



**ISIPS2023**

*17th International collaboration Symposium on Information, Production and Systems  
and Workshop on Next-generation Technology for AI Application Strategic Industries:  
MASR 2023 (Mobility, AI, Semiconductor, and Robot)  
16<sup>th</sup> – 17<sup>th</sup> November 2023, Kitakyushu, Fukuoka, JAPAN*

**PROCEEDINGS OF THE  
WORKSHOP ON**

**NEXT-GENERATION  
TECHNOLOGY FOR AI  
APPLICATION STRATEGIC  
INDUSTRIES**

**MASR 2023 (MOBILITY, AI,  
SEMINCONDUCTOR, AND  
ROBOT)**



## Greeting Message

Dear All,

I am truly pleased to hold the 2023 International collaboration Symposium on IPS Waseda University (ISIPS) and the Special Academic Workshop MASR2023 amid the autumn atmosphere.

Above all, I would like to express my gratitude to Professor Hee-hyeol Lee of IPS Waseda University and Professor Jae-hyun Park of Inha University for providing this workshop. In addition, I would like to welcome the researchers who attended the MASR2023.

Now we are facing a new paradigm of change and bold innovation in the high-tech future industry.

To respond boldly to this demand, DSC Platform and Korea University are playing a role as institutions that explore the future, including Mobility, AI, Semiconductor and Robots create value for high-tech future industries.

But the complex challenges demanded by this era cannot be solved alone, and not by one university and country.

I hope that through the MASR2023, research exchanges will be conducted with Waseda, Inha University and DSC Platform to contribute to the country and future industries, and that the challenges faced will be solved.

Thank you.

**Choong-ho Cho**

Professor at Korea University,

Head of the Daejeon-Sejong-Chungnam (DSC) Regional Innovation Platform

Autonomous Driving System Project Division



The following are the proceedings of the Workshop on Next-Generation Technology for AI Application Strategic Industries: MASR 2023 (Mobility, AI, Semiconductor and Robot). The workshop took place on the 17<sup>th</sup> of November 2023 in Kitakyushu, in conjunction with the International collaboration Symposium on Information, Production and Systems (ISIPS 2023).

The workshop was supported by

- the Regional Innovation Strategy (RIS) through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (MOE) (2021RIS-004)

in addition to the support for ISIPS 2023 by

- Waseda University Information, Production and Systems Research Center (IPSRC),
- the International Office of Waseda University under a Super Global Universities (SGU) grant for organisation of international symposia and
- a FAIS convention grant for the organisation of symposiums.

The workshop welcome 17 papers by authors from institutions in Korea, and Japan, presented during 2 sessions.

These proceedings being automatically generated in conjunction with ISIPS 2023 proceedings, the numbering of pages starts after the last page of ISIPS proceedings.

We thank again all participants in MASR 2023. We wish you a pleasant reading.

Kitakyushu, the 15th of November 2023

MASR 2023 Programme Chairs

Choong-ho CHO, Korea University, Korea

Hee-hyol LEE, Waseda University, Japan

Jae-hyun PARK, Inha University, Korea



<b>W1-1: A study on the development of a driving-level model for autonomous driving based on ODD scenarios</b> <i>Dae-Kug Lee, Min-Jae Jeon, Dong-Hoon Lee, Dae-Gyu Lee, Ki-Hyeok Jeong, Choong-Ho Cho</i> (Korea University) .....	11
<b>W1-2: Survey of Utilizations, Challenges and Thinking of Neural Networks in Solving Path Planning Problem</b> <i>Yuan Huang, Cheng-Tien Tsao, Hee-Hyol Lee</i> (Waseda University) .....	15
<b>W1-3: Fire/Non-fire false alarm reduction method using weighted cost function</b> <i>Saeme Byun, Kim Haneul, Changhyun Lee, Hunkee Kim</i> (Inha University) .....	21
<b>W1-4: Optimum Driving Algorithm to Correspond with Various Outdoor Environment</b> <i>Yeong Ho Ha, Hyun Seo Jang, Young Eun Song</i> (Korea University) .....	25
<b>W1-5: Theoretical Back Grounds for Efficiency of Robust Design in Multi-Level Optimization</b> <i>Masao Arakawa</i> (Waseda University) .....	29
<b>W1-6: Traffic Context Augmentation Dataset for Roadway Environment Situation Awareness of Autonomous Driving</b> <i>Sunghwan Kim, Ir. Made Sudarma, Deden Witarsyah Jacob, Minho Jo</i> (Korea University) .....	35
<b>W1-7: Portable ECG Monitor on Edge Device using TensorFlow Lite</b> <i>Cheongwan Hwang, Nagyoun An, Jaehyun Park</i> (Inha University) .....	41
<b>W1-8: Multi-Input Shape CNN-based Application Traffic Classification</b> <i>Ui-Jun Baek, Jee-Tae Park, Jeong-Woo Choi, Myung-Sup Kim</i> (Korea University) .....	45
<b>W1-9: Optimal Logistics Distribution Algorithm for Autonomous Delivery for Local Logistics Hub System</b> <i>Yeong-Ho Ha, Seong-Min Kim, Hyun-Seo Jang, Young-Eun Song</i> (Korea University).....	49

<b>W2-1: Robot Car Motion Control Leveraging Hand Gesture Recognition with MediaPipe</b> <i>Jianan Xie, Yue Yi, Kenji Hashimoto</i> (Waseda University).....	53
<b>W2-2: A Dynamic Human Following Method Based on Ultra-Wideband System for Mobile Robots: Application and Strategies in Librarian Assistance</b> <i>Yun Kang, Taehoon Kim, Chengyan Gu, Jinchul Roh, Joono Cheong</i> (Korea University) .....	57
<b>W2-3: An Advanced Motor-less Walking Assist Harnesses Power by Integrated Dual-Slider Lever Mechanism</b> <i>Xiuyuan Wu, Keisuke Osawa, Yuntian Wang, Eiichiro Tanaka</i> (Waseda University) .....	61
<b>W2-4: Development of a Social Coexistent Two-Wheeled Inverted Pendulum Balancing Robot</b> <i>Chengyan Gu, Taehoon Kim, Joono Cheong</i> (Korea University) .....	65
<b>W2-5: Development of a Stepper Motor Powered Lower Limb Exoskeleton with 3-DOF Kinematic-Biological Matched Hip Joint Structure</b> <i>Yuntian Wang, Yifan Fang, Xiuyuan Wu, Keisuke Osawa, Eiichiro Tanaka</i> (Waseda University) .....	69
<b>W2-6: Computer vision based surface defect inspection of injection molded products</b> <i>Hee Min Noh, Pyeon MinJi, Chang Hyun Lee, Hun Kee Kim</i> (Inha University).....	73
<b>W2-7: Graphene nanostructures grown on stainless steel by plasma chemical vapor deposition</b> <i>Siyu Jia, Jun Kameoka, Kenji Ueda</i> (Waseda University) .....	77
<b>W2-8: A study on the high-aspect-ratio oxide etching characteristics using a 2 MHz bias power ICP system</b> <i>Kwang-Ho Kwon</i> (Korea University).....	81



# **SESSION W1: ORAL PRESENTATIONS**

## **CHAIRS:**

**PROF. YOUNG-EUN SONG, KOREA UNIV.**

**PROF. MASAO ARAKAWA, WASEDA UNIV.**

**17 NOV 2023 (FRI.) 13:10-14:50**

**ROOM S153-154**



# A study on the development of a driving-level model for autonomous driving based on ODD scenarios

Dae-Kug Lee, Min-Jae Jeon, Dong-Hoon Lee, Dae-Gyu Lee, Ki-Hyeok Jeong, Choong-Ho Cho

Department of Computer and Information Science

Korea University, Republic of Korea

[daekuglee@korea.ac.kr](mailto:daekuglee@korea.ac.kr), [93jmj@korea.ac.kr](mailto:93jmj@korea.ac.kr), [kinno21@korea.ac.kr](mailto:kinno21@korea.ac.kr),  
[sayin@korea.ac.kr](mailto:sayin@korea.ac.kr), [knj6464@korea.ac.kr](mailto:knj6464@korea.ac.kr), [chcho@korea.ac.kr](mailto:chcho@korea.ac.kr)

## Abstract

In this paper, safe driving ODD scenarios were developed based on the threat factors of unexpected situations on real roads in Sejong City, and based on the derived ODD scenarios, the road situation risk and autonomous driving system implementation difficulty were analyzed by object/event. Based on the analysis results, we developed a model to display the level of operation for each segment of the autonomous shuttle service route.

**Topics:** information and communication model, computational intelligence

## Introduction and background

Self-driving cars initially replaced the driver's role with machines. To realize this, the ability to respond to the driving environment of the car and objects and situations in the driving environment is essential. For this, it is necessary to define Operational Design Domain and Object and Event Detect and Response.[1] Operational Design Domain(ODD) is a scope to define the specific operating conditions of the autonomous driving system(ADS). It is an essential concept for securing the safety of autonomous vehicles. Object and Event Detection and Response (OEDR) refers to the ability of ADS to respond in a specific situation with an object while driving. ADS's function is to accurately recognize dynamic and static objects within the range of ODD, such as vehicles (including emergency vehicles), pedestrians, cyclists and motorcyclists, animals, and road installations, and respond to possible situations.[2] Therefore, before developing an autonomous vehicle, ODD and OEDR for the road driving environment must be defined to ensure the safety of passengers and the safety of the autonomous vehicle.

## Experiments, results, analysis of the results

The U.S. National Highway Traffic Safety Administration (NHTSA) divided autonomous driving systems' operational design domain (ODD) into six major categories: road infrastructure, operating constraints, objects/things, environmental conditions, connectivity, and operating zones.[3]

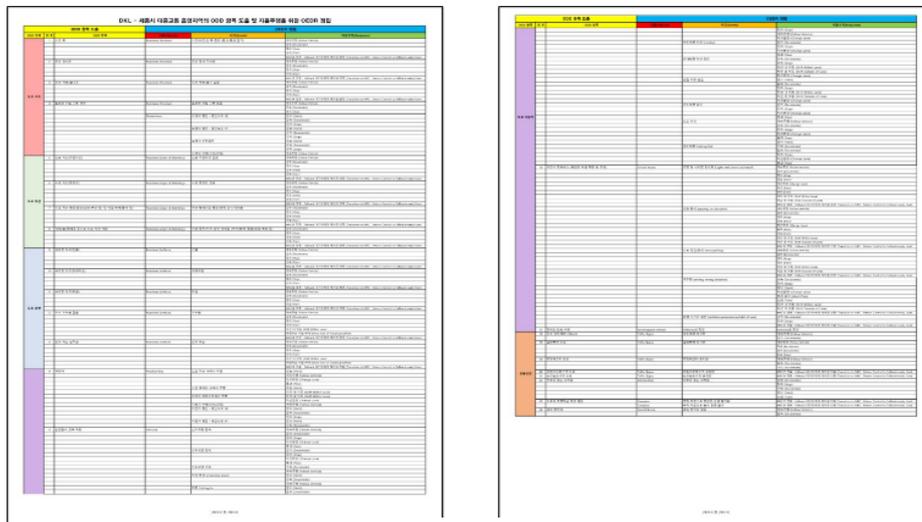


Fig.1: ODD information collection and analysis of Sejong City

Referring to this, as shown in Figure 1, ODD items were derived by collecting and analyzing road traffic environment information in the shaded area of public transportation in Sejong City. We defined ODD (Operation Design Domain) and OEDR (Object and Event Detection and Response) by collecting and analyzing threat factors for unexpected situations on real roads through road field surveys by selecting five public transportation shaded areas in Jochiwon-eup, Sejong City.[4]

We developed ten safe operation ODD scenarios for a safe self-driving shuttle service. We supplemented safe operation ODD scenarios using traffic accident images based on road traffic environment information. Figure 2 shows the procedure for deriving a safe driving scenario based on the threat factors of an emergency on an actual road in the shaded area of public transportation in Sejong City. [4]



Fig.2: Procedure for deriving safe driving scenarios based on road emergency threat factors

To classify the shaded areas of public transportation in Sejong City based on the ODD scenario, we analyzed the risk of road conditions by object/event and the difficulty of autonomous driving system implementation. We classified road conditions into event prediction difficulty, traffic accident risk, event frequency, and violating traffic laws. Furthermore, we divided the autonomous driving system implementation field into recognition, judgment, and control, and we analyzed the difficulty for each category. We divided difficulty, risk, or frequency of occurrence into three levels: high (10 points), average (5 points), low (3 points), and 10 points for violations of traffic laws. Table 1 shows the scores calculated for each analysis item.

Cat.	Item	Road Condition				Implementation of ADS			Score	Level
		Difficulty of Predicting Events	Risk of Traffic Accident	Event Frequency	Violation of Traffic Laws	Cognitive Difficulty	Judgment Difficulty	Control Difficulty		
Road User	Reverse Driving(Motorcycle)	10	10	3	10	10	10	10	63	4
	Reverse Driving(Bicycle)	10	10	10	10	10	10	10	70	4
	Reverse Driving(Kickboard)	10	10	10	10	10	10	10	70	4
	Reverse Driving(Cart)	10	10	3	10	10	10	10	63	4
	Illegal Parking(On Road)	5	5	10	10	5	10	5	50	3
	Illegal Parking(Bus Stop)	5	5	10	10	5	5	5	45	3
	Violation of Priority at no-signal intersection	10	5	10	10	10	10	10	65	4
	Violation of Priority at Roundabouts	10	5	10	10	10	10	10	65	4
	Driver's License Training Vehicle	10	5	3	0	5	5	5	33	2
	School Bus	10	5	5	0	5	5	5	35	2
Person	Vehicle for the Disabled	10	5	3	0	5	5	5	33	2
	Jaywalking	10	10	10	10	10	10	10	70	4
	Frequent Pedestrian Access	5	3	10	0	5	5	10	38	2
	Child Protection Area (Children jump out)	10	10	10	10	10	10	10	70	4
Road Environment	Elderly Protection Area (Slow pedestrian Crossing)	10	10	5	10	10	10	10	65	4
	Road Lane Occlusion	5	5	10	0	5	5	5	35	2
	Road Obstacle	10	5	5	0	10	10	10	50	3
	Road Damage	10	5	5	0	10	10	10	50	3
	Speed bump	5	5	10	0	5	5	5	35	2

Table 1: Road situation risk analysis by object/event and system implementation difficulty analysis

Utilizing the analysis results on the degree of risk of road conditions by object/event and the difficulty of implementing the self-driving system, the self-driving shuttle service route being demonstrated in Jochiwon-eup, Sejong-si, is classified as level 1 (green), level 2 (yellow), level 3 (orange) and level 4 (red) as shown in Figure 3 along with colors.

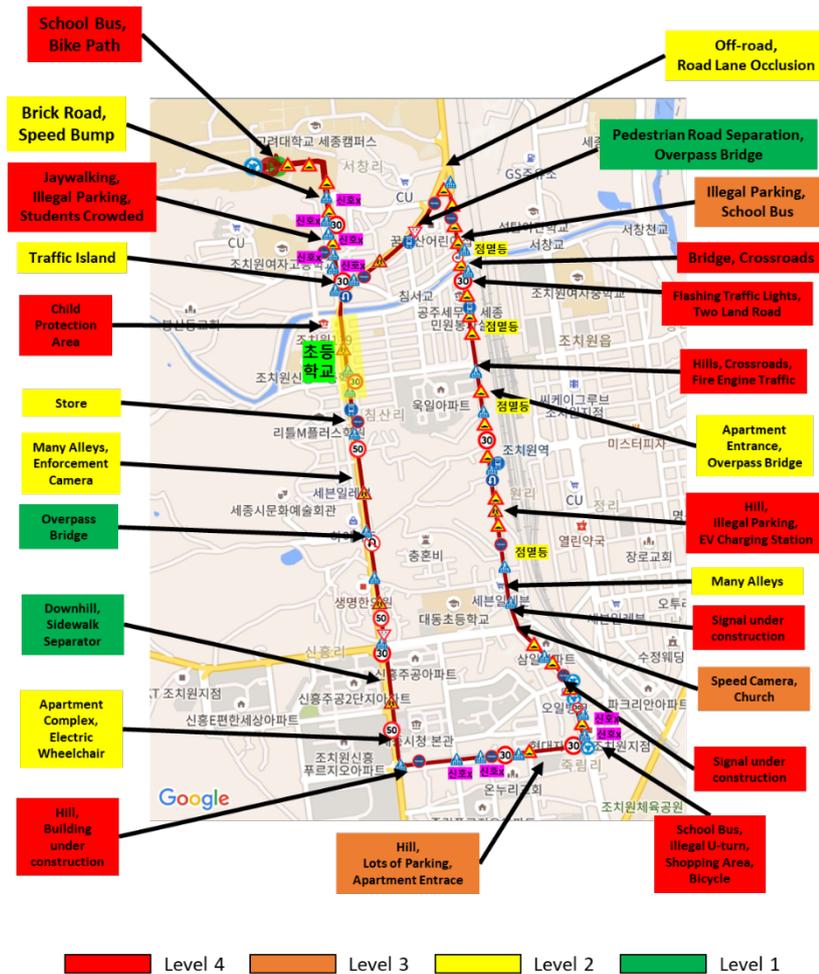


Fig.3: A map showing the level of autonomous driving

### Conclusion

In this paper, a safe operation ODD scenario was developed based on the threat factors of unexpected situations on real roads to develop and demonstrate safe on-demand autonomous driving shuttle service technology in the shaded area of public transportation in Sejong City.

Based on the derived ODD scenarios, we analyzed the risk of road conditions and the difficulty of autonomous driving system implementation by object/event to establish criteria for classifying shaded public transportation areas in Sejong City and segmenting autonomous driving shuttle service routes using the classification criteria. We developed a model to display the level of operation for each level.

### Acknowledgment

This research was supported by the “Regional Innovation Strategy (RIS)” through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (MOE) (2021RIS-004).

### References

[1] SAE International, Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles, SAE International, USA, J3016, Jun. 2018.

[2] Jinha Kim and Seokcheol Ki, A study on ODD and OEDR guidelines based on self-driving demonstration cases in Sejong City, *Journal of the Korean Society of Automotive Engineers*, 28(10), p. 659—668, 2020.

[3] Staplin, L., Mastromatto, T., Lococo, K. H., Kenneth W. Gish, K. W., and Brooks, J. O., A Framework for Automated Driving System Testable cases and Scenarios, NHTSA, Washington, DC, USA, Rep. DOT HS 812 623, Sep. 2018.

[4] Dae-Kug Lee, Eun-Young Cho, Min-Jae Jeon and Choong-Ho Cho, Research on Sejong City Public Transportation Shaded Area Classification Based on ODD Scenario and Development of Driving-level application model, *KICS Fall Conference 2022*, p. 724—725, 2022.



