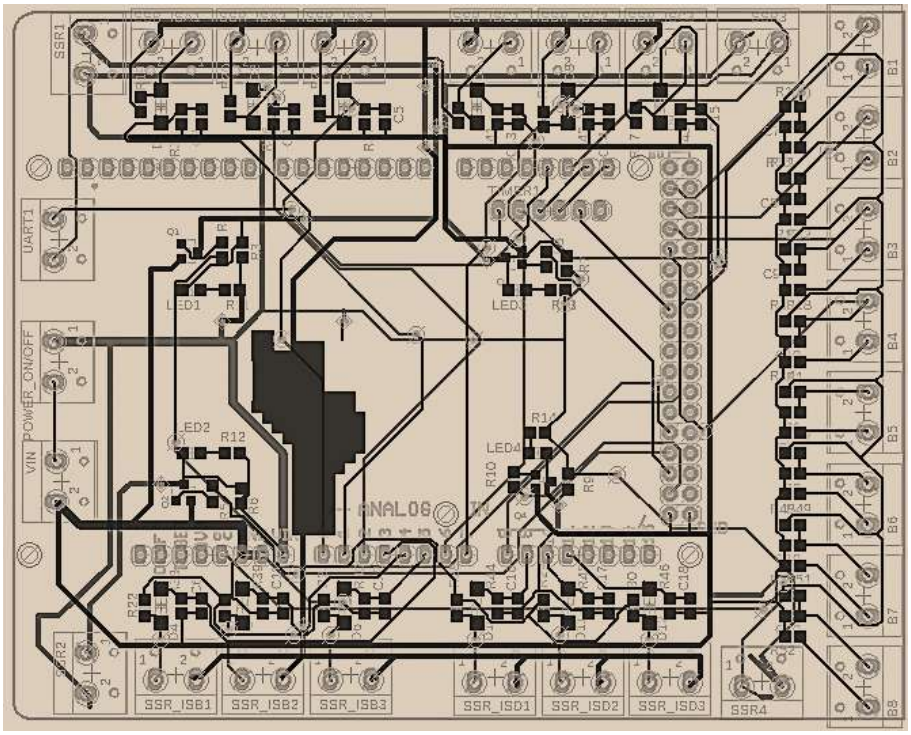


Figure 12. PCB driver board of the driver module.

The circuit diagram design of the driver module is shown in Figure 11. All ports connected are according to the corresponding ports of the ATmega 2560. Block A is keyboard circuits. There are an anti-vibration circuits for 8 buttons. A pair of capacitors 0.1uF and resistor 10kΩ are used for filtering noise in anti-vibration circuits. Block B, the driver circuits using MOSFETs and resistor 4.7kΩ to switch on/off four three-phase SSRs. Block C includes 12 signal acquisition circuits from three-phase current sensors. The signal from the phase current sensors is passed through the rectifier diode and then through the filter circuit to measure the average amplitude value. This value is used to interpolate the phase current value. A pair of capacitors 3.3uF and resistor 50kΩ are used for filtering 50Hz AC power line. The design of the PCB circuit as shown in Figure 12. The layout of components is the same position as in the driver module structure in Figure 11.

2.4 Design of the embedded software for Arduino Atmega 2560

The control embedded software is designed to have the following functions:

- Read the updated data from the cloud and transmit control commands to the actuators such as water fans, supply pumps, aeration system, waste siphon pumps, automatic shrimp feeders.
- Scan the keyboard and process changed states of the buttons. There are 8 control buttons, four buttons for four three-phase motors, three buttons for sub-devices and a button to set AUTO or MANUAL mode. In AUTO mode, the MFC will control in automatically all the actuators. In MANUAL mode, users can directly control all actuators through the keyboard. This mode is used in case the Wi-Fi connection is lost.
- Allow to reconfigure the nominal parameters of voltage and power for the actuators on the mobile app. These parameters will be updated to the MFC during operation time.
- Read temperature sensor values and handle overheating. The temperature sensor measures the temperature of the heat sink for SSR. The allowable working temperature range of SSR is shown in Appendix 2. In case of overheating, over 60 Celsius degrees, the controller will warn with an alarm buzzer and overheat status on the smartphone app.
- Read the phase current sensors of the SSRs and handle overload or failure. The handling include:

a. In case the phase current is greater than ten times the rated current of the device at start-up time, the controller will automatically cut off the device power, warn with an alarm buzzer and report the overcurrent

status on the smartphone app.
b. In case the phase current is more than two times the rated current of the device in operation time, the controller will alert the high current status on the smartphone app.
c. In case the value of phase sensor current is less than ½ of the rated current value, the motor may spin weakly or stop working; the system will alert the alarm buzzer and report a malfunction on the smartphone.
• Read data sent from other controllers over the Lora/Zigbee link and update all device state variables to database on the cloud for extending control connectivity.

2.5 Design of smartphone control app

Blynk.io is a powerful IoT platform, it supports for storing database on cloud and developing smartphone app in iOS or android with a reasonable cost. So, we selected Blynk.io as a tool to design, store database and create the mobile control app.
The smartphone control app is design as Figure 13 with the following fucntions:
- **OPEN SYS** button has the function of controlling ON/OFF operation of the entire system.
- **Wifi** icon has the function to display the status of Wi-Fi. If it is white, there's no Wi-Fi connection and if it's yellow, it's connected Wi-Fi.
- **Alarm** presents a alarm. If it's white, there's no alarm and if it's yellow, a buzzer alarm is done.
- Input texts **Voltage**, **Power** and **Phase** are used for setting the voltage, power and phase of actuator devices.
- **Dev1**, **Dev2**, **Dev3** and **Dev4** are the function buttons of switch on/off power for the three-phase actuator devices.
- **I_Dev1**, **I_Dev2**, **I_Dev3** and **I_Dev4** are the phase current indicators on the actuator devices.
- The two-line text box shows the command execution. The top line shows the progress of the control command execution, and the bottom line shows the actual status of the devices. The actual device state is represented by a five-char array. The first four weights represent the status of 1 (ON) or 0 (OFF) for the respective devices, and the last weight represents the AUTO (A) or MANUAL (M) mode. For example: In Figure 13, on the left, the text box shows that only Dev1 is turned on and the system is operating in the MANUAL mode.
- **Temp1**, **Temp2**, **Temp3** and **Temp4** display the temperature of the SSRs.
- **pH**, **DO**, **TU** and **Sa** display the pH, dissolved oxygen, turbidity, and salinity values of the water environment of shrimp pond.

Figure 10. A capture of the smartphone control app for the MFC. From left to right, a capture of MANUAL mode, AUTO mode and the lower part of the app page.

Besides, there are two setting types of automation control on the smartphone apps. “Schedule” setting allows the system to automatically turn on or off the devices according to a preset schedule in daytime and “Device state” setting allows the system to automatically turn on and off the device when the pH, DO and salinity values are above or below the thresholds. For example, we can set the threshold of dissolved oxygen in pond water to be above 5 mg/l. If the dissolved oxygen is below this threshold, the system will alert on smartphone and alert buzzer for users. The manipulation of these settings can be done easily on the smartphone app.

2.6 Installation of the control cabinet

The controller is installed in an electrical cabinet with dimensions of 30 cm x 60 cm x 120 cm. For ventilation and heat dissipation, the electrical cabinet is designed with two vents on two sides above and four vents for four cooling fans at the bottom as shown in Figure 14. This arrangement will create a convection stream for the air adsorbed from the bottom fans to go up and out through the two upper vents.