## A Survey of Utilizations, Challenges and Thinking of Neural Networks in Solving Path Planning Problem

Yuan Huang<sup>1</sup>, Cheng-Tien Tsao<sup>2,3</sup>, and Hee-Hyol Lee<sup>3</sup>

Graduate School of Information, Production and Systems,  
Waseda University, Japan

[gakki@toki.waseda.jp](mailto:gakki@toki.waseda.jp), [ct.tsao@akane.waseda.jp](mailto:ct.tsao@akane.waseda.jp), [hlee@waseda.jp](mailto:hlee@waseda.jp)

### Abstract

In this work, we introduce current utilizations of neural networks in the path planning field, including a segmentation task, a classification task, and a regression task. The main aspect of the utilizations is to predict a promising region that contains an optimal solution, which is regarded as a segmentation task. We further extend the utilization into classification and regression tasks, and then describe the challenge and our effort to enhance the performance. Furthermore, we propose an example of the utilizations by combining the segmentation task with the classification task, namely region and guideline prediction (RGP) model, which produces regions and guidelines for accomplishing non-uniform sampling, thereby improving the sampling efficiency, and reducing the calculation time of a sampling-based path planner, RGP-RRT\*. The performance is validated by simulations in respect of the accuracy of the prediction results, and calculation time in path planning.

**Topics:** robotics, machine learning, intelligent control

### Introduction and background

Path planning problem focuses on finding a collision-free path to connect a starting point with a goal point. Previous research developed sampling-based algorithms to tackle the path planning problem rapidly and reliably by generating samples. The samples are employed to explore a feasible configuration space and construct a graph like probabilistic roadmap [1] or a tree structure like rapidly exploring random tree (RRT) [2]. However, low efficiency of sampling impedes the development of the sampling-based algorithms, especially in complex environments (as shown in Fig.1). On the one side, redundant samples and exploration in useless areas lead to a long calculation with expensive solution costs. On the other side, samples without heuristic information struggle to escape a local trap, such as narrow passages.

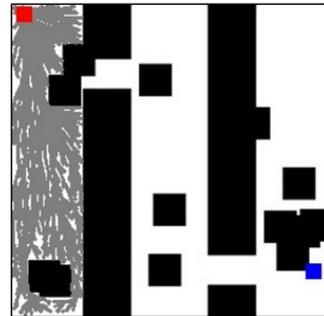


Fig.1:A sampling-based planner, RRT. in complex environments.

Recently, a new perspective to enhance the sampling efficiency is introduced by utilizations of neural networks. As shown in Fig.2, neural networks predict a promising region as a key to improve sampling efficiency. Specifically, the region is defined as an area that contains optimal solutions in the feasible configuration space. Subsequently, sampling-based path planners employ the region domain as heuristics to realize non-uniform sampling, guaranteeing samples are in close proximity to the optimal path. The prediction process is regarded as a segmentation task, allowing convolution neural network architectures [3-9], generative adversarial networks [10,11], and transformers [12] to be integrated with the path planners. Nevertheless, the previous research only consider the segmentation task to extract the region information instead of other spatial and temporal information like the approximate length and the accurate position of the optimal path, which are rarely discussed before.

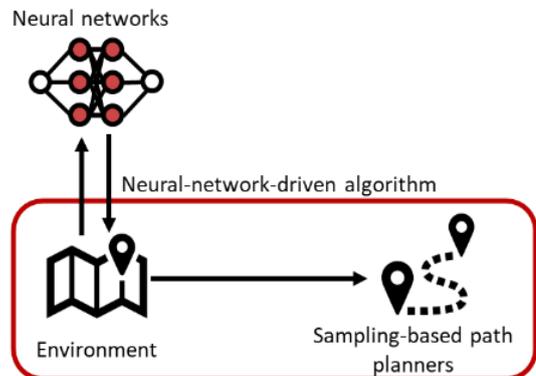


Fig.2: Flow of neural-network-driven algorithms.

In this work, we extend and divide the utilization of neural networks into four tasks, and accordingly discuss the current challenges to further improve the performance of the neural-network-driven algorithms. Subsequently, we introduce the effort and extension we have made to overcome the challenges. Besides, we propose a CNN-based model as an example to illustrate the procedure of neural-network-driven algorithms. Simulations are conducted to verify the efficiency of the proposed method compared with other state-of-the-art algorithms. Finally, we present our thinking on the utilizations of neural networks in path planning.

### Utilizations and challenges in path planning

we extend the utilization of neural networks. The utilization can be subdivided into three tasks referring to research targets.

- 1) Segmentation task: As an essential application of neural networks, the segmentation task aims at extracting a promising region to construct a region sampling domain [4-11]. Sampling-based path planners greatly benefit from the region prediction results to implement a non-sampling strategy and enhance the sampling efficiency. However, the predicted region can be inaccurate and even disconnected, which mislead the exploration.
- 2) Classification task: This task is designed to generate a guideline from the segmentation results, which can be regarded as the central line of the predicted region. We first introduce this guideline in [13]. It is more accurate to represent the position of the optimal solution relative to the region. Thus, the samples can explore the configuration space near the optimal solution efficiently. It should be noted that the samples generated in the guideline domain are far from each other since the guideline is narrow. This observation indicates that the attempts to connect each sample are time-consuming due to the blocking between samples far apart.
- 3) Regression task: This task is first proposed in [14], and we employ a regression model to estimate the path length between the starting point and the goal point. In the previous research, we utilized the regression result to establish a graph for a multi-goal path planning problem, and the visiting order is computed based on the complete graph.

Table 1 Tasks for neural networks in path planning

Neural network task	Previous research	Internal challenge	External challenge
Segmentation	[4-11]	accuracy, connectivity	Multi-task learning
Classification	[13]	accuracy	
Regression	[14]	accuracy	

Unlike previous neural-network-driven algorithms with a single task, we extend the neural-network-driven algorithm involving two or more neural network tasks for path planning. It allows the neural network to extract rich information simultaneously, which is beneficial to enhancing the efficiency of path planners referring to the research targets. Consequently, an external challenge of multi-task learning should be considered. Due to the individual losses with different gradients and scales, it is significant to balance the individual loss from different tasks in joint learning.

### Effort to overcome the challenges

To overcome the challenges, we proposed various methods in [13,14]. We outline our major contributions as follows:

- 1) To enhance the accuracy and connectivity of the predicted region, we designed a UNet-like model, region prediction neural network (RPNN), involving a segmentation task as shown in Fig.3. Significantly, we devised a patch-level loss function, namely purity loss, to evaluate the significance of a pixel in maintaining the connectivity of the predicted region. Simulation results demonstrate a 3.46-5.84% improvement in accuracy, compared with the region prediction model, NEED [7], and MPT [12].

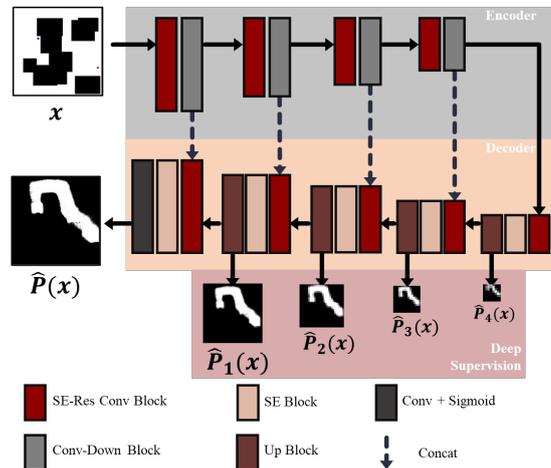


Fig.3: Illustration of the RPNN.

- 2) We implemented a multi-task learning model in [14] to tackle the multi-goal path finding problem. A regression module is integrated with a segmentation module, while a predicted region and an estimation of the length of the path are produced orderly by the proposed model, S&Reg, as displayed in Fig.4. The estimation is utilized for establishing a complete graph to compute a visiting order, while the region is employed to guide the search of a local path planner. The estimation and segmentation results are assessed by a mean squared error MSE and an intersection over union (IOU), respectively. The S&Reg achieves a 73.49 MSE score and an 81.60 IOU score.

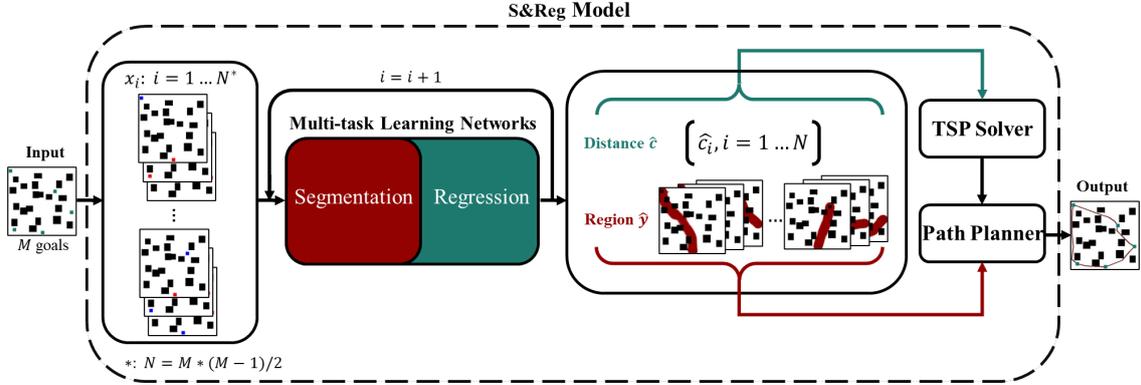


Fig.4: Illustration of the S&Reg model.

#### Example of neural-network-driven algorithm

In this section, we introduce an example by integrating the segmentation task and the classification task to realize rapid path planning. As shown in Fig.5, a region and guideline prediction model (denoted by RGP) is designed, including a region prediction sub-model and a guideline prediction sub-model. We designed a multi-scale convolution module to enlarge the receptive field, allowing the decoder to extract more features to recover the resolution accurately. Additionally, a plain U-Net [15] fuses the original map and the output of the region prediction to extract the central line as the guideline. The outputs of the RGP are employed as a region domain and a guideline domain for sampling. Subsequently, a sampling-based path planner, RGP-RRT\*, employs the sampling domains to realize a non-uniform sampler. Unlike other neural-network algorithms with a fixed bias probability, the RGP-RRT\* utilizes an adaptive bias probability to select the sampling domain in different stages of the search procedure.

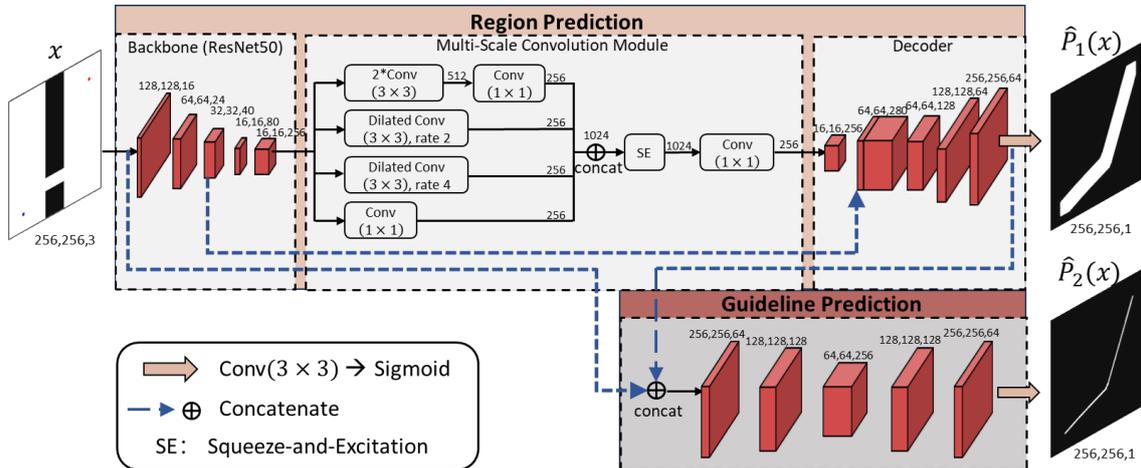


Fig.5: Illustration of the RGP model.

We conduct simulations to demonstrate the superiority of the RGP model and the RGP-RRT\* in respect of prediction accuracy and calculation time to find an optimal path. Specifically, we utilize a dataset with 16000 path planning cases from <https://github.com/RTPWDSM/PPD>, and a test dataset with 2400 scenarios including seen scenarios and unseen scenarios. The results are listed in Table 2. The RGP outperforms the other prediction neural networks (MPT [12] and NEED [7]) in region prediction. Because the MPT and NEED lack a classification sub-model to generate the guideline, we only present the results of the RGP.

To verify the performance of the RGP-RRT\* in path planning, we compare the RGP-RRT\* with the RGP-RRT\* without (w/o) AB, the NEED-RRT\* [12], the MPT-RRT\* [7], and the plain RRT\* [16] based on 50 independent tests on four scenarios (as shown in Fig.6). The prediction results are displayed with the original map for better understanding. The predicted regions and guidelines are in deep red and orange, respectively. The exploration results of sampling-based path planners are illustrated in Fig.7. We can observe that the RGP generates guidelines to connect the starting point and the goal point although the region is disconnected as shown in

scenario 3. Benefitting from the accurate region and guideline, the RGP-RRT\* achieves a 7.2-80.1% reduction in the calculation time with a 2.0-58.1% reduction in the sampler number.

Table 2 Accuracy of neural networks in region prediction and guideline prediction

	Seen Scenarios		Unseen Scenarios	
	Region	Guideline	Region	Guideline
RGP	0.9281	0.5816	0.5974	0.3582
MPT(2021)	0.8770	\	0.4975	\
NEED(2022)	0.9234	\	0.5916	\

Fig.8 shows boxplots of the calculation time in 50 independent tests on scenarios 1 to 4. The performance of all neural-network-driven algorithms greatly exceeds the plain RRT\*, resulting from the non-uniform sampler based on the outputs of the neural networks. Notably, the NEED-RRT\* shows worse performance in scenario 2 on the score of misleading searching by the inaccurate region prediction results. In scenario 3, the RGP-RRT\* outperforms the other neural-network-driven algorithms with the guideline sampling, which accounts for the fewer search areas (denoted by gray in Fig.7).

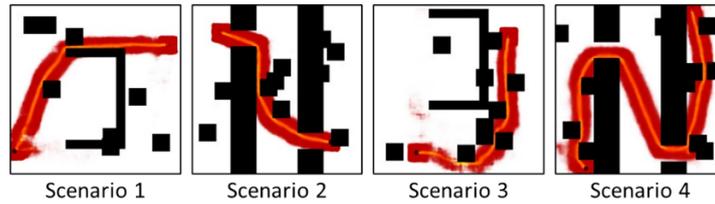


Fig.6: Illustration of the RGP model.

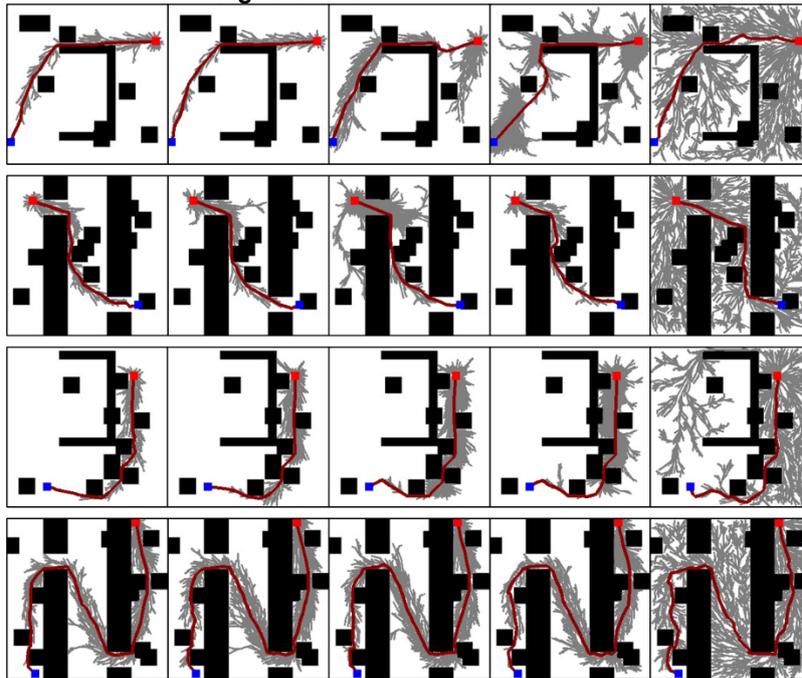


Fig.7: Illustration of the RGP model.

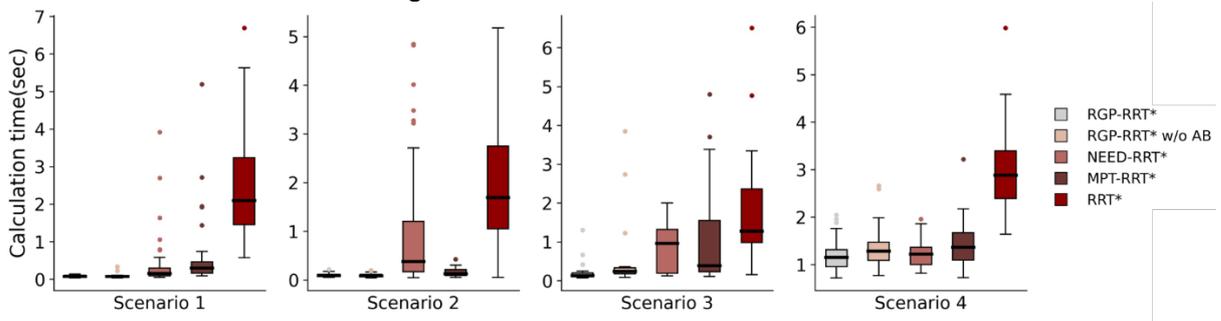


Fig.8: Illustration of the RGP model.

Conclusion and thinking

In this work, we introduced the utilizations and challenges of neural networks in path planning. Different from the mainstream utilization as a segmentation task, we categorized utilizations into a segmentation, a classification, and a regression task. We further presented the effort we had made in improving the performance of neural-networks-driven algorithms. Besides, we proposed an example utilization by combining the segmentation and classification tasks to enhance the sampling efficiency and reduce calculation time. The performance of the proposed model, RGP, and sampling-based path planner, RGP-RRT\*, were validated by simulations.

In my opinion, the neural networks can be extended with specific tasks for path planning corresponding to the research target, e.g. utilizing the regression model to predict the length of the optimal path. Moreover, balancing the difference among individual losses and designing an efficient non-uniform sampler based on the outputs of neural networks are significant to further enhance the performance of the neural-network-driven algorithms to tackle the path planning problem.

## References

- [1] Kavraki, Lydia E., et al. "Probabilistic roadmaps for path planning in high-dimensional configuration spaces." *IEEE transactions on Robotics and Automation* 12.4 (1996): 566-580.
- [2] LaValle, Steven M., and James J. Kuffner Jr. "Randomized kinodynamic planning." *The international journal of robotics research* 20.5 (2001): 378-400.
- [3] Qureshi, Ahmed Hussain, et al. "Motion planning networks: Bridging the gap between learning-based and classical motion planners." *IEEE Transactions on Robotics* 37.1 (2020): 48-66.
- [4] Ichter, Brian, James Harrison, and Marco Pavone. "Learning sampling distributions for robot motion planning." *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018.
- [5] Wang, Jiankun, et al. "Neural RRT\*: Learning-based optimal path planning." *IEEE Transactions on Automation Science and Engineering* 17.4 (2020): 1748-1758.
- [6] Liu, Jianbang, et al. "Learning-based Fast Path Planning in Complex Environments." *2021 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 2021.
- [7] Wang, Jiankun, et al. "Robot path planning via neural-network-driven prediction." *IEEE Transactions on Artificial Intelligence* 3.3 (2021): 451-460.
- [8] Ma, Han, et al. "Enhance connectivity of promising regions for sampling-based path planning." *IEEE Transactions on Automation Science and Engineering* (2022).
- [9] Wang, Jiankun, et al. "Deep neural network enhanced sampling-based path planning in 3D space." *IEEE Transactions on Automation Science and Engineering* 19.4 (2021): 3434-3443.
- [10] Zhang, Tianyi, Jiankun Wang, and Max Q-H. Meng. "Generative adversarial network based heuristics for sampling-based path planning." *IEEE/CAA Journal of Automatica Sinica* 9.1 (2021): 64-74.
- [11] Ma, Nachuan, et al. "Conditional generative adversarial networks for optimal path planning." *IEEE Transactions on Cognitive and Developmental Systems* 14.2 (2021): 662-671.
- [12] Johnson, Jacob J., et al. "Motion Planning Transformers: A Motion Planning Framework for Mobile Robots." *arXiv preprint arXiv:2106.02791* (2021).
- [13] Huang, Yuan et al. "PKE-RRT: Efficient Multi-Goal Path Finding Algorithm Driven by Multi-Task Learning Model." *arXiv preprint arXiv: 2308.07972* (2023).
- [14] Huang, Yuan et al. "S&Reg: End-to-End Learning-Based Model for Multi-Goal Path Planning Problem." (2023). *arXiv preprint arXiv: 2308.04160* (2023).
- [15] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. Springer International Publishing, 2015.
- [16] Karaman S, Frazzoli E. Sampling-based algorithms for optimal motion planning. *The International Journal of Robotics Research*. 2011;30(7):846-894. doi:10.1177/0278364911406761



## Fire/Non-fire false alarm reduction method using weighted cost function

Saeme Byun, Kim Haneul, Changhyun Lee and Hunkee Kim\*

AI system laboratory

Inha university, Republic of Korea

[saeme1030@inha.edu](mailto:saeme1030@inha.edu), [metalneul77@inha.edu](mailto:metalneul77@inha.edu), [95222304@inha.edu](mailto:95222304@inha.edu), [hunkeekim@inha.ac.kr](mailto:hunkeekim@inha.ac.kr)

### Abstract

The subject of this study is an algorithm and complex sensor system to accurately and quickly determine fire and non-fire. In the event of a fire, a false alarm may cause a large-scale fire by mistaking the fire for a non-fire, and a false alarm in a non-fire situation may cause loss of human life and property by mistaking the non-fire for a fire. Therefore, it is very important to accurately and quickly determine fire and non-fire.

One of the main reasons of false alarm of fire and non-fire is to determine a fire only by the amount of smoke and not analyze the components of the smoke. Smoke from materials such as plastic, wood, and clothing in the home is likely to be a fire, while cigarette smoke and cooking smoke are likely to be non-fire. Based on this idea, the characteristics of combustible combustion gas were collected and analyzed through a complex sensor that can analyze the composition of combustible combustion gas, and an algorithm for determining fire and non-fire was developed using a weighted cost function according to the characteristics of the components.

The new fire/non-fire decision system developed through this study is expected to quickly determine fire and non-fire situation based on high reliability, extinguish the fire early before it spreads, and reduce social costs by reducing non-fire false alarms.

**Topics:** IoT, machine learning, sensor system

### Introduction and background

The problem of fire/non-fire false alarm of fire detectors is increasing every year and is becoming a serious social issue. These false alarms of fire detectors can cause non-fire alarms, resulting in unnecessary wastage of manpower and money, and loss of reliability. This loss of reliability can lead to a state of safety numbness that prevents early extinguishment of a real fire, resulting in property and life loss [1,2]. Typical non-fire alarm triggers in the home include cooking gas emissions(e.g., pork belly, fish, etc.) and tobacco-related gases, and there is currently a lack of research on how to accurately determine fire/non-fire and early fire suppression [3].

To solve these problems, we propose an algorithm that can determine fire/non-fire by analyzing the components of gas, unlike gas detectors that determine fire only by the amount of gas. To develop an algorithm that can accurately determine fire/non-fire in an indoor environment and extinguish it early, we set the fire situation as gas from wood, clothing, plastic, etc. and the non-fire situation as gas from cooking (pork belly, fish). Gas composition data of CO<sub>2</sub>, CO, NO<sub>2</sub>, and NH<sub>3</sub> generated during combustion by substance for 200 seconds, which is the slow-growth fire time, is extracted, and then combustion characteristics by fire/non-fire situation are identified based on the data of combustion gas changes by substance to obtain a cost function that is weighted differently according to the characteristics of the substance.

The integration of gas component analysis into existing fire and non-fire alarm systems could be a promising solution to mitigate false alarms and enhance overall system accuracy. Notably, existing research predominantly relies on image processing algorithms for fire detection, including methods like fire color analytics through CNN-based image processing, analysis of optical flow patterns for video smoke recognition, texture analysis of smoke, and fire detection based on flame color dispersion and similarity in consecutive frames. In contrast, our approach introduces gas component analysis as an additional and complementary technique. We believe that by combining these diverse methods in an ensemble approach, we can significantly improve the effectiveness of fire and non-fire alarm systems, ultimately leading to more precise and reliable fire detection and early suppression capabilities.

**Goals, proposed method, novelty**

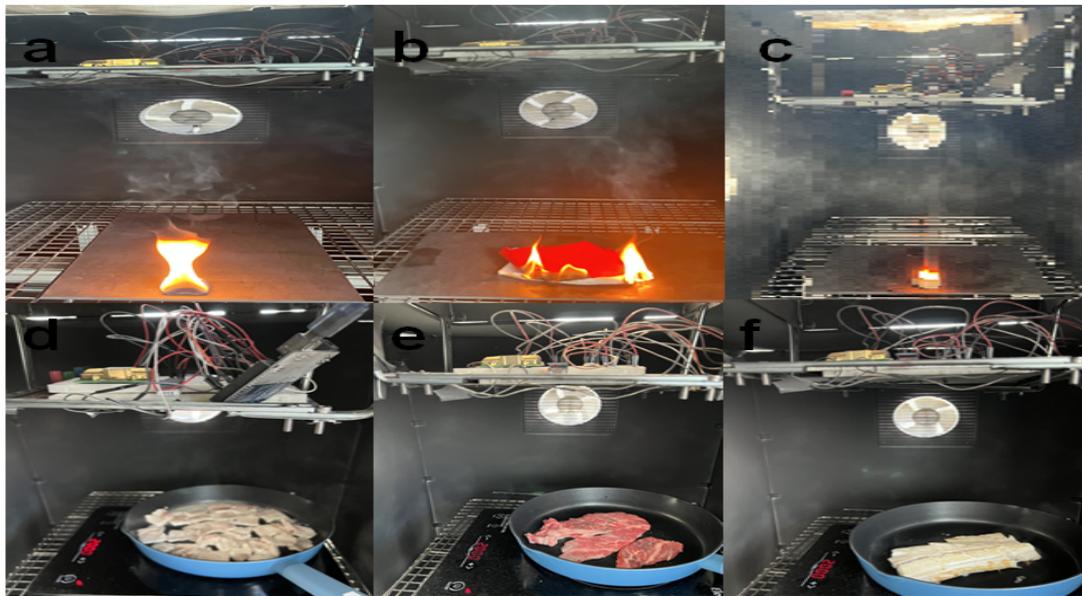
In our study, we aimed to investigate combustion gases produced during fires in indoor environments, with a particular focus on slow-growth fires that are commonly observed in homes. To accomplish this, we conducted experiments utilizing a semiconductor CO<sub>2</sub> sensor known for its high gas detection sensitivity and an electrochemical CO, NO<sub>2</sub>, NH<sub>3</sub> multi-channel gas sensor.

The fire conditions in our experiments simulated typical household scenarios involving materials such as wood, clothing, and plastic. In contrast, we used 'meat gas' to represent non-fire conditions. For consistency and to ensure the reproducibility of our results, we maintained the same mass for both fire and non-fire items. Therefore, we employed 5 grams for the fire items and 200 grams, which corresponds to a serving size, for items like 'meat gas'.

To capture relevant data, we set the flue gas measurement time to 200 seconds, a standard duration for early detection of low-growth fires. During the experiments, we ignited the fire items using a torch, while the non-fire item ('meat gas') was ignited using a highlight. Additionally, we took measures to block any interference from gases other than the flue gas, as illustrated in Fig 1.

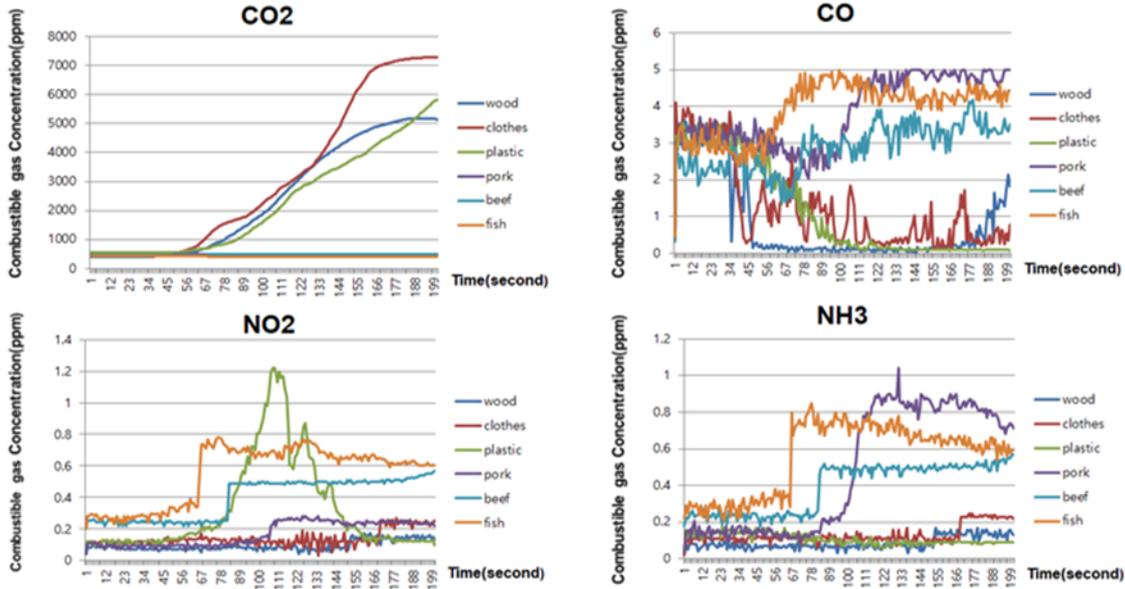
The item	Weight	Key Ingredients
Wood	5g	Solid wood : 100%
Clothing	5g	Cotton : 100%
Plastic	5g	ABS : 100%
Pork Belly	200g	Pork : 100%
Beef Sirloin	200g	Beef : 100%
Mackerel	200g	Fish : 100%

**Table 1** : Experiment Material Properties



**Fig 1** : Combustion Gas Measurement Experimental Process

**Experiments, results, analysis of the results**



**Fig 2 :** Graph of Combustible gas concentration by substance

Based on the experimental data results, we found that the combustion gases that are effective in determining fire and non-fire are CO<sub>2</sub>, CO. In the case of fire, CO<sub>2</sub> increases rapidly, CO has a negative slope when gas is generated. For non-fires, the CO<sub>2</sub> concentration increases slightly below 1000 ppm, with a positive slope for CO. This analysis of combustible composition can be used as an effective way to apply algorithms for determining fire and non-fire.

**Cost Function and Algorithm Development**

Based on experimental data, we conducted an analysis of gas composition data for various combustible materials and identified CO<sub>2</sub>, CO, and temperature as crucial variables for distinguishing between non-fire and fire conditions. Consequently as shown as Equ 1 and Equ 2, we developed two cost functions: one incorporating CO<sub>2</sub> and CO, denoted as i and j, with varying weights assigned based on component characteristics, and another cost function solely based on temperature.

$$S_1(x) = (i * w_1) - (j * w_2)$$

**Equ 1 :** The cost function related to CO<sub>2</sub> and CO

$$S_2(x) = \frac{1}{1 + e^{-(T-72)}}$$

**Equ 2 :** The cost function related to temperature

First, for Equation 1, the flue gas concentration values of CO<sub>2</sub>, CO obtained from the experimental data are normalized by converting them to values between 0 and 1, respectively, and then the weights are set for CO<sub>2</sub>, CO which play an important role in distinguishing between fire and non-fire. CO<sub>2</sub> is assigned a high weight because it has a strong influence on fire judgment, CO is assigned a low weight because its negative slope helps distinguish fire. Therefore, in this study, a weight of 0.7 was assigned to CO<sub>2</sub>, and a weight of 0.3 was assigned to CO. By multiplying the normalized value of each combustion gas, denoted as i and j, by their respective weights, a weighted value is calculated. Fire and non-fire conditions are determined based on the resulting cost function value

Following Equation 1, Equation 2 is designed as a temperature-related cost function. It is configured to have values between 0 and 1 by setting the indoor sprinkler activation temperature to 72 degrees Fahrenheit.

And so, if either of the two functions yields a value exceeding 0.5, it is considered a fire event. It's important to note that the weights and the threshold temperature are arbitrarily set and may require optimization in accordance with specific situations in the future.

Using these two cost function, you can develop an algorithm that takes gas composition data as input to determine whether there is a fire. The algorithm enables accurate fire/non-fire judgment and improves the accuracy of early suppression and alarm generation in actual fire situations.

## **Conclusion**

To solve the problem of non-fire alarms caused by malfunctioning fire detectors, this study proposes a new algorithm for determining fire/non-fire by analyzing the components of gas. The algorithm is expected to improve the accuracy of fire detection and contribute to preventing the cost and loss of reliability due to malfunction.

In order to develop an algorithm that can accurately determine fire/non-fire in an indoor environment and extinguish it early, we conducted an experiment to analyze the composition of smoke and found that the characteristics and temperature of CO<sub>2</sub> and CO combustion gases can reliably distinguish between fire and non-fire. Based on this, a weighted cost function for CO<sub>2</sub> and CO and a cost function related to temperature can be constructed to develop an algorithm that determines whether a fire is present based on smoke composition and temperature.

In the future, we will conduct experiments to extract combustion gas data other than CO<sub>2</sub>, CO, NH<sub>3</sub>, and NO<sub>2</sub> to see if there are other gas components that can be applied to the fire/non-fire discrimination algorithm, and we plan to construct a comprehensive cost function that incorporates multiple gas components.

The proposed algorithm is expected to serve as a useful tool to improve the accuracy of early suppression and alarm generation in real fire situations, thereby reducing the occurrence of unnecessary non-fire alarms, improving the reliability of fire detection, and minimizing the damage to property and people.

## **References**

- [1] Byung Keun Seo and Sang Gu Nam, Study of the Improvement of False Fire Alarms in Analog Photo electric Type Smoke Detector, Disaster of Science ,Graduate School, University of Seoul \*R&D, Tyco Korea Co. Ltd, 2016.
- [2] Euy-Hong Hwang, Sung-Eun Lee and Don-Mook Choi, Statistics and Management Systems of Unwanted Domestic and Foreign Fire Alarms, Graduate Student, Dept. of Equipment System·Fire Protection Engineering, Gachon Univ, Professor, Dept. of Fire and Disaster Prevention Engineering, Hoseo Univ, Professor, Dept. of Equip-ment System·Fire Protection Engineering, Gachon Univ, 2020.
- [3] Sung-Ho Hong, A Experimental Study on the Heat Release Rate to activate Fire Detection Sensor, The Transac-tions of the Korean Institute of Electrical Engineers, 2012.

## Optimum driving algorithm to correspond with various outdoor environment

Seung Min Kim, Yeong Ho Ha, Hyun Seo Jang and Young Eun Song

The College of Science & Technology

Korea university, Korea

songarak27@korea.ac.kr, haho9658@korea.ac.kr, kelvin926@korea.ac.kr, tdsong@korea.ac.kr

### Abstract

Autonomous driving for mobile robot is basically rely on [1] Highly Definition map (HD map), But mobile robot's driving is quite different from car. This map is not enough to overcome variable situations mobile robot face. In this paper we review how to optimize process of driving on slope in 3 steps entry process, driving on slope and exit process. This extended ability makes it possible mobile robot to drive on more widespread area.

**Topics:** Computational Intelligence, Planning and Scheduling, Intelligent and process control system, robotics, Wireless communication.

### Introduction and background

- Mobile robot has been developed for years and we can see some mobile robots inside of the buildings but we want to make mobile robots driving on both inside and outside. For the experiment [2] we use Ackermann based mobile robot (Fig. 1) which can change wheel's direction separately.



Fig.1: Ackermann based mobile robot

- We first divide the process into 3 steps entry process, driving on the slope and exit process. Figure 2 shows the 3 steps. We don't include getting out from the downhill in the process. Because it is totally same to the first step. For better driving, if we don't slow down the robot when it starts to drive on the slope, mobile robot's wheel could be damaged a lot and in worst cast robot could rolled over. So it need to divide the process and optimize each step.

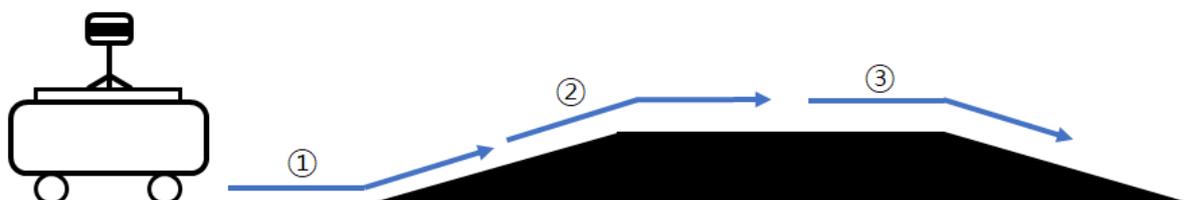


Fig.2: Mobile robot's driving on slope process

**Goals, proposed method, novelty**

To accomplish our goal we use the Ackermann based mobile robot which has big flexible mobility on driving environment. While the mobile robot driving on the slope based on information the path planning algorithm provide it will store some informations related with IMU, Motor speed and software errors. After that we analyze the data which component is crucial.

Ackermann can steer it's four tires angle so we can change Center Of Mass(COM) which means it has flexibility of mobility But it is very hard to how steer the each tire based on given data from controller. Fig3) So we make the Flowchart of mobile robot's control system.