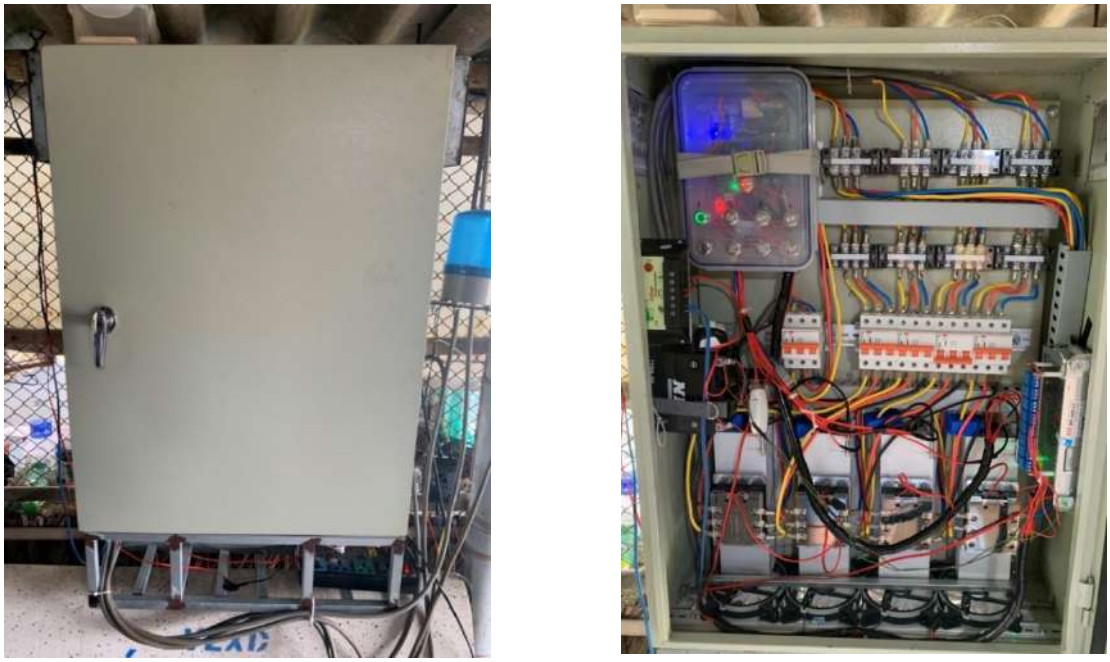


Figure 14. On the left is outside picture and on the right is inside picture of the cabinet.

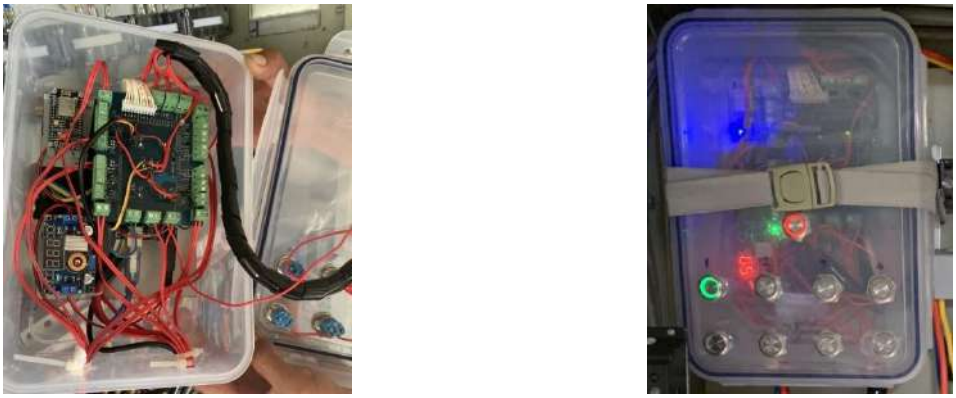


Figure 15. Image of the central control block. The left image shows the installation of components modules, and the right side shows the front panel of the central control box with the keyboard.

The MFC system is shown in Figure 14. From the left to the right and from the top to the bottom, it includes Wi-Fi module, central control box, terminal blocks connecting the actuators to the SSRs output below, the terminal blocks connecting the mains input power to the MCBs below, the outputs of the MCBs connected to the SSRs below and 4 cooling fans in the bottom. Besides, the left side middle attaches the solar charging module and the battery 7A. The right-side middle attaches the sub-devices module.
On the left of Figure 15 shows the modules installed in the central control block. They include: IoT gateway module, power module, driver module superimposed on the Arduino Atmega 2560 module. On the right of Figure 15 shows the keyboard with 9 buttons: the top button is used to set the AUTO/MANUAL mode, the four upper buttons are used to control the large power actuators and the lower 4 buttons is used to control the sub-devices.

3 RESULTS

Figure 16. Photo of the field test of a multi-function controller for shrimp pond equipment in Binh Dai district, Ben Tre province, Vietnam.

The MFC was experimentally operated at a HTSF as shown in Figure 16. The shrimp pond applied with the MFC has the following actuators: 4 aeration fans with a power of 5HP (3.75KW), 1 shrimp feeding machine, 1 siphon machine.

The MFC is evaluated for operability with the following results:

- The smartphone app works fine. It allows user to reconfigure and control actuators according to the actual requirement. AUTO and MANUAL modes are good working.
- The phase sensor current and temperature values of the SSRs are displayed accurately. It makes user easily to monitor the quality of the equipment.
- The overload protection and warning functions are evaluated by setting the power of 3HP motor, but we connected the output SSR to the 5 HP power device. The results show that it is working properly, the controller shows an overloaded device error and a buzzer alarm.
- When the activating motor is forced power off by removing 1 or 3 output phase wires of the corresponding SSR, the controller will report equipment failure on smartphone and buzzer alarms. In HTSF, with high stocking density, just stopping aeration for few hours is enough to make shrimp die from lacking oxygen. On day or during the night, when shrimp farmer sleeps, he cannot detect when the water fans or the aerators are damaged or stopped. In this case, the MFC will help him to detect and promptly warn whether the motors are overloaded or not working. This function is especially important and useful.
- During the installation and operation of the controller at field, the farmer discovered that a 5HP motor has a problem because it has a consumption current of over 5A, instead of the rated current value of this device is 3.3A (can be up to 4A when it has load). After replacing the good motor, this motor operates at a steady current level the same as remaining motors of 4A.

Through the testing process from April 2022 up to now, the shrimp farm owner assessed that the MFC has working well, installing, and configuring the device was simple with many convenient functions such as: allowing surveillance, warning damage, programming automatic control in diversely and easily. The testing process is performed with three 5HP devices and a 7HP device for 30 experimental times. We use multimeter Uni-T UT201+ for measuring phase current. It is shown in statistics of Table 2. The average operating current in range from 3.5A to 3.38A for 5HP devices and 5.2A for 7HP device. These values are reasonable because all devices have been rusted follow the time. So, the operating current of devices are different. The tolerance of current measurement between the real current measured by the multimeter Uni-T UT201+ with the MFC is about ±0.5A. Because the displayed value is rounded one unit for convenience of observation, so the tolerance of current measurement is reasonably about ±0.5A.

Table 2. Statistic of the phase current sensors

Device	Power	Reated current	Average operating current	Tolerance of measurement
Dev1	5HP	3.3A	3.8A	±0.5A
Dev2	5HP	3.3A	3.7A	±0.5A
Dev3	5HP	3.3A	3.5A	±0.5A
Dev4	7HP	4.6A	5.2A	±0.5A

4 CONCLUSIONS

The goal of designing the MFC is to perfect the control features to meet the practical requirements for high-tech shrimp farms. The proposed controller has some improvements to compare with the existing controllers such as:

- Using SSR instead of magnetic starter makes the control process safer, more stable and increases the durability of the device.
- Using phase current sensors for each actuator motor helps user to detect damage in time and avoid loss power during the operation.
- The controller allows user to reconfigure according to the actual actuators for a shrimp pond. So, it is very flexible and easy to install and operate for user who is not technician in automation.

APPENDIX

1 List of components are assembled in the MFC

Figure 17. Kit RF UART Lora SX1278.

Figure 18. Kít RF UART Zigbee CC2530+PA V2.

Figure 19. Arduino mega 2560 module.

Figure 20. NODEMCU ESP 8266 module.

Figure 21. Modem Wifi 4G, Alcatel MW40.

Figure 22. Solid state realy HSR-3D704Z.

Figure 11. T-Sensor module, DS18B20.

Figure 24. I- Sensor module, YHDC STC013.

Figure 25. Power supply module.

Figure 26. Solar charging module.

Figure 27. Heat sink fan for SSRs.

Figure 28. Relay module for sub-devices.

Figure 29. MCB LS BKN 16A – 3 phases.