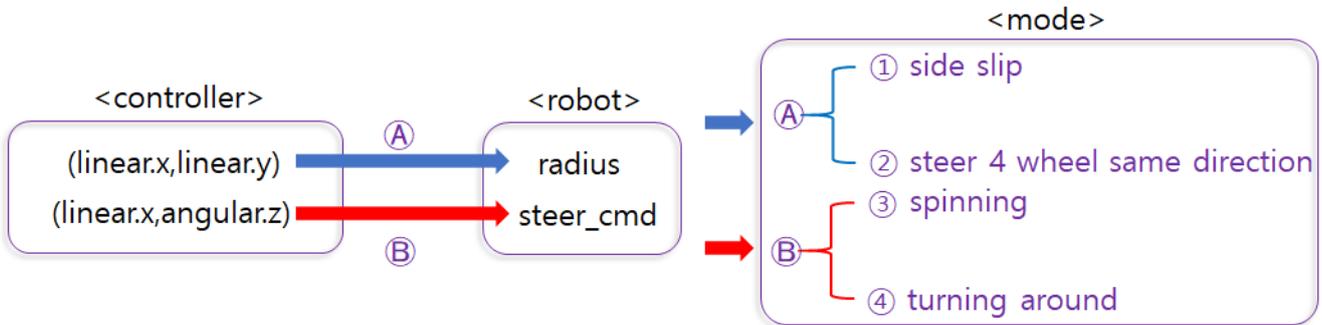


Fig.3: Flowchart of mobile robot's control system

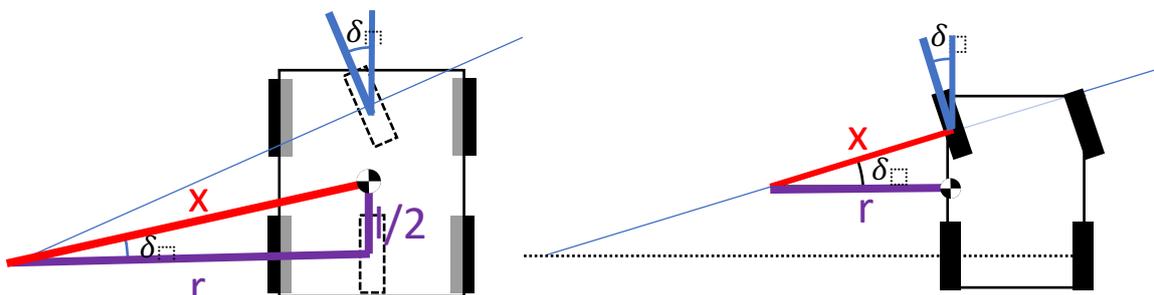
Table1: Classification of mobile robot's each mode

Mode	Side slip	steer 4 wheel same direction	spinning	turning around
Controller	(linear.x,linear.y)	(linear.x,linear.y)	(linear.x,angular.z)	(linear.x,angular.z)
Calculation base	steer_cmd	steer_cmd	radius	radius
Wheel direction				

As shown in table1 each mode has properties, so we can use 4 modes at different situation

**Experiments, results, analysis of the results**

- In this experiment we first did kinematics of Ackermann not only for the case COM is middle of robot also for the case COM is middle of each side. Fig4 represents example when COM is locate at different location.



**Fig.4:** when COM is located at center of mobile robot(left) when COM is located at middle of left side(right)

Where  $r$ =linear/angular,  $l$ =wheelbase,  $x$ =turning radius,  $\delta$ =steer angle. In this experiment we hope to optimize Ackerman based mobile robot's control process to correspond with gradient in different condition. This experiment is meaningful as initial step for future works. In the future works we will include various situations like obstacles, road condition and tremor.

#### References

- [1] Bin Yang, Ming Liang, Raquel Urtasun  
Proceedings of The 2nd Conference on Robot Learning, PMLR 87:146-155., p. 1-5, 2018.
- [2] W. C. Mitchell, A. Staniforth, and I. Scott, "Analysis of Ackermann steering geometry," SAE Tech. Paper 2006-01-3638., p. 1-3, 2006.



## Theoretical Back Grounds for Efficiency of Robust Design in Multi-Level Optimization

**Masao ARAKAWA**

Department of Information, Production and System

Waseda University, JAPAN

[arakawa.masao@waseda.jp](mailto:arakawa.masao@waseda.jp)

### Abstract

There are a lot of ways to obtain good results in large scale optimization, which have been studied in multidisciplinary optimization. Among them, multi-level optimization is one of the most efficient methods. We need to go through decomposition of the main problem, and make levels of small-scale optimizations, and finally recompose them to get final solutions. If and only if we can avoid feedbacks of decisions among levels, it will be so efficient to be real use. It is said that if we can set efficient side constraints to the following levels, we can limit these feedbacks. Theoretically, this idea seems to work. But there are no effective ways to give side constraints. We have applied robust design that it can control deterioration of behavior functions with handling width of design variables. It means we can treat design variables as interval variables. We have shown effectiveness of the proposed method through numerical example. In this paper, we are going to show its back grounds why it works so well.

**Topics:** Machine Design, Design Engineering, Multi-level Design, Robust Design

### Introduction and background

Needs for using optimization in large scale problems become urgent these days. Besides, their problems become more reality in industry, problems need to treat complicated nonlinearity especially in constraints. Although, there are great developments of PC powers, and also with methodological algorithms in optimization, it still does not catch up with the size problem of real use. To make problem in effective size, there are multi-level optimization which decompose the problem in small size, and make loops to recompose the problem. As optimization size becomes small, we can carry out each optimization quite effectively. However, we need to solve compromise problems between decisions in levels. Otherwise, it is difficult to converge forever. To avoid the situations, there are recommended to set side constraints to the next levels to limit range of decisions to keep your own decisions. However, there is no significant method to give side constraints, therefore these recommendations are quite difficult to be used. On the other hand, Robust Design has been studied for a few decades. It can control deterioration of both objective functions and constraints with design variables and their perturbations. Therefore we can treat interval variables and we can determine side constraints for each design variables. Unesaki[1] had shown effectiveness of the proposed method through numerical example. We actually do not need to carry out any feedbacks at all. In this paper, theoretical back grounds why it works so well are going to be discussed.

### Multi-level Optimization in MDO

One of the most successful study in Multidisciplinary optimization (MDO) had done by Kroo[2], and it theoretically leads how to decompose main problems and how to make efficient levels of sub-problems. Figure 1 shows just decomposition and good decomposition. In this figures, upper right connections mean there are some design variables to be succeeded to the levels that has connections. Lower left connections mean that there are mutual design variables between both levels. Which are called feedbacks that we would like to avoid. As there are feedbacks of design variables which are involved in different levels and if they have compromised decisions, we need to carry out sub-problems optimization again and again when levels were decomposed which has larger loops. They also recommended to set side-constraints to cut these loops. If and only if we can give effective side-constraints decomposition might be (C) in Fig. 1. Thus, we can reduce a number of small optimizations to get results finally. In order to make ideal decomposition, we need to give side-constraints efficiently, so that we do not need to go through feedbacks. For that purpose, we need to treat design variables which are included in different levels as interval variables. As for behavior variables, we need to determine deteriorations of them for those intervals.

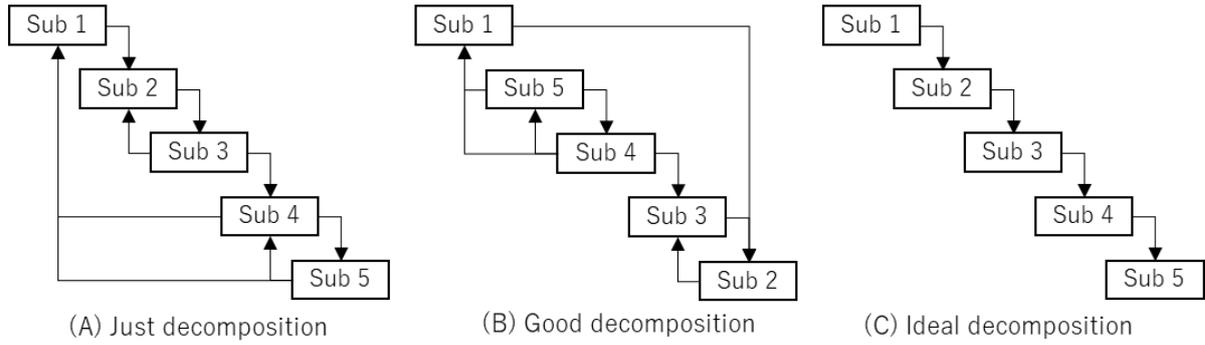


Figure 1 MDO decomposition of the problem

### Robust design with fuzzy numbers

Robust design is one of the major topics in the field of optimization for decades. It treats perturbation of behavior variables under a sort of uncertainty in design variables and its parameters. Most of them treat these uncertainty as probability and which are determined priori to the design. Arakawa[3] proposed to use fuzzy numbers as design variables, and make it possible to control width of design variables. This idea is preferable to the problems that we are going to treat. We recommended to use optimization of worst case for robust design[4] and made formulation according to ideas of fuzzy numbers, we give formulations as follows.

$$\min_{\mathbf{x}} \max_{\mathbf{p}} f(\mathbf{x}, \mathbf{p}) \quad (1)$$

Subject to

$$\min_{\mathbf{p}} g_k(\mathbf{x}, \mathbf{p}) - g_k^a \geq 0 \quad (2)$$

Where  $\mathbf{x}$  is vector of design variables,  $\mathbf{p}$  is vector of parameters and it includes fuzzy parameters,  $f$  is objective function and  $g_k$  is constraints. Equation (1) means worst case can be determined by maximization of parameters  $\mathbf{p}$ , and we would like to minimize it with design variables  $\mathbf{x}$ . Equation (2) means that we can keep constraints even with worst case of constraints. Clearly, it is composed by two loop of optimization. Which is difficult to be solved. Therefore we are going to introduce release variable  $z$  to avoid this problem and formulate it as multi-objective optimization as follows.

$$\min f(\mathbf{x}, \mathbf{p}) \quad (3)$$

$$\min z \quad (4)$$

$$\max \mathbf{p} \quad (5)$$

Subject to

$$f(\mathbf{x}, \mathbf{p}) \leq (1 + z)f(\mathbf{x}) \quad (6)$$

$$g_k(\mathbf{x}, \mathbf{p}) - g_k^b \geq 0 \quad (7)$$

It seems a little strange, but Eq. (4) and (5) always have conflict so that it will be equivalent with Eq.(1) but does not need to have inner and outer loop, but trade-off analysis. With relax variable  $z$ , we can solve robust design within allowable numbers of function calls.

### Golinski's speed reducer

There are a lot of bench mark problems that AIAA provides. Among them Golinski's speed reducer has mid-size problems with mid-size constraints (7 design variables and 11 constraints). If the number of design variables are less equal 4, no matter how complex the problems are, it is relatively easy to solve. Formulation of Golinski's speed reducer are followings.[5]

Minimize

$$f = C_{f1}x_1x_2^2(C_{f2}x_3^2 + C_{f3}x_3 - C_{f4}) - C_{f5}(x_6^2 + x_7^2)x_1 + C_{f6}(x_6^3 + x_7^3) + C_{f1}(x_4x_6^2 + x_5x_7^2) \quad (8)$$

Subject to

$$g_1 = \frac{C_{g1}}{x_1x_2^2x_3} \leq 1.0, g_2 = \frac{C_{g2}}{x_1x_2^2x_3^2} \leq 1.0, g_3 = \frac{C_{g3}x_4^3}{x_2x_3x_6^4} \leq 1.0, g_4 = \frac{C_{g4}x_5^3}{x_2x_3x_7^4} \leq 1.0$$

$$g_5 = \sqrt{\frac{C_{A12}^2 x_4^2}{x_2^2 x_3^2} + C_{A1}} \leq 1.0, g_6 = \sqrt{\frac{C_{A12}^2 x_5^2}{x_2^2 x_3^2} + C_{A2}} \leq 1.0, g_7 = \frac{x_2 x_3}{C_{g7}} \leq 1.0, g_8 = \frac{C_{g8} x_2}{x_1} \leq 1.0$$

$$g_9 = \frac{x_1}{C_{g9} x_2} \leq 1.0, g_{10} = \frac{C_{g10} x_6 + C_{g245}}{x_4} \leq 1.0, g_{11} = \frac{C_{g11} x_7 + C_{g245}}{x_5} \leq 1.0$$

$$2.6 \leq x_1 \leq 3.6, 0.7 \leq x_2 \leq 0.8, 17 \leq x_3 \leq 28, 7.3 \leq x_4 \leq 8.3, 7.3 \leq x_5 \leq 8.3,$$

$$2.9 \leq x_6 \leq 3.9, 5.0 \leq x_7 \leq 5.50$$

Where

$$C_{f1} = 0.7854, C_{f2} = 3.333, C_{f3} = 14.9334, C_{f4} = 43.0934, C_{f5} = 1.5079, C_{f6} = 7.477$$

$$C_{g1} = 27.0, C_{g2} = 397.5, C_{g3} = 1.93, C_{g4} = 1.93, C_{g5} = 1100, C_{g6} = 850, C_{g7} = 40.0, C_{g8} = 5.0, C_{g9} = 12.0$$

$$C_{g10} = 1.5, C_{g11} = 1.1, C_{A12} = 745.0, C_{A1} = 1.69 \times 10^7, C_B = 0.1, C_{A2} = 1.575 \times 10^8, C_{245} = 1.9$$

Table 1 shows attempts from a number of researchers. There might be a lot more to be listed, but I picked several of them to show its difficulties. Some of them, I do not have precise data. Most likely, Lin and Agent PSO finally reached to global solutions.

Table 1 Comparison of the results

Methods	Type	x1	x2	x3	x4	x5	x6	x7	f	Error in Const
Ku et al	Metaheuristics	3.600000000000	0.700000000000	17.000000000000	7.300000000000	7.800000000000	3.400000000000	5.000000000000	2876.219475000000	<100
Akhtar et al	Metaheuristics	3.506122000000	0.700000600000	17.000000000000	7.549126000000	7.859330000000	3.365576000000	5.289773000000	3008.197440000000	<10 <sup>-6</sup>
Rao and Xiong	Metaheuristics	3.500000000000	0.700000000000	17.000000000000	7.300000000000	7.800000000000	3.360000000000	5.290000000000	3000.959715000000	<10 <sup>-6</sup>
Gagnina et al	Metaheuristics	3.500000000000	0.700000000000	17.000000000000	7.300000000000	7.800000000000	3.350214000000	5.286683000000	2996.347849000000	<10 <sup>-6</sup>
Jaberipour and Khorram	Metaheuristics	3.500000000000	0.700000000000	17.000000000000	7.300000000000	7.715332338339	3.350215092568	5.286664035455	2994.477531000000	<10 <sup>-6</sup>
Li and Papalambros	Metaheuristics	3.500000000000	0.700000000000	17.000000000000	7.299999000000	7.715317000000	3.350541000000	5.286654000000	2994.553869000000	<10 <sup>-6</sup>
Tosserams et al.	Deterministic	3.500000000000	0.700000000000	17.000000000000	7.300000000000	7.720000000000	3.350000000000	5.290000000000	2996.645783000000	<10 <sup>-3</sup>
Zummurd	Deterministic	3.500000000000	0.700000000000	17.000000000000	7.300000000000	7.715319900000	3.350214700000	5.286654500000	2994.355026100000	0
Lin et. Al	Deterministic	3.500000000000	0.700000000000	17.000000000000	7.300000000000	7.715319000000	3.350282000000	5.286654000000	2994.471921000000	<10 <sup>-6</sup>
Agent PSO	Metaheuristics	3.500000000000	0.700000000000	17.000000000000	7.300000000000	7.715319911478	3.350214666096	5.286654464980	2994.355026112850	0

MDO test suites used to prepared multi-level optimization of this problem as in figure 2. And its formulations are followings.

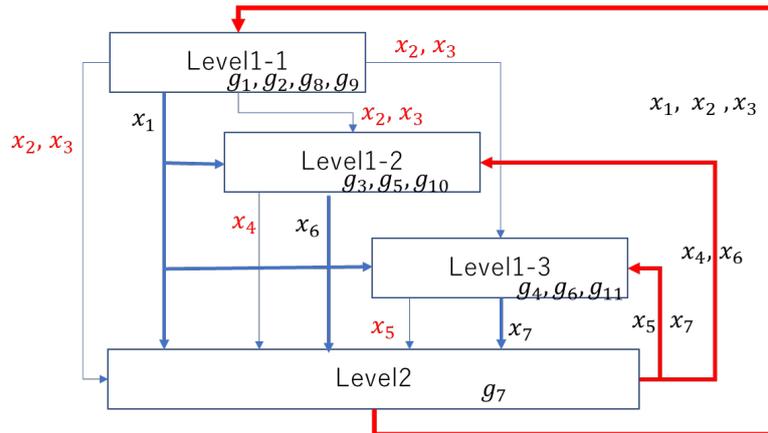


Figure 2 Structure of multi-level optimization of Golinski's Speed Reducer

In this figure, black variables in left bottom area show their decisions are fixed and to be used in the following levels. Red variables need to be feedback from the following levels. Therefore, if there are compromise between them, we have to make decision all over again.

Level1-1: Design variables are  $x_1, x_2, x_3$

$$\text{minimize } f_1 = C_{f1} x_1 x_2^2 (C_{f2} x_3^2 + C_{f3} x_3 - C_{f4}) \quad (9)$$

Subject to  $g_1, g_2, g_8$  and  $g_9$

Optimum results are  $x_1=3.5, x_2=0.7, x_3=17$  and  $f_1=1581.46$ .

Level1-2: Design variables are  $x_2, x_3, x_4, x_6$  and use results of  $x_1$  in Level1-1

$$\text{minimize } f_2 = -C_{f5} x_1^* x_2^2 + C_{f6} x_6^3 + C_{f1} x_4 x_6^2 \quad (10)$$

Subject to  $g_3, g_5$  and  $g_{10}$

Optimum results in all independent case are  $x_1=3.6, x_2=0.8, x_3=28, x_4=7.3, x_6=3.345304$  and  $f_2=283.333$ .

Level1-3: Design variables are  $x_2, x_3, x_5, x_7$  and use results of  $x_1$  in Level1-1

$$\text{minimize } f_3 = -C_{f5} x_1^* x_2^2 + C_{f6} x_7^3 + C_{f1} x_5 x_7^2 \quad (11)$$

Subject to  $g_4, g_6$  and  $g_{11}$

Optimum results in all independent case are  $x_1=3.6, x_2=0.8, x_3=28, x_5=7.3, x_7=5.28568$  and  $f_3=1112.68$ .

Level2: Design variables are x2, x3, x4, x5 and use results of x1 in Level1-1, x6 in Level1-2 and x7 in Level1-3

$$\text{minimize } f_4 = C_{f1}x_1^*(C_{f2}x_3^2 + C_{f3}x_3 - C_{f4}) - C_{f5}x_1^*(x_6^{*2} + x_7^{*2}) + C_{f6}(x_6^{*3} + x_7^{*3}) + C_{f7}(x_4x_6^{*2} + x_5x_7^{*2}) \quad (12)$$

Subject to g7

Optimization of all independent case are non-sense in this level. Results are going to be discussed more detail later. But you can see there are conflicts in x1, x2 and x3.

**Proposed Method**

Problems in MDO lie in conflicts in design variables in each level. From figure 2, we can understand we should use decision of x1 in Level1-1. Therefore, x2 and x3 might be major conflicts for all decisions. Same problems should exist in x4 and x5. Thus, we would like to set upper and lower bound for these design variables. Meaning we would like to treat them as interval design variables.

Let them set as

$$x_i = [x_i - |p_i|, x_i + |p_i|] \quad (13)$$

Then, clearly pi gives you upper and lower bound of xi, and for robust design, we are going to treat them as perturbation of design variable xi. For MDO, we would like to have larger perturbation in order to give more freedom of decisions for the designer of next level, at the same time, to keep deteriorations of the designers' behavior variables of their own level. Robust Design ensures you that within perturbations you have determined they are going to keep deterioration of objective functions within the amount of you have decided and also to keep constraints. Therefore, designers does not need to care for feedbacks from following levels. Besides, designers can show their decisions of deteriorations to the following levels to explain how much efforts they did for them. This information will be precious for following designers to be accept their decisions. Moreover, they can expect that decisions the following levels would not be the worst case for them.

**Results**

Results are shown in Table 2. Assume no body knows conventional results, but they all start from initial stages for each level, and they are shared. And all level1 designers knows their own optimum results but does not share their information. From level1-1, x1 is determined. Most likely, this decision is not accepted directly by level1-2 and level1-3. Most likely, they insist to get a little closer to their decisions. But even after they accept x1 and optimize the result shown in table, gaps between their own optimum are about 2 and 4. Same as x2 and x3, Level1-1 can insist their decision will be better than level1-2 and Level1-3 get a little worse even if they accept x2 and x3. Therefore, it might not be so hard to explain and make them accepted. Difficulty still stays there in Level1-3 and Level 2. As they break constraint g11 in final. But Level 2 does not have responsibility for g11. And Level1-3 does not know how to put side constraints on x5.

In proposed method, Table 3 shows trade-off analysis of Level1-1. At initial aspiration levels, we can see there are enough room for x2 and x3, but did not reach to the satisfaction level of objective functions. There are two ways to make function f1, that are giving more lower aspiration level or to make ideal value worse. We took later one in this time. In 2<sup>nd</sup> trial, it seems we have achieved satisfaction level on objective function. Therefore, we tried 3<sup>rd</sup> settings as to keep function f1 and try to have a little wide range for x2 and x3.

Table 2 Comparison of the results

	x1	x2	x3	x4	x5	x6	x7	f	f1	f2	f3	g	g1	g2	g3	g4	g5	g6	g7	g8	g9	g10	g11	
Conventional	3.5	0.7	17	7.3	7.715341	3.350215	5.286655	2994.356				0	0	0	0	0	0	0	0	0	0	0	0	0
Initial Stage	3.1	0.75	22.5	7.8	7.8	3.4	5.25	4144.84				0.2304	0	0	0	0	0	0.0207	0	0	0	0	0	0
Initial for Level1-1	3.1	0.75	22.5						2712.227			0.65	0	0	0	0	0				1	0	0	0
Initial for Level1-2	3.1	0.75	22.5	7.8		3.4				310.6569		0					0							0
Initial for Level1-3	3.1	0.75	22.5		7.8		5.25				1121.956	0.1496						0.036						0.1136
Initial for Level2	3.1	0.75	22.5	7.8	7.8	3.4	5.25	4144.84				0												
Set#1 Level1-1	3.5	0.7	17						1581.464				0	0	0	0	0					0	0	0
Set#1 Level1-2	3.5	0.8	28	7.3		3.345304				285.0207		0		0			0							0
Set#1 Level1-3	3.5	0.8	28		7.3		5.286656				1116.888	0			0			0						0
MDO Set #1	3.5	0.7	17	7.3	7.3	3.345304	5.286656	2983.373																
Check Set #1	3.5	0.7	17	7.3	7.3	3.345304	5.286656	2983.373				0.0616	0	0	0	0	0.0044	0.0005	0	0	0	0	0	0.0567
Set#2 Level1-1	3.5	0.7	17						1581.464				0	0	0	0	0					0	0	0
Set#2 Level1-2	3.5	0.7	17	7.3	7.8	3.350215	5.25	2973.066		286.2702		0		0			0							0
Set#2 Level1-3	3.5	0.7	17		7.71537		5.286656				1116.888	0			0			0						0
MDO Set #2	3.5	0.7	17	7.3	7.71537	3.350215	5.286656	2985.24																
Check Set #2	3.5	0.7	17	7.3	7.3	3.350215	5.286656	2985.24				0.0569	0	0	0	0	0	0	0	0	0	0	0	0.0569
	x1	x2	x3	x4	x5	x6	x7	f	f1	f2	f3	g	g1	g2	g3	g4	g5	g6	g7	g8	g9	g10	g11	
Proposed Level1-1	3.500209	0.700022704	17.00005649						1581.614			0	0	0	0	0	0					0	0	0
Proposed Level1-2	3.500209	0.700028378	17.00021727	7.644193		3.35148			改悪上限	1606.432		0			0		0							0
Proposed Level1-3	3.500209	0.700031907	17.00005649	-0.34227	7.848952		5.286761				1129.61	0			0			0						0
Proposed	3.500209	0.700022704	17.00005649	7.30192	7.715543	3.35148	5.286761	2994.961				0							0					
Final Results	3.500209	0.700022704	17.00005649	7.30192	7.715543	3.35148	5.286761	2994.961	1581.614	286.4518	1126.682	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 3 Results of trade-off analysis in Level 1-1

	f	p1	p2	z	gap	x1	x2	x3	p1	p2	z
1st	Ideal	0	0.1	1.5	0.005	5000					
	Aspiration	800	0.06	1	0.01	4000					
	Scalar	2.754853	2.340413	2.384278	2.754852	2.754854					
	Result	2203.882	0.006383	0.307861	0.018774	2245.146	3.598081	0.708147	19.77142	0.006383	0.307861
2nd	Ideal	500	0.1	1.5	0.005	5000					
	Aspiration	800	0.06	1	0.01	4000					
	Scalar	2.754853	2.340413	2.384278	2.754852	2.754854					
	Result	1581.614	0.000115	0.001977	0.011645	1581.639	3.501406	0.700125	17.00199	0.000115	0.001977
3rd	Ideal	600	0.15	2.1	0.009	5000					
	Aspiration	800	0.1	0.5	0.05	3300					
	Scalar	4.908069	2.999894	1.312379	0.16231	2.010801					
	Result	1581.614	5.32E-06	0.000194	0.015655	1581.639	3.500167	0.700015	17.00093	5.32E-06	0.000194

Unfortunately, we cannot get wider ranges so we decided to quit trade-off analysis. From given x1 and given upper and lower bound for x2 and x3, Level1-2 is formulated as x1 is fixed, x2,x3 usual design variables with given upper and lower bound from Level1-1, x4 is interval variable, x6 is usual design variables with given original upper and lower bounds. Level1-3 is formulated as x1 is fixed, x2, x3 usual design variables with given upper and lower bounds from Level1-1, x5 is interval variable and x7 is usual design variable with original upper and lower bounds. Finally Level2 is formulated as x1 is fixed, x2, x3 is usual design variable with original upper and lower bounds, x4 is usual design variables with upper and lower bounds from Level1-2, x5 is usual design variables with upper and lower bounds from Level1-3, x6 is fixed from Level1-2, x7 is fixed from Level1-3. Difference between usual MDO and proposed method can be seen in x5. As usual MDO did not have any lower bound to keep constraint g11 within active area, Level 2 in usual MDO gave x5=7.3 as lower bound. On the other hands proposed method as lower bound from robust design, therefore is lower bound are limited to keep constraint g11. As a results, we do not need to carry out any feedback at all. Moreover, if you look at the final results, you can see they are always better than the worst case decisions. Thus, they do not need any feedbacks. With power of robust design we can well control behavior variables even with interval variables.

### Conclusion

There are requirements to treat large scale optimization in industries. However, size that we can treat effectively in real use might be up to 6 or 7 design variables, even in simple behavior function cases. In order to extend their size, it might be good choice to make formulation in MDO, multi-level optimization. Theoretical idea is great, but at the same time, there are a lot of compromise decisions between levels so that we need to carry out a number of feedbacks. It will occasionally never converge. To avoid it, it is recommended to give upper and lower bound for design variables for following levels to cut the feedback loop. But there was no specific method to give these bounds effectively. In this paper, reason why robust design works effectively in MDO, multi-level design. With robust design we can treat interval variables. Interval variables mean giving upper and lower bound for design variables. Within its perturbations, we can control deterioration of objective functions and also ensure to keep constraints for behavior functions. Its effectiveness are shown through famous benchmark problem call Golinski's speed reducer.

### Acknowledgements

I have been working on Robust design ever since my supervisor Prof. Hiroshi Yamakawa at Waseda University told us it might be interesting to be researched. Prof. Hiroyuki Sugimoto told me that although we have been working on robust design, he didn't see any good application of it. Therefore, I have been looking to see good application of robust design. Their advice is greatly acknowledged.

### References

- [1] Unesaki, H., Arakawa, M., Hiramatsu, S., Kondo, S., Butsuen, T., Robust design Using Interval Design Variables for MDO, Proc. Asian Congress of structural and Multidisciplinary Optimization, 2020. Paper No. P00228
- [2] Kroo, I., Altus, S., Braun, B., Gage, P. and Sobieski, J., Multidisciplinary Optimization Methods for Aircraft Preliminary Design, Proc. the 5th AIAA/ USAF/ NASA/ ISSMO Symposium on Multidisciplinary Analysis and Optimization, AIAA 94-4325,(1994).
- [3] Arakawa, M., Yamakawa, H., Ishikawa H., Robust Design Using Fuzzy Numbers, Proc. of Design Technical Conference (in CD-ROM), ASME, St. Louis,1998.9
- [4] Kobayashi, T., Arakawa, M., Butsuen, T., Comparison between worst case optimization and conventional robust design, Prof. Asian Congress of Structural and Multidisciplinary Optimization, 2018.5, Dailia
- [5] <http://www.eng.buffalo.edu/Research/MODEL/mdo.test.orig/class2prob4/descr.html>
- [6] Ku,K.J., Rao, S.S. and Chen, L., Taguchi-aided search method for design optimization of engineering systems, *Engineering Optimization*, vol. 30, no. 1, pp. 1-23, 1998.
- [7] Akhtar,S. , Tai,K. and Ray.T., A socio-behavioral simulation model for engineering design optimization, *Engineering Optimization*, vol.34, no. 4, pp.341-354, 2002

- [8] Rao, S.S. and Xiong, Y., "A hybrid genetic algorithm for mixed-discrete design optimization," *Journal of Mechanical Design*, vol.127, no. 6, pp. 1100-1112, 2005
- [9] Cagnina, L.C., Esquivel S.C., and CoelloCoello, C.A., "Solving engineering optimization problems with the simple constrained particle swarm optimizer," *Informatica*, Vol. 32, No.3, pp.319-326, 2008
- [10] Tosserams, S., Etman, L. F. P., and Rooda, J. E., "An augmented Lagrangian decomposition method for quasi-separable problems in MDO," *Structural and Multidisciplinary Optimization*, Vol. 34, No. 3, pp.211-227, 2007
- [11] Zummund, A., "Design and Optimization of Golinski's Speed Reducer by using the five methods of Optimization (SQP, GA, SA, PSO, and Hybrid Algorithm)," *European International Journal of Science and Technology*, Vol. 9, No. 10, pp.29-60., 2020
- [12] Lin, M-H., Tsai, J-F., Hu, N-Z. and Chang, S-C, "Design Optimization of a Speed Reducer Using Deterministic Techniques," *Mathematical Problems in Engineering*, Vol. 2013, Article ID 419043, 2013
- [13] Arakawa, M., Shoda, K., Kitayama, S., "Development of Multi-Objective Agent PSO," *CJK-OSM5*, Jeju, South Korea, 2008.6

## Traffic Context Augmentation for Roadway Environment Situation Awareness of Autonomous Driving

Sunghwan Kim<sup>1</sup>, Ir. Made Sudarma<sup>2</sup>, Deden Witarsyah Jacob<sup>3</sup>, Minho Jo<sup>1</sup>

<sup>1</sup>Department of Computer Convergence Software, Korea University, South Korea

<sup>2</sup> IT Track, Department of Electrical Engineering, Udayana University, Indonesia

<sup>3</sup>School of Industrial and Systems Engineering, Telkom University, Indonesia

[hwann@korea.ac.kr](mailto:hwann@korea.ac.kr), [msudarma@unud.ac.id](mailto:msudarma@unud.ac.id), [dedenw@telkomuniversity.ac.id](mailto:dedenw@telkomuniversity.ac.id), [minhojo@korea.ac.kr](mailto:minhojo@korea.ac.kr)

### Abstract

In this paper, we propose a multi annotated traffic context augmentation dataset that includes class, behavior, and location information of objects in driving environment. Class of objects refer to the type properties of each object. Location of objects refer to the semantic location of the area where each object exists, and behavior of objects refer to the action that each object is currently performing. We developed an annotation program for multi-annotation of behavior and location for multi-objects. We proposed a situation awareness model for object's location and behavior statement detection to validate the proposed dataset. We had experiments on YOLOX and YOLOv6 based situation awareness model and showed a comparison table to prove that the proposed situation awareness model performs better.

**Topics:** computational intelligence machine learning, vision/image processing

### Introduction

In recent years, autonomous driving has emerged as a fast-growing research area. Deep learning-based systems are being researched in the mainstream and showing successful performance. While self-driving cars rely on deep learning to make decisions in real-world situations, major reasons for concern remain in terms of safety, ethics, cost, and reliability [1]. In terms of safety, self-driving cars need to understand the roadway environment and make appropriate decisions accordingly. Therefore, situational awareness and the ability to accurately understand the behavior of other road users are required for autonomous driving in a complex road environment in the city.

Training dataset plays a very important role for successful deep learning model training and development. Although various datasets for autonomous driving are being released by various institutions, it is difficult to learn a reliable deep learning model because the platform used to collect the dataset and the environment in which the actual experiment is conducted are different.

There are many things to consider for self-driving cars in urban areas. Even the same object in an autonomous driving environment has different risks depending on the location and behavior of the object. Existing autonomous driving dataset only provide information on object, lane and semantic segmentation information [2][3]. There is a need to research and develop a dataset that define location and behavior of surrounding objects [4].

In this paper, we propose a multi annotated traffic context augmentation dataset that includes behavior and location information for objects in driving environment. We developed an annotation program for multi-annotation of behavior and location for multi-objects. We trained a model for object's location and behavior statement detection to validate the proposed dataset.

### Traffic Context Augmentation Dataset

To detect situation awareness of campus roadway environment, this paper uses two-dimensional multi-object detection technique. The two-dimensional multi-object detection technique recognizes the locations of target objects in the input RGB image as two-dimensional boxes and classifies the classes of the objects. In addition to detect the position and class of a box in a conventional multi-object detector, we define and classify its location and behavior properties.

Class of objects refer to the type properties of each object. Location of objects refer to the semantic location of the area where each object exists, and behavior of objects refer to the action that each object is currently

performing. We defined that an object's class, location, and behavior can only have one attribute, with no overlap. The roadway environment situation awareness result is shown in the Figure 1. The position of detected objects is represented by a box form, the class is represented by the color of the box, the location and the behavior is represented the property information as letters with index.

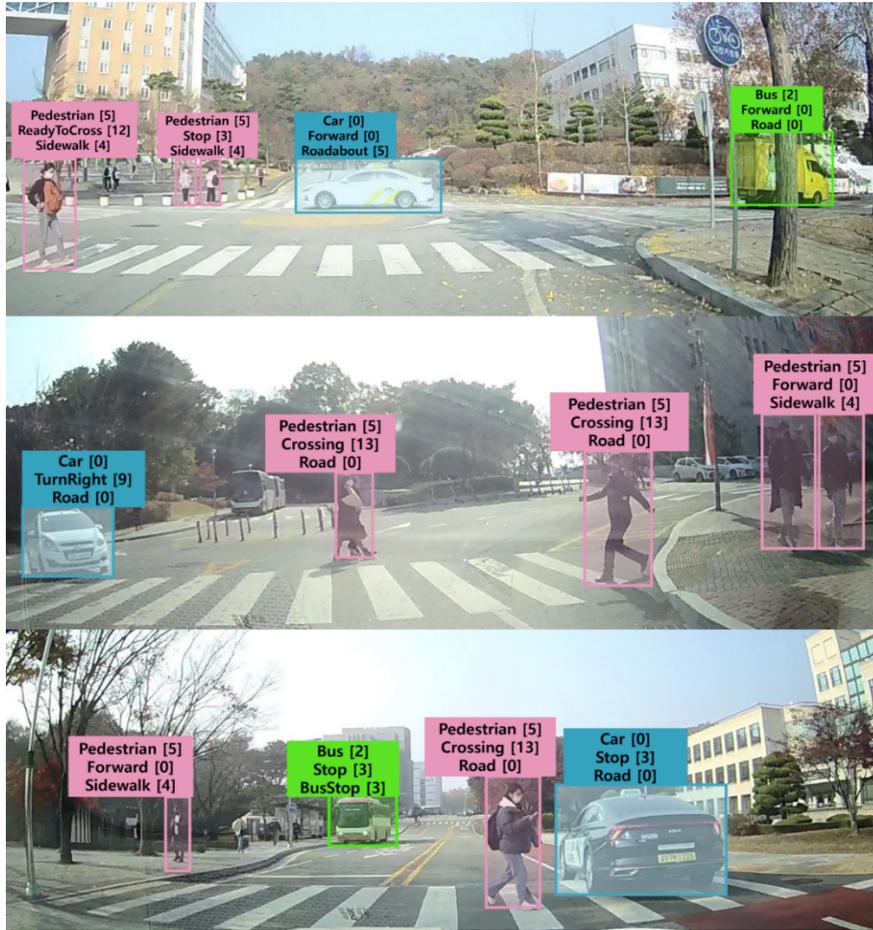


Fig 1: Multi-annotation results of roadway environment situation awareness

*Class of objects.* In the campus road environment, the objects or people able to perform actions which can influence the decision made by the autonomous vehicle are classified : car, bus, electric kickboard, motorbike, bicycle, pedestrian.

*Behavior of objects.* We have defined movements that can be performed by cars, vehicles, and pedestrians. For cars, we defined the following actions: moving forward, backward, slowing down, stopping, left and right turn signal, hazard lights, turning left, turning right, and turning. In addition, pedestrian movements are defined as moving forward, preparing to cross, and crossing, and vehicle movements were defined as moving forward, backward, slowing down, stopping, turning left, turning right, and turning.