2 Load current characteristics of HSR-3D704Z

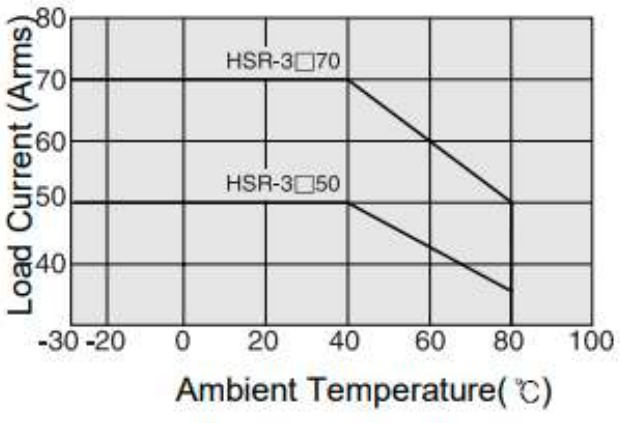


Figure 30. Response graph of load current vs. temperature of HSR-3D704Z.

ACKNOWLEDGMENT

It is acknowledgement that Mr. Tran Hoang Vinh, director of Thanh Phu farm, Huy Thuan Co., LTD, Ben Tre province. He has provided technical advice to the author for shrimp farming. The author also would

like to thank Mr. Nguyen Van Dien, the owner of a shrimp farm in Binh Dai district, Ben Tre province for supporting the installation and commissioning of the project's equipment.

REFERENCES

[1] J.-H. Huh, "PLC-based design of monitoring system for ICT-integrated vertical fish farm," *Human-centric Computing and Information Sciences, Springer,* vol. 7, no. 20, 2017.

[2] N. Mahalik and K. Kim, "Aquaculture Monitoring and Control Systems for Seaweed and Fish Farming," *World Journal of Agricultural Research,* vol. 4, no. 2, pp. 176-182, 2014.

[3] D. S. Simbeye, . J. Zhao and S. Yang, "Design and deployment of wireless sensor networks for aquaculture," *Computers and Electronics in Agriculture, Elsevier,* vol. 102, pp. 31-42, 2014.

[4] S. P.S. and R. V.S. , "Automatic monitoring and control of shrimp aquaculture and paddy field based on embedded system and IoT," in *2017 International Conference on Inventive Computing and Informatics (ICICI), IEEE,* Coimbatore, India, 2017.

[5] R. D.R. and A. Gunawan, "Proposed Design of an Automatic Feeder and Aerator Systems for Shrimps Farming," *International Journal of Materials, Mechanics and Manufacturing*, vol. 6, no. 4, pp. 277-280, 2018.

[6] N. T. Duy, N. D. Tu, T. H. Son and L. H. Khanh, "Automated monitoring and control system for shrimp farms based on embedded system and wireless sensor network," in *IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, Coimbatore, India, 2015.

[7] T. Khaoula, R. A. Abdelouahid, I. Ezzahoui and A. Marzak, "Architecture design of monitoring and controlling of IoT-based aquaponics system powered by solar energy," in *Procedia Computer Science, Elsevier*, 2021.

[8] J. Capelo, E. Ruiz, V. Asanza, T. Toscano-Quiroga and N. N. Sánchez-Pozo, "Raspberry Pi-based IoT for shrimp farms Real-time remote monitoring with automated system," in *2021 International Conference on Applied Electronics (AE), IEEE*, Pilsen, Czech Republic, 2021.

[9] HydroNeo, "https://hydroneo.net/," HydroNeo. [Online]. [Accessed 22 June 2022].

[10] "https://www.senect.de/en/," SENECT. [Online]. [Accessed 22 June 2022].

[11] "https://tepbac.com/tin-tuc/full/tu-dien-cabinext-nguoi-nuoi-tom-co-can-mua-khong-32526.html," TEPBAC. [Online]. [Accessed 22 June 2022].

[12] V. K. Mehta and R. Mehta, Principles of Electrical Machines, S. Chand, 2006.

ISBN: 978-604-920-163-9

SÁCH KHÔNG BÁN