*Location of objects.* We propose to label the location of each object from the perspective of the autonomous vehicle. For example, a pedestrian can be found on the sidewalk, while crossing on the road. Locations of objects are defined : road, parking lot, bike lane, bus stop, sidewalk and roundabout.

Label name	Description
Car	A car up to the size of a multi-purpose vehicle
Bus	A single or double-decker bus or coach
Electric kickboard	A person is riding electric kickboard
Motorbike	Motorbike, dirt bike, scooter with 2/3 wheels
Bicycle	A person is riding a push/electric bicycle
Pedestrian	A person including children

Table 1: List of Traffic Context Augmentation Dataset Objects classes with description

Label name	Description
Moving forward	Object moving in the direction from it is currently going

Moving backward	Object moving in the opposite direction from it is currently going
Slowing down	Object slows down with rear brake lights shining
Stopping	Object stationary but in ready position to move
Left and right turn signal	Object indicating left/right by flashing left/right indicator light
Turning left and right	Object is turning in left/right direction
Turning	Object is turning at roundabout
Preparing to cross	Agent on a pavement, stationary, facing in the direction of the road
Crossing	Agent crossing road

**Table 2:** List of Traffic Context Augmentation Dataset Behavior classes with description

Label name	Description
Road	Object is on the road
Parking lot	Object is in parking lot
Bike lane	Object is in bike lane
Bus stop	Object is at bus stop
Roundabout	Object is at roundabout
Sidewalk	Object is on sidewalk

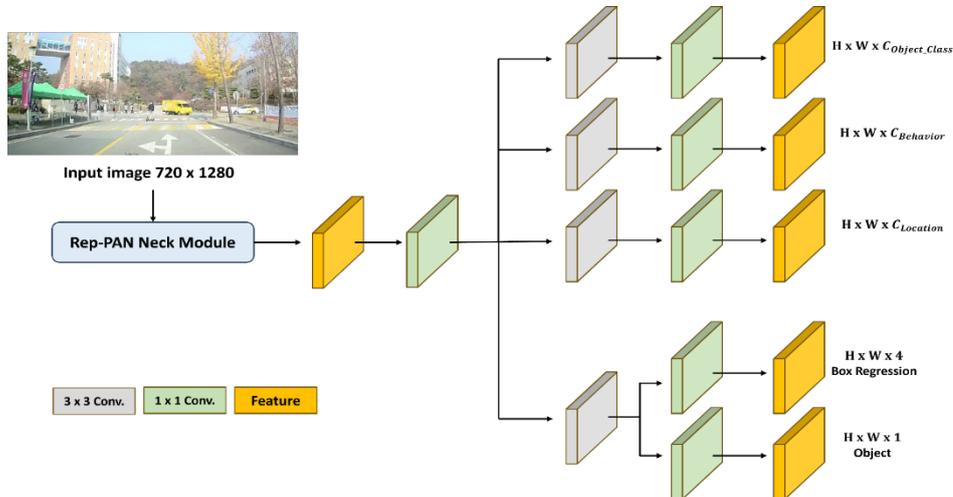
**Table 3:** List of Traffic Context Augmentation Dataset Location classes with description

### Multi-annotation program for multi-objects

We developed a program to describe the class, behavior, and location properties of objects using the campus roadway environment data collected on campus. Existing annotation programs do not provide functions to assign more than one attribute to an object, so we developed a program for our purpose. Our program was developed to be cross-platform, working in various platform environments: Windows, MacOS, and Linux. The annotated data is saved as txt and json files in the form <class id>, <action id>, <location id>, <x\_center>, <y\_center>, <width>, <height>.

### Situation Awareness Model

YOLOv6 model structure to train the road environment situational awareness dataset proposed in this study, we modified the previously published YOLOv6 model and proposed situation awareness model [5]. YOLOv6 model uses Rep-PAN(Path Aggregation Network) neck module to improve accuracy and computation cost. Rep-PAN module improves localization performance by effectively passing lower layer feature information to the top layer through bottom-up path augmentation. Since the existing YOLOv6 model has a structure that can only classify the class of objects, we modified the structure as shown in Figure 2 to further detect the behavior and location of objects. The model was modified to have separate branches and heads because the class, location, and behavior of objects are expected to have different characteristics and areas that need to be referenced, whereas the existing model had one branch and head. The decoupled head structure provides faster convergence and improved AP(Average Precision) than the traditional coupled head method [6].



**Fig 2:** Proposed YOLOv6 based situation awareness model architecture

### Experiments

To build a traffic context augmentation dataset that further defines the class, behavior, and location of objects, we collected data from Korea University Sejong Campus in South Korea, collecting a total of 80,972 uncensored photos. The data was mainly collected during daylight hours with good visibility and includes rainy weather conditions. We selected and annotated 4800 photos where the class, location, and behavior of the objects are well represented, taking into account shake, viewing angle, and speed of the vehicle.

From the generated dataset of 4800 images, 3600 images were used as training data and 1200 images were used as validation data to train and validate the situational awareness model. The results are evaluated in terms of frame-level bounding box detection, and the evaluation measure is mean average precision(mAP).

Table 4 shows the mAP comparison for the YOLOX and YOLOv6 models on the proposed traffic context augmentation dataset. The proposed YOLOv6 model shows that it detects the class, location, and behavior of objects better than the YOLOX model. In particular, the proposed model is better at detecting the behavior of objects in our dataset because it is simpler than the existing method to characterize the behavior of objects based on their location.

Model	Attributes	mAP
YOLOX based Situation Awareness Model	Class of Objects	78.5%
	Location of Objects	57.5%
	Behavior of Objects	45.2%
YOLOv6 based Situation Awareness Model	Class of Objects	82.4%
	Location of Objects	58.6%
	Behavior of Objects	53.4%

**Table 4:** Comparison mAP on YOLOX and YOLOv6 based situation awareness model for proposed traffic context augmentation dataset

Table 5 shows the mAP comparison by behavior class of objects for the proposed YOLOv6 based situation awareness model. This table shows mAP of each behaviors to find out which behaviors are not well detected. Moving forward, Preparing to cross and Crossing classes show higher mAP results than other behaviors. Since our proposed dataset consists of image data, our model had difficulty detecting objects in situations where objects are moving backward, slowing down, or stopping in still images. On the other hand, the performance was relatively good when the objects are moving forward and crossing. When objects are turning, the performance is poor due to the resolution of the image and the ambiguity of the situation.

Behavior Class	Validation	Test
Moving forward	72%	72.3%
Moving backward	34.7%	30.4%
Slowing down	30.6%	33.2%
Stopping	45.8%	42.3%
Left and right turn signal	56.5%	54%
Turning left and right	55.4%	53.2%
Turning	63.8%	62%
Preparing to cross	64.3%	62.3%
Crossing	67.5%	68.3%

**Table 5:** Comparison of mAP by behavior class of objects for the proposed YOLOv6 based situation awareness model

## Conclusion

In this paper, we propose a multi annotated traffic context augmentation dataset that includes class, behavior and location information for objects in driving environment. Class of objects refer to the type properties of each object. Location of objects refer to the semantic location of the area where each object exists, and behavior of objects refer to the action that each object is currently performing. We developed an annotation program for multi-annotation of behavior and location for multi-objects which is not existing. proposed a situation awareness model for object's location and behavior statement detection to validate the proposed dataset. We had experiments on YOLOX and YOLOv6 based situation aware model and showed a comparison table to prove that the proposed model performs better. In future research, we will change the structure of YOLO's backbone network and head network to improve the performance of object location and behavior.

## **References**

- [1] M. e. a. Maurer, *Autonomous driving: technical, legal and social aspects*. Springer Nature, 2016.
- [2] Fisher Yu, Haofeng Chen, Xin Wang, Wenqi Xian, Yingying Chen, Fangchen Liu, Vashisht Madhavan, Trevor Darrell, "BDD100K: A Diverse Driving Dataset for Heterogeneous Multitask Learning", *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 2636-2645.
- [3] Pei Sun, Henrik Kretzschmar, Xerxes Dotiwalla, Aurelien Chouard, Vijaysai Patnaik, Paul Tsui, James Guo, Yin Zhou, Yuning Chai, Benjamin Caine, Vijay Vasudevan, Wei Han, Jiquan Ngiam, Hang Zhao, Aleksei Timofeev, Scott Ettinger, Maxim Krivokon, Amy Gao, Aditya Joshi, Yu Zhang, Jonathon Shlens, Zhifeng Chen, Dragomir Anguelov, "Scalability in Perception for Autonomous Driving: Waymo Open Dataset", *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 2446-2454.
- [4] G. Singh, et al., "ROAD: The Road Event Awareness Dataset for Autonomous Driving" in *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 45, no. 01, pp. 1036-1054, 2023.
- [5] Chuyi Li, Lulu Li, Hongliang Jiang, Kaiheng Weng, Yifei Geng, Liang Li, Zaidan Ke, Qingyuan Li, Meng Cheng, Weiqiang Nie, Yiduo Li, Bo Zhang, Yufei Liang, Linyuan Zhou, Xiaoming Xu, Xiangxiang Chu, Xiaoming Wei, Xiaolin Wei, "YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications", *arXiv:2209.02976*, 2022.
- [6] Ge, Zheng and Liu, Songtao and Wang, Feng and Li, Zeming and Sun, Jian, "YOLOX: Exceeding YOLO Series in 2021", *arXiv:2107.08430*, 2021.



## Portable ECG Monitor on Edge Device using TensorFlow Lite

\*Cheonghwan Hwang, \*\*Nagyoum An and \*Jaehyun Park

\*Department of Electronics and Computer Engineering

\*\*Department of Computer Engineering

Inha University, Korea

chhwang@emcl.org, ymj08073@gmail.com, jhyun@inha.ac.kr

### Abstract

In recent years, the healthcare and wellness industry has grown dramatically due to the increased focus on health. Early detection of cardiovascular disease through electrocardiogram monitoring is becoming increasingly important. Recent advances in artificial intelligence have made great strides in the early detection of cardiovascular disease through electrocardiogram. This paper presents a method to monitor electrocardiogram in real time by implementing artificial intelligence technology on edge device using IoT technology, unlike the existing server-based artificial intelligence technology. Since the edge device used in this paper has limited resources, the TensorFlow Lite that is specially designed to support real-time machine learning (ML) inference on low-power and resource-constrained edge devices was used. The performance of the pre-trained model and the lightweight model was compared through TensorFlow Lite to confirm whether the application was suitable.

**Topics:** Embedded System, Healthcare, electrocardiogram, Deep Learning, TensorFlow Lite Micro

### Introduction

Health and wellness are paramount concerns in modern society, leading to significant growth in the healthcare industry. Among various health threats, cardiovascular disease (CVD) stands out as one of the leading causes of death globally. This trend, exacerbated by an aging population, underscores the importance of early and accurate diagnosis. The primary diagnostic tool for detecting heart rhythm disorders, a leading cause of cardiovascular disease, is the electrocardiogram (ECG) which records the heart's electrical activity at specific moments [1], [2].

Myocardial cells, stimulated by the heart's conductive system, undergo cycles of contraction and relaxation. An electrocardiogram captures this electrical activity, with waveforms typically divided into P, QRS Complex, T, and (U) sequences. A normal electrocardiogram cycles in the P-QRS-T-(U) order. The QRS complex represents the electrical activity of the ventricles: The Q waveform starts with the first negative waveform, the R waveform is the first positive waveform representing the polarization of the ventricle, and the S waveform is the second negative waveform representing the non-polarization of the ventricle.

Traditional algorithmic methods utilized in conventional electrocardiogram monitoring systems have been proven as a standard diagnosis method. However, while they showed relatively good performance for the standardized scenarios and data, they showed the limited capacity to the range of an individual's attributes such as heart condition, or changes in response to certain disease states. Recent research has explored the use of machine learning techniques, including deep learning, to identify and analyze individual characteristics and variability in electrocardiogram data with improved accuracy. Deep learning enables automatic identification of complex patterns from large datasets, which results in precise characterization of an individual's electrocardiogram and faster detection of early signs or abnormalities in heart disease [3].

Advancements in computing devices and high-speed mobile networking have enabled global deployment of deep learning applications via the cloud services. Nevertheless, large-scale distributed applications have not been suitable for cloud computing due to the network capacity [4] and latency produced when the application is far from the cloud [5]. To overcome these issues of cloud computing, the edge computing was proposed [6], which carries out tasks at the edge of the network. Machine learning applications including deep learning also likely to use these edge computing concept to expand the real-time and on-line services. One of the widely used deep learning frameworks is TensorFlow, an open source based deep learning framework. To optimize inference speed and enhance performance with reduced memory usage to improve efficiency and minimize size, TensorFlow Lite uses FlatBuffers rather than the standard protocols used in TensorFlow. TensorFlow Lite Micro

is designed for ultra-low power microcontrollers and other devices with limited memory, enabling machine learning models to run effectively with only a few kilobytes of memory [7].

To adopt a deep learning algorithm for cardiovascular disease diagnosis, we utilized the MIT-BIH arrhythmia database as the electrocardiogram dataset for our experiment. This database has been the standard for arrhythmia detector evaluation since 1980. It has been employed in various studies, including foundational research on cardiac mechanics. The data in this database was collected from 47 subjects using two lead configurations: V1 (which can occasionally be V2, V4, or V5, depending on the subject) and Modified Limb Lead II (MLII)[8]. V1 is the precordial (or thoracic) lead, which is in the fourth intercostal space to the right of the sternum. The electrical activity of the heart can be seen in the precordial position. Modified limb leads II (MLII) records the electrical difference between the right arm (cathode) and left leg (anode).

In this paper, we aimed to detect arrhythmia by implementing CNN-based deep learning models specialized in electrocardiogram analysis for embedded devices. Lighten the model for use in microcontrollers, deploy the model with Arduino Nano 33 BLE Sense as Edge Node, and analyze the signal measured by the electrocardiogram sensor in real time.

### Proposed method

The proposed hardware system configuration is shown in Figure 1. The electrocardiogram monitoring device utilizes an Arduino Nano 33 BLE Sense supplied by Nordic Semiconductors. The device has a nRF52840 SoC powered by a 32-bit ARM Cortex™-M4 architecture CPU with a floating-point unit, capable of running at 64 MHz [9]. It is energy-efficient, making it ideal for low power usage, and is compatible with the TensorFlow Lite Micro deep learning framework to infer electrocardiogram data as an edge device. With one megabyte of internal flash memory, the device can execute models that are 500-600 KB in size, not including the API programs necessary to operate deep learning models, thus classify three types of images.

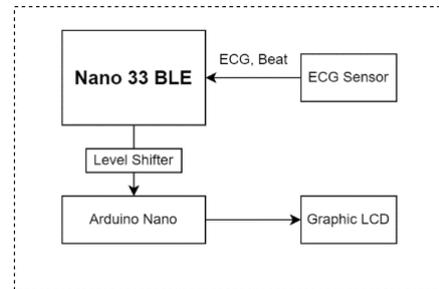


Fig.1: System Block diagram

The MIT-BIH arrhythmia dataset consists of Normal(N), Atrial premature(S), Premature ventricular contraction(V), Fusion of ventricular and normal(F), and Paced(Q) depending on the arrhythmia state [8]. Due to the limited size of the model that can be deployed due to the constraints of device memory space, separate data were used for training and inference. The selected training data was extracted from the MIT-BIH arrhythmia database, categorized as noise data including steady-state electrocardiogram (N) and ventricular premature contraction (PVC) data during arrhythmia and time series data. The QRS waveform is a crucial feature in electrocardiogram. data, providing essential information on the heart's electrical activity. In our dataset, we identified QRS waveforms using vertices with heights ranging from 1250 to 1500 and a 200-unit distance. To optimize the data input size when the model is run on the hardware device, we narrowed our focus to the QRS and P waves by cropping 85 pixels from the left and 15 pixels from the right around the vertices of the identified QRS waveforms. To facilitate data analysis, the image was resized to 96 x 96 using the matplotlib module. Additionally, the data was processed to emphasize the distinguishing characteristics of normal and arrhythmic patterns.

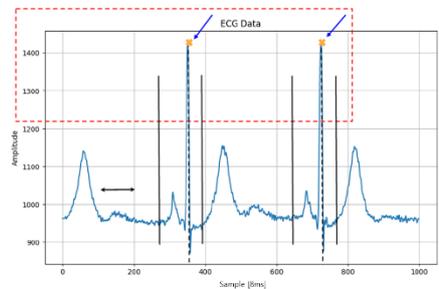


Fig.2: Data Preprocessing

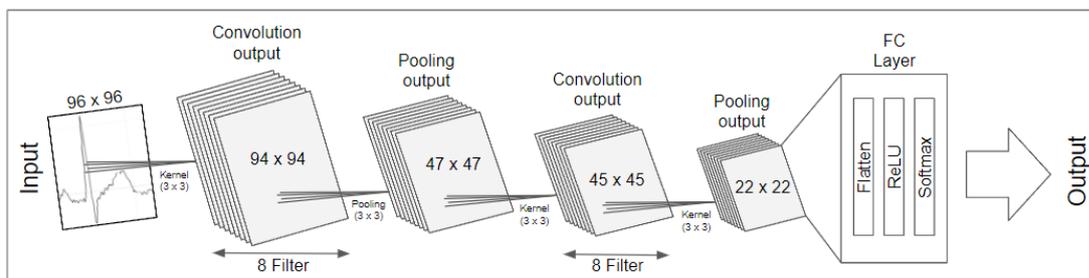


Fig.3: Implemented Model Configuration

The CNN model implemented comprises two convolutional layers, two pooling layers, and one fully connected layer. The fully connected layer is made up of Flatten, ReLU activation function, and Softmax classification. The ReLU activation function was chosen to address the gradient vanishing issue, as it can return 0 for negative input values and its corresponding value for positive ones. While some nodes in the hidden layer may have zero gradients, this does not seem to be a significant issue as the majority have values higher than zero.

As a technique for model optimization, we selected the Adam optimization method, which combines the strengths of Momentum and RMSprop to offer a more effective learning process. We opted for 'categorical\_crossentropy' as the loss function, which is appropriate for multi-class classification, considering the scalability of the project to classify input data into more precise categories. 'categorical\_crossentropy' is a frequently used loss function that trains a model by minimizing the difference between the probability distributions of each class.

The commonly implemented TensorFlow model has capacity issues that make it difficult to apply to small edge devices using microcontrollers, so there is a need to apply model lightweighting and optimization techniques. When we initially implemented the model with a focus on performance and used 32 filters, the size of the converted TensorFlow Lite model was 2.4MB, which exceeded the flash memory. To solve this problem, we limited the size of the input layer and adjusted the number of filters in the internal convolution layer to implement a model that can be accommodated by limited resources. This approach risks impacting performance, so we conducted several experiments to find the optimal balance between lightweight design and optimal performance, setting the input image to 96 x 96 and using three filters to achieve the best compromise between model accuracy and lightweight structure.

The model was trained using TensorFlow and converted to TensorFlow Lite for microcontrollers for inference on the Arduino Nano 33 BLE Sense. In this process, the training model, consisting of weights and biases in floating point format, is converted to a TensorFlow Lite FlatBuffer file. This file has the weights converted from traditional floating-point weights to an array of C bytes in 8-bit integer format with quantization [10]. The implemented model was 503 KB and deployed normally to the device.

In addition, to improve the model implemented in this way, re-learning was conducted using individual electrocardiogram data. Personal data was collected in a variety of environments and conditions, which were used to improve the generalization ability of the model. The model after re-learning evaluated the difference in accuracy compared to the original model.

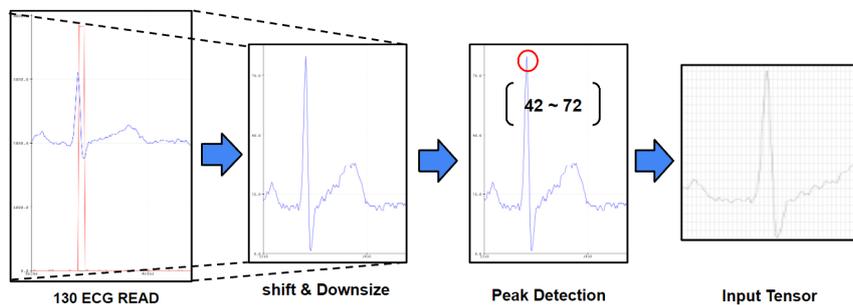
	Base model	Re-trained model
	Accuracy	Accuracy
32 Layer	99.23%	99.99%
8 Layer	88.72%	99.99%
TF Lite	53.75%	83.55%

**Table 1:** Model Performance Comparison

Table 1 displays the accuracy achieved during the model lightening process and its conversion to TensorFlow Lite, as well as the improved accuracy obtained after retraining with individual electrocardiogram data. The results underscore the potential to enhance the model's predictive capabilities in electrocardiogram analysis by incorporating personal data for retraining.

### Experiments

The MIT-BIH electrocardiogram database includes two types of leads: Lead I/II (VI) and Modified Lead II (MLII). MLII is a frequently used type that allows for observation of the heart's overall electrical characteristics by measuring electrical activity in a transverse direction. Therefore, we utilized the MLII lead type to capture and analyze the characteristics of the QRS complex more effectively.



**Fig.3:** Implemented Model Configuration

Electrocardiogram data were sampled using the MLII lead method. Each sample included 130 values at intervals of 8.3ms (70 SPS). To adjust the range of the electrocardiogram data, the minimum value was shifted closer to zero. The 12-bit resolution data values were then reduced from a maximum of 4095 to 96 to boost processing efficiency. The R-wave position of the electrocardiogram signal was processed to be centered by verifying whether the signal values above 70 were between 48 and 72 and centered accordingly. To enhance computational efficiency, the electrocardiogram values at the  $(x + \text{peak} - 48)$ th electrocardiogram array location were converted to grayscale. This reduced the data complexity and minimized the model computation. Using this experiment method, the data was processed in a way that captured the ventricular arrhythmias' features and maximized the model's computational efficiency.

Since direct experiments on arrhythmia patients are limited, we developed a validation process utilizing genuine patient electrocardiogram data. The following section provides a detailed description of the methodology used to validate the model. To minimize the data size while maintaining the original electrocardiogram waveform characteristics, the 96 x 96 electrocardiogram waveform images (.bmp) utilized for training were converted to 1-bit values using a tool. Instead of using the signal from the actual electrocardiogram sensor, we directly input the 1-bit converted electrocardiogram waveform data that was preprocessed above into the model. This serves as a proxy for the absence of real patient data, and we use it to validate the model's ability to make predictions on the data it was trained on. We cross-fed the arrhythmia and normal heart rate data to evaluate the model's performance and ensure its ability to make predictions in diverse situations. This provided an initial evaluation of the functionality and user-friendliness of the model in a practical context on an edge device.

## Conclusion

In this paper, we proposed a deep learning-based edge computing node that can be used for the diagnosis of Cardiovascular disease with real-time monitoring an electrocardiogram. To implement real-time inference, the proposed system utilized TensorFlow Lite Micro engine on a Arduino Nano 33 BLE. We used a CNN-based model to convert electrocardiogram waveforms into images of heart rate units and the converted images were trained. The learned model was optimized for a tiny edge computing device where the real-time electrocardiogram signal was captured, and inference of diagnosis was performed using TensorFlow Lite Micro.

Since the electrocardiogram was learned via a method based on deep learning, it can easily analyze diverse environments and complex data patterns with high accuracy, depending on the model and trained data. Furthermore, by retraining the individual's heartbeat data, it is expected to exhibit high accuracy that is specific to the user.

## Acknowledgment

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the Innovative Human Resource Development for Local Intellectualization support program(IITP-2023-RS-2023000259678) supervised by the IITP(Institute for Information & communications Technology Planning & Evaluation)

The authors would like to express their appreciation to Hoyoung Son and Minsub Kim at Inha University for helping experiments to improve the quality of this paper.

## References

- [1] D. Lloyd-Jones "Heart disease and stroke statistics 2009 update, A report from the American Heart Association statistics committee and stroke statistics subcommittee", *Circulation*, vol. 119, no. 3, pp. 21 - 181 2009.
- [2] D. Buxi et al., "Correlation Between Electrode-Tissue Impedance and Motion Artifact in Biopotential Recordings," in *IEEE Sensors Journal*, vol. 12, no. 12, pp. 3373-3383, Dec. 2012.
- [3] P. Rajpurkar, et., "Cardiologist-level arrhythmia detection with convolutional neural networks," arXiv preprint arXiv:1707.01836, 2017.
- [4] X. Ke, W. Hou and L. Meng, "Research on Pet Recognition Algorithm With Dual Attention GhostNet-SSD and Edge Devices," in *IEEE Access*, vol. 10, pp. 131469-131480, 2022.
- [5] C. A. A. Valencia, R. S. S. Suliva and J. F. Villaverde, "Hardware Performance Evaluation of Different Computing Devices on YOLOv5 Ship Detection Model," 2022 IEEE 14th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), Boracay Island, Philippines, pp. 1-5, 2022.
- [6] N. T. Ha et al., "Leveraging Deep Learning Model for Emergency Situations Detection On Urban Road Using Images From CCTV Cameras," 2022 International Conference on Engineering and Emerging Technologies (ICEET), Kuala Lumpur, Malaysia, pp. 1-5, 2022.
- [7] TensorFlow Lite for Microcontrollers, Aug 2023, [online] Available: <https://www.tensorflow.org/lite/microcontrollers>.
- [8] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH Arrhythmia Database," in *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45-50, May-June 2001.
- [9] nRF52840, Aug 2023, [online] Available: <https://www.nordicsemi.com/products/nrf52840>. jhyun@inha.ac.kr

## Multi-Input Shape CNN-based Application Traffic Classification

Ui-Jun Baek, Jee-Tae Park, Jeong-Woo Choi  
Seung-Woo Nam, Jae-Won Park, Gyeong-Min Yu, Myung-Sup Kim

Computer Information and Science  
Korea University, Korea

{pb1069, pj5846, choigoya97, nam131119, 2018270614, rudals2710, tmskim}@korea.ac.kr

### Abstract

Due to the rapid advancement of the internet and online applications, network traffic classification has become a crucial topic in the field of network management. Recently, many traffic classification methods based on AI, particularly deep learning, have been proposed. However, these methods often focus solely on altering the hierarchical structure of deep learning models and do not consider the shape of the incoming traffic data. We propose a classification method that utilizes various input shapes that can be derived from fixed-length packet bytes. Our proposed approach achieves superior classification performance compared to previous research that only utilized two-dimensional square-like input shapes and one-dimensional linear input formats from publicly available datasets.<sup>1</sup>

**Topics:** traffic classification, traffic identification, convolutional neural network, DL-based classification.

### Introduction and background

Traffic classification is a key technology in network traffic monitoring and analysis, involving grouping similar or related traffic and classifying it into predefined categories. This technique serves various crucial purposes. Firstly, it aims to troubleshoot network issues, such as locating faulty network devices, hardware/software misconfigurations, and points of packet loss within the network. Secondly, it ensures the overall acceptability of applications by managing quality of service (QoS), including bandwidth resources and cloud service usage. Thirdly, it plays a pivotal role in network security, enabling the distinction between normal and malicious traffic for security measurement and intrusion detection. Historically, widely used traffic classification methods include port-based classification and payload-based classification. However, these methods face limitations due to the application of dynamic ports and payload encryption. Although machine learning-based classification methods have been actively proposed, they fall short in analyzing the diverse traffic patterns in complex network environments. Recently, AI, especially deep learning-based classification methods, have gained attention. High-performance models based on Convolutional Neural Networks (CNN), which have shown remarkable performance in computer vision, have been applied in research.

Paper No.	Category	DL Method	Features	Shapes
[1] 2018	IDS	CNN	Header, Payload	Sqaure
[2] 2019	APP TC	CNN	Payload	Sqaure
[3] 2020	APP TC	CNN	Header, Payload	Sqaure
[4] 2021	APP TC	CNN, LSTM	Header	Linear
[5] 2022	APP TC	CNN	Header	Linear

**Table 1:** CNN-based TC studies

CNN is widely utilized for traffic classification in deep learning. It involves truncating packets to a fixed length before entering the learning model. These truncated packets are then transformed into 2D square-shaped or 1D linear-shaped vectors. Previous research has predominantly focused on reshaping raw packet bytes into 2D matrices resembling squares. From these matrices, spatial features are extracted for network traffic classification using CNN. These studies are consolidated in Table 1. Our investigation reveals the historical predominance of the two-dimensional square-like input shape. However, a recent shift has occurred towards the consistent adoption of 1-D linear input shapes. This shift is presumably informed by empirical studies indicating superior performance of linear input shape-based traffic classification models compared to their square-like counterparts [6]. Notably, the distinct characteristic of raw packet bytes is their lack of clear adjacency among data points, unlike conventional images where pixel relationships are evident. Unfortunately, most studies have not

---

This results was supported by "Regional Innovation Strategy (RIS)" through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(MOE)(2021RIS-004)

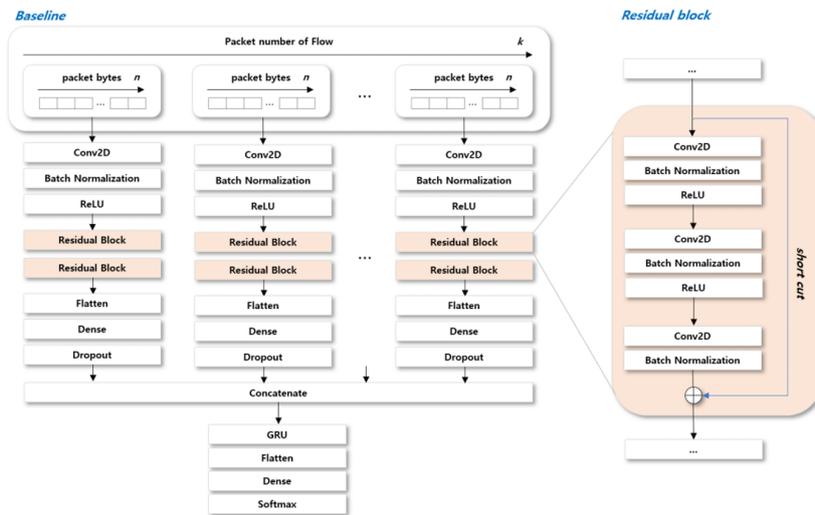
adequately addressed the consideration of input shapes derived from raw packet bytes. In light of this gap, we propose an innovative traffic classification approach that capitalizes on various input shapes derived from fixed-length packet bytes. Our proposed method begins with the assumption that there may be alternative models that better represent packets than linear or rectangular shapes. Furthermore, the combination of various perspectives looking at the same input can be expected to have a similar effect to using ensemble techniques. Our method's efficacy is assessed using the publicly accessible ISCX VPN-non VPN 2016 dataset. Impressively, our approach outperforms prior research in classification performance. This holds true for both 2D square input shapes and 1D linear input shapes.

**Goals, proposed method, novelty**

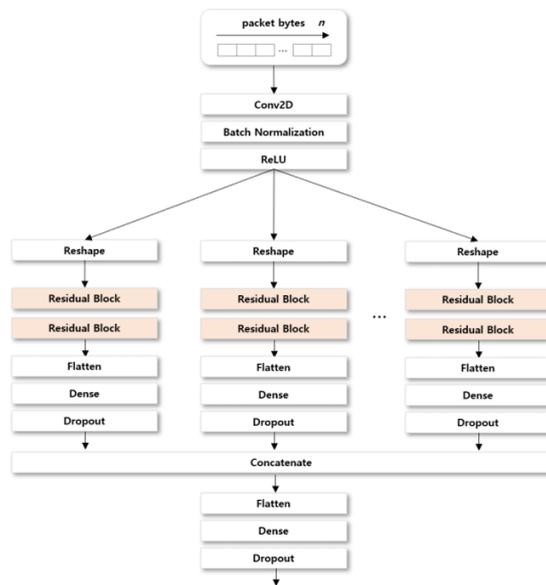
We propose a method to extract distinctive features of packets through various input shapes that can be derived from fixed-length packet bytes, which are not limited to the aforementioned square-like or linear shapes. Assuming a fixed packet length of 784 bytes, there are a total of 15 possible shapes that can be extracted from it, as detailed in Table 2. The number of shapes is equal to the number of divisors of the packet length, and since the last shapes represent the same structure as the first shapes, it is excluded. Therefore, the final count of shapes is 14, which is one less than the number of divisors.

Shape.	Kernel Size	Shape	Kernel Size
(784, 1)	(6, 1)	(16, 49)	(2, 2)
(392, 2)	(6, 1)	(14, 56)	(2, 2)
(196, 4)	(4, 1)	(8, 98)	(1, 4)
(98, 8)	(4, 1)	(4, 196)	(1, 4)
(56, 14)	(2, 2)	(2, 392)	(1, 6)
(49, 16)	(2, 2)	(1, 784)	(1, 6)
(28, 28)	(2, 2)		

**Table 2:** Shapes that can be derived from a packet with a length of 784 bytes.



**Fig.1:** Baseline



**Fig.2:** MISCNN scheme

The baseline model employed in this research comprises a blend of CNN and Gated Recurrent Unit (GRU), as depicted in Figure 1. Instead of aggregating them into a single input, utilizing each packet as an individual input ensures optimized performance. Consequently, the baseline processes multiple packets and subsequently consolidates them into a unified output at the GRU layer. To extract features from each packet-based input, a residual block is strategically positioned [7]. Each residual block is composed of three Convolution Layers, followed by the insertion of a Batch Normalization layer after every convolution operation, facilitating adjustment of output mean and variance. MISCNN is underpinned by a straightforward yet potent concept. Prior to entering the initial residual block, the input is partitioned into multiple shapes. The count of shapes for each input corresponds to the number of divisors in the packet vector size, denoted as 'n'. For instance, the frequently employed 7841 input configuration in earlier studies encompasses 15 distinct shapes (784\*1, 392\*2, 196\*4, ... , 1\*784). At this juncture, the shape 1\*784 is excluded due to its structural similarity to the shape 7\*841 (though shapes like 2\*392 and 392\*2 differ noticeably). A overview of the MISCNN architecture is illustrated in Figure 2. These reshaping techniques find limited application in the realm of image recognition and computer vision, as they compromise the 2D spatial information of neighboring (vertical, horizontal) pixels inherent in typical images. On the contrary, since raw packets lack 2D spatial information, reshaping methods can be explored, enabling the training model to perceive the same input from diverse perspectives.

### Experiments, results, analysis of the results

For evaluation, we employ the publicly accessible ISCX VPN-nonVPN 2016 dataset. This dataset, denoted as "ISCX VPN-nonVPN 2016," comprises raw pcap files featuring diverse applications, which we utilize to evaluate classification performance [6]. It encompasses human-generated traffic of varied types, coupled with information on the correlated applications. This data is garnered from both regular sessions and sessions encapsulated via VPN. This setup permits us to assign a three-view label (specifically, encapsulation, traffic type, and application) to any segmentation of raw network traffic, effectively forming a generic TC object. Each of these three-view labels corresponds to a distinct TC task that must be addressed. Each task is employed within our proposed model and subjected to comparative experiments. Packets extracted from the raw pcap files are amalgamated into flows based on the 5-tuple attributes (source IP, destination IP, source port, destination port, protocol) and reconstructed into bidirectional flows, taking directionality into account. Notably, nearly 60 percent of the aggregated two-way flows consist of only one UDP packet, leading to disruptions in the proper learning process. These instances have been filtered out, resulting in 27.8k bidirectional flows. Each flow sample contains the initial-k packets from the entire collection. When the number of packets in a flow falls below the pre-defined value k, an empty object filled with zeros is appended. The first-n bytes are extracted from the packets within each flow. If the packet size is less than the predetermined value n, the remaining vacant space within the packet object is zero-padded. The original shape of the packet data is  $k \times 1$ , which is then transformed to a  $p \times q$  shape through the Reshape layer. Details of the hardware and software environment employed during the training process are outlined in Table 3. During model compilation, the Learning Rate was set to  $25 \times 10^{-5}$ . We employed Categorical Cross-entropy as the loss function and utilized the Adam Optimizer.

	List	Specification
Hardware	CPU	Intel(R) Xeon(R) CPU E5-2630 v4 @ 2.20GHz
	GPU	NVIDIA GP102 [TITAN Xp]
	RAM	120G
Software	Nvidia Driver	440.33.01
	CuDNN	cuDNN/7.6 for cuda 10.1
	Cuda	cuda/10.1
	Python	cuda/10.1
	Keras	keras 2.4.0

**Table 3:** Shapes that can be derived from a packet with a length of 784 bytes.

In order to assess the efficacy of the proposed approach, we conducted a comparative analysis between MISCNN, four prior studies, and baseline models across three distinct TC tasks. The resulting comparison outcomes are detailed in Table 4. Notably, MISCNN exhibited superior performance compared to its predecessors. Across the three tasks, it displayed notable improvements – a 4.83% enhancement in the Encapsulation task, a surpassing of 10% in the Category task, and a commendable 7% advancement in the Application task, widely regarded as the most challenging. Remarkably, the F-measure displayed substantial growth when contrasted with previous studies. Another salient point is that the Baseline's performance also exceeded that of earlier TC models. The Baseline, enriched with the skip-connection technique that has demonstrated excellence in the domain of image recognition and computer vision, exhibited superior performance in two metrics compared to previous TC models. This substantiates the utility of the skip-connection technique within the TC domain. Lastly, it is worth noting that the 1D CNN-based TC model more effectively extracts traffic characteristics compared to its 2D CNN-based counterpart, aligning with findings from prior investigations.

	Encapsulation (%)		Traffic Type (%)		Application (%)	
	Acc	F1	Acc	F1	Acc	F1
1D-CNN [6]	87.4	83.5	73.1	71.1	72.7	61.3
1D-CNN [8]	82.3	76.2	56.0	54.7	56.5	40.8
2D-CNN [8]	87.4	83.5	71.8	69.7	71.4	59.2
Distiller [9]	93.7	91.9	80.7	78.7	77.6	66.4
Baseline(1D)	97.4	95.2	91.2	91.3	81.0	78.7
Baseline(2D)	94.1	94.0	87.2	85.2	78.0	77.6
MISCNN	98.5	98.5	93.5	93.2	85	85.2
MISCNN (gain)	4.8	6.6	12	14.4	7.3	18.8

**Table 4:** comparison of MISCNN with previous studies

## Conclusion

In this paper, we introduce a deep learning scheme aimed at comprehensively analyzing packets from multiple perspectives using diverse shapes derived from a single input. Our focus is on packet data, which lacks two-dimensional spatial information unlike typical images. We propose that reshaping packets into various forms offers a means to effectively observe and extract valuable features from raw packets. To validate our approach, we evaluate it using the publicly available ISCX VPN-nonVPN 2016 dataset, and we establish a baseline incorporating the skip-connection method. Subsequently, we implement the proposed method, termed MISCNN, atop the designed baseline, and compare its traffic classification (TC) performance against previous studies.

Our proposal contributes in three key ways. Firstly, in terms of overall comparison, MISCNN demonstrates a remarkable 14% accuracy improvement and an 18% enhancement in f-measure over state-of-the-art TC methods. This enhancement underscores that adopting diverse perspectives for packet observation mitigates overfitting in TC models and yields substantial performance gains. Secondly, performance comparisons based on experimental parameter variations reveal that employing multiple shapes, except for the Encapsulation task, contributes to accurate classification. Notably, considering that the Encapsulation task holds less significance and versatility in real-world networks compared to other tasks, integrating a broader array of shapes into the TC model can prove advantageous in common scenarios. Lastly, we perform comparative experiments within a novel dataset that addresses data imbalances, closely resembling real-world network conditions. This attests to the practical viability of the proposed MISCNN in actual network environments.

We also outline three areas for future exploration based on this study. The first pertains to reducing complexity. MISCNN, while accurate, can be excessively resource-intensive, limiting its practicality. Secondly, we suggest exploring the application of state-of-the-art techniques. Lastly, we plan to conduct research on XAI (Explainable AI) to discover improved justifications compared to existing methods.

## References

- [1] W. Wang et al., "HAST-IDS: Learning hierarchical spatial-temporal features using deep neural networks to improve intrusion detection," IEEE access, vol. 6, pp. 1792–1806, 2017.
- [2] H.-K. Lim, J.-B. Kim, J.-S. Heo, K. Kim, Y.-G. Hong, and Y.-H. Han, "Packet-based network traffic classification using deep learning," in 2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC), 2019, pp. 046–051.
- [3] J. Li and Z. Pan, "Network traffic classification based on deep learning," KSII Transactions on Internet and Information Systems (TIIS), vol. 14, no. 11, pp. 4246–4267, 2020.
- [4] X. Hu, C. Gu, and F. Wei, "CLD-Net: a network combining CNN and LSTM for internet encrypted traffic classification," Security and Communication Networks, vol. 2021, 2021.
- [5] S. Izadi, M. Ahmadi, and R. Nikbazzm, "Network traffic classification using convolutional neural network and ant-lion optimization," Computers and Electrical Engineering, vol. 101, p. 108024, 2022.
- [6] W. Wang, M. Zhu, J. Wang, X. Zeng, and Z. Yang, "End-to-end encrypted traffic classification with one-dimensional convolution neural networks," in 2017 IEEE International Conference on Intelligence and Security Informatics (ISI), Jul. 2017, pp. 43–48. doi: 10.1109/ISI.2017.8004872.
- [7] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.
- [8] S. Rezaei and X. Liu, "Multitask Learning for Network Traffic Classification," in 2020 29th International Conference on Computer Communications and Networks (ICCCN), Aug. 2020, pp. 1–9. doi: 10.1109/ICCCN49398.2020.9209652.
- [9] G. Aceto, D. Ciunzono, A. Montieri, and A. Pescapé, "DISTILLER: Encrypted traffic classification via multimodal multitask deep learning," Journal of Network and Computer Applications, vol. 183, p. 102985, 2021.

## Optimal Logistics Distribution Algorithm for Autonomous Delivery for Local Logistics Hub System

Yeong-Ho Ha, Seong-Min Kim, Hyun-Seo Jang, Young-Eun Song

The College of Science & Technology

Korea University, Korea

[haho9658@korea.ac.kr](mailto:haho9658@korea.ac.kr), [songarak27@korea.ac.kr](mailto:songarak27@korea.ac.kr), [kelvin926@korea.ac.kr](mailto:kelvin926@korea.ac.kr), [tdsong@korea.ac.kr](mailto:tdsong@korea.ac.kr)

### Abstract

With the development of e-commerce worldwide, many companies are being created, and a lot of research is being conducted on last-mile delivery. We propose e-commerce last-mile delivery as an alternative to self-driving delivery robots linked to local logistics bases. Most self-driving delivery robots can deliver only one item at a time. However, by developing a robot that can deliver multiple items, an optimal logistics allocation algorithm is needed to set up efficient delivery and delivery routes for delivery robots. This paper presents an efficient delivery method using VRPTW (Vehicle Routing Problem with Time Window) when autonomous robots deliver goods from logistics bases in various forms.

**Topics:** planning and scheduling, robotics, machine learning

### Introduction and background

As the global e-commerce market grows, research on the last mile of the delivery chain is becoming more active. Last-mile delivery, the final part of the delivery process, "involves a series of activities and processes necessary for the delivery process from the last transit point to the final drop point of the delivery chain" [1].

In this paper, we designate the last transport point in the delivery chain as a local logistics base, and the process from the local logistics base to the final drop point is proposed as an autonomous delivery robot.

According to a study conducted in Korea, the delivery time, excluding travel time, of the time when a courier driver loaded each delivery item into a car and delivered the entire delivery quantity for the day occurred from 26% to 32% of the total delivery time [2].

To solve the delivery time for logistics delivery it is solved by using an autonomous delivery robot linked to a local logistics base. A local logistics base is a delivery robot that delivers goods to the door by creating a base for each region instead of the previous delivery method, the last transportation point of the existing delivery chain. These delivery robots can deliver efficiently through route settings.

Routing of delivery robots was first proposed by Dantzing et al. in 1959 [3] as the Vehicle Routing Problem (VRP) and is based on VRP. VRP applies various constraints to the Traveling Salesman Problem (TSP), among which VRPTW (Vehicle Routing Problem with Time Window) is applied. VRPTW considers the total quantity of goods the vehicle carries and its capacity. Due to time constraints, it can apply the current courier delivery system, which requires all deliveries to be completed in one day. Therefore, this paper aims to solve the problem by applying VRPTW to various delivery destination types.

### Goals, proposed method, novelty

To apply VRPTW, delivery must be completed within a fixed time through time constraints in basic VRP.

A. Gupta and S. Saini[4] make four assumptions to construct VRPTW.

- 1) A client must be served only once.
- 2) Each vehicle's route must start and arrive at the garage.
- 3) The sum of demand for all vehicle routes must not exceed the maximum capacity.
- 4) The client is delivered within the given time.

The main objective of VRPTW is to minimize the total cost of travel.

The basic VRPTW is applied based on the vehicle. Our self-driving delivery robot can move freely indoors and outdoors. To apply such self-driving delivery robots, there are various considerations. Critical situations cannot be considered for a delivery robot with smaller wheels than a vehicle, such as road pavement and inclination angle.

Our goal is to set the optimal path for robots in these situations. It is to set the optimal route considering all conditions, such as the shortest route from the inside, pavement condition of the road from the outside, angle of inclination, and obstacles.

### Experiments, results, analysis of the results

To apply the VRPTW described above, we applied it to various forms. It can be divided into situations that require multiple deliveries on one floor, such as large offices and large hospitals, and problems that require various deliveries on multiple floors.

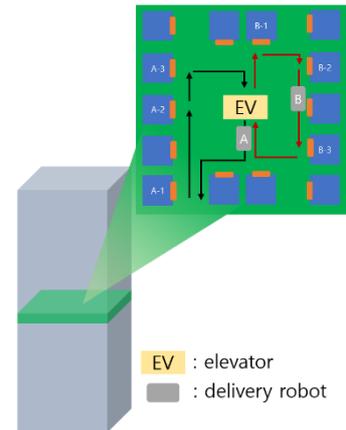


Fig.1: Many delivery destinations must be considered on one floor

In the case of large offices and large hospitals, as shown in Fig.1, many delivery destinations must be considered on one floor. In this form, priority should be given to the efficient movement of the robot by prioritizing the goods to the nearest delivery destination.

When multiple deliveries must be made to various floors, as shown in Fig.2, adjacent floors must be considered.

Also, in common, the maximum loading number should be considered. Although the number of loads is large, costs must be regarded in remote areas, and although the size and weight of delivery boxes vary, the capacity of the robot is limited. The load weight is also fixed, so this must also be considered.

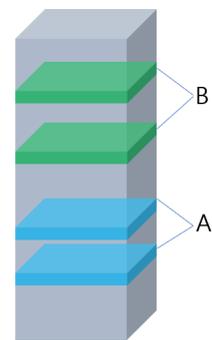


Fig.2: Multiple deliveries must be made to various floors

### Conclusion

In this paper, VRPTW is applied to various delivery destinations for efficient delivery to logistics bases and self-driving delivery robots capable of carrying multiple loads in one delivery while reducing the delivery time that takes up much time in parcel delivery from courier drivers to local logistics bases. Delivery and delivery by self-driving robots to each destination from local logistics bases can reduce the overall delivery time of couriers.

### References

- [1] Lindner, J. "Last mile logistics capability: A multidimensional system requirements analysis for a general modelling and evaluation approach." *Technical University of Munich*, 2011
- [2] Park, Sung-Mee, and Gee-Ju Moon. "A Study on the Walking Time to Drop off Parcels on the Design of VRP Heuristic for Parcel Delivery Services." *Journal of the Society of Korea Industrial and Systems Engineering* 35.2 , 189-195, 2012
- [3] Dantzig, G. B. & Ramser, J. H., "The truck dispatching problem," *Management Science*, Vol. 6, No.1, 80-91, 1959
- [4] A. Gupta & S. Saini, "An Enhanced Ant Colony Optimization Algorithm for Vehicle Routing Problem with Time Windows," *2017 Ninth International Conference on Advanced Computing (ICoAC)*, 267-274, 2017



# **SESSION W2: ORAL PRESENTATIONS**

## **CHAIRS:**

**PROF. HUN-KEE KIM, INHA UNIVERSITY  
PROF. KENJI HASHIMOTO, WASEDA  
UNIVERSITY**

**17 NOV 2023 (FRI.) 15:05-16:45**

**ROOM S153-154**



## Robot Car Motion Control Leveraging Hand Gesture Recognition with MediaPipe

Jianan Xie, Yue Yi, and Kenji Hashimoto

Graduate School of Information, Production and Systems  
Waseda University  
Japan

[xiejn@akane.waseda.jp](mailto:xiejn@akane.waseda.jp), [ekiesi123@moegi.waseda.jp](mailto:ekiesi123@moegi.waseda.jp), [kenji.hashimoto@waseda.jp](mailto:kenji.hashimoto@waseda.jp)

### Abstract

Currently, many studies aim to improve the control methods of robots. Compared to traditional methods such as mouse, keyboard and remote controller, visual gesture-based methods allow users to interact more naturally with robots. To achieve robot control, many researchers use neural network-based methods to recognize specific gestures, such as numbers, letters, etc. However, specific gestures only possess specific semantic information, losing the rich movement information of the hand. Robots cannot achieve complex movements through specific gestures. In this study, we leverage a real-time hand-tracking technology provided by MediaPipe to introduce a novel gesture-based method for controlling the motion of a robotic car. Specifically, the pitch and rotation of the palm are analyzed to facilitate forward/backward movement and turning of the vehicle.

**Topics:** robotics, gesture recognition, motion control, Gesture-based control

### Introduction

In the realm of operating robots, a wide array of methods have been extensively adopted, including physical remote controllers, game controllers [1], and control based on a Graphic User Interface (GUI) using a mouse and keyboard. However, in outdoor or dynamic settings, these physical control methods may not be the best as they need to be carried at all times and controlled with both hands. Given the ubiquity and innate portability of smartphones, utilizing the cameras to capture and recognize gestures for one-handed control of robots presents significant practical value.

While gesture recognition technology has made significant strides across various domains, it still faces technical challenges in aspects such as illumination variations, complex backgrounds, hand occlusions, camera depth perception, and performance requirements. MediaPipe Hands [2] offers a real-time solution suitable for mobile devices, characterized by its independence from specific hardware and its ease of deployment. This study primarily explores the use of the MediaPipe Hands solution to generate motion commands for robotic cars through gesture recognition.

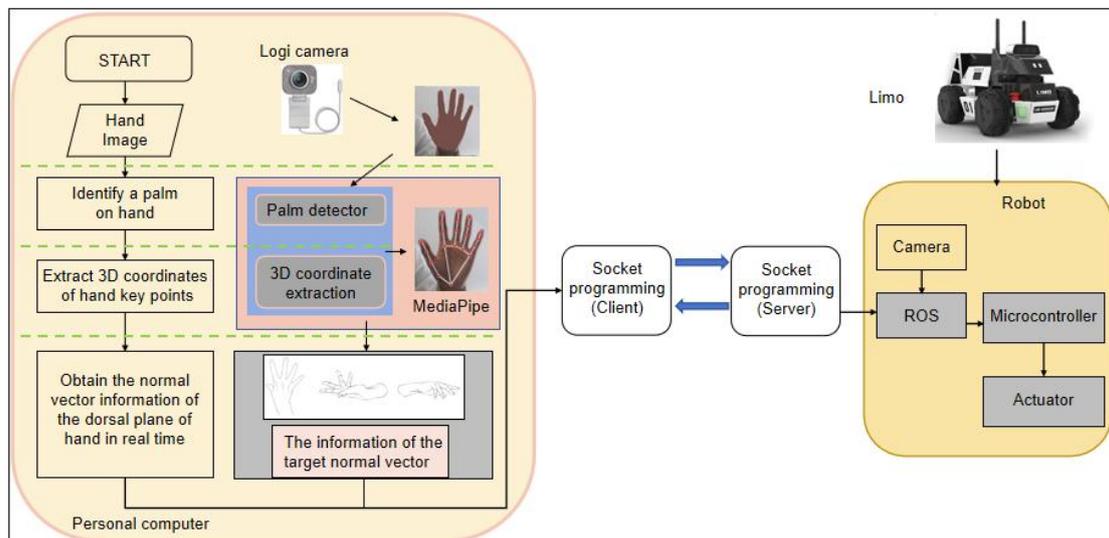


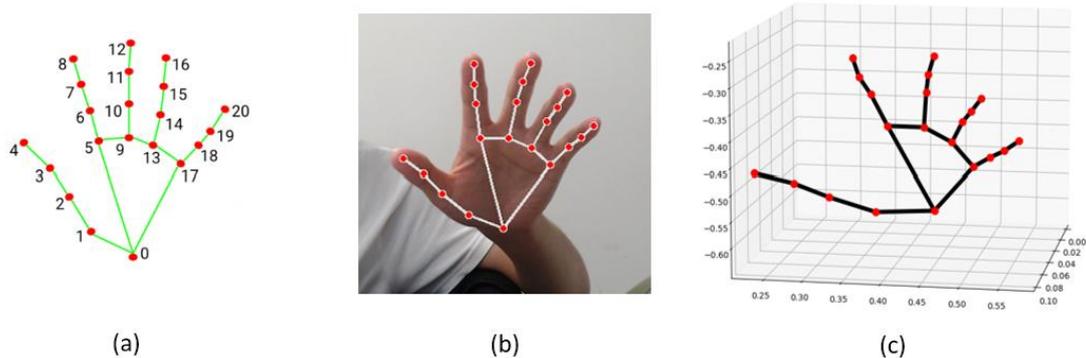
Fig.1 Overview of gesture-based control system

## Method

MediaPipe Hands presents a real-time on-device hand-tracking solution that can predict the hand skeleton of a human through a single RGB camera. It consists of two models.

- A palm detector that functions on a complete input image and identifies palms using an oriented hand bounding box.
- A hand landmark model that functions on the cropped hand bounding box supplied by the palm detector and returns high-definition 2.5D landmarks.

Through the above two models, it is possible to accurately locate the hand landmarks based on input images. The crux of this study lies in utilizing these spatial coordinates to extract and interpret motion cues.

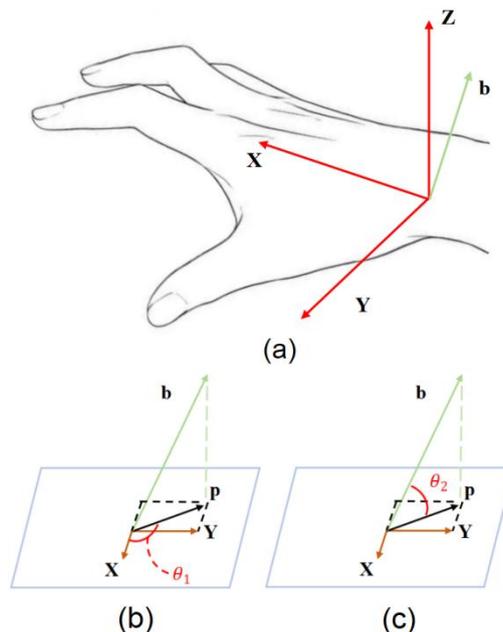


**Fig.2** Gesture recognized by MediaPipe Hands

As illustrated in Fig.2, Fig.2 (a) reveals the indexing of hand landmarks in the recognition results, enabling a clear correspondence between the physical joints and the data. Fig.2 (b) displays the skeletal diagram derived from the recognition data, and Fig.2 (c) presents the three-dimensional joint diagram constructed based on the data from (b).

Since the unique virtual camera-based coordinate system is used in the design of the MediaPipe algorithm [3], we need to use the rotation matrix to transform the hand landmark coordinates in this coordinate system into the coordinate system commonly used in robot control.

With the coordinate information of hand landmarks, specific gestures such as numbers from 1 to 5, a clenched fist, and the letter V can be accurately identified following a set of rules. In this study, we adopt an intuitive approach: controlling the vehicle's movement based on the normal vector of the dorsum of the hand.



**Fig.3** Illustration of how normal vector on the back of the hand takes effect on car motion

As illustrated in Fig.3, These vectors provide us with an indication. Fig.3 (a) represents the palm, the Cartesian coordinate system and the normal vector on the back of the hand  $b$  in the space. When the palm rotates around the center of the palm, Fig.3 (b) dictates the direction of the car (Vector  $p$  is the projection of vector  $b$  on the horizontal plane). We can conveniently adjust  $\theta_1$  to varying between 0 and 359 degrees. As the palm tilts upwards or downwards, the parameter  $\theta_2$  shown in Fig.3 (c) increases or decreases respectively. A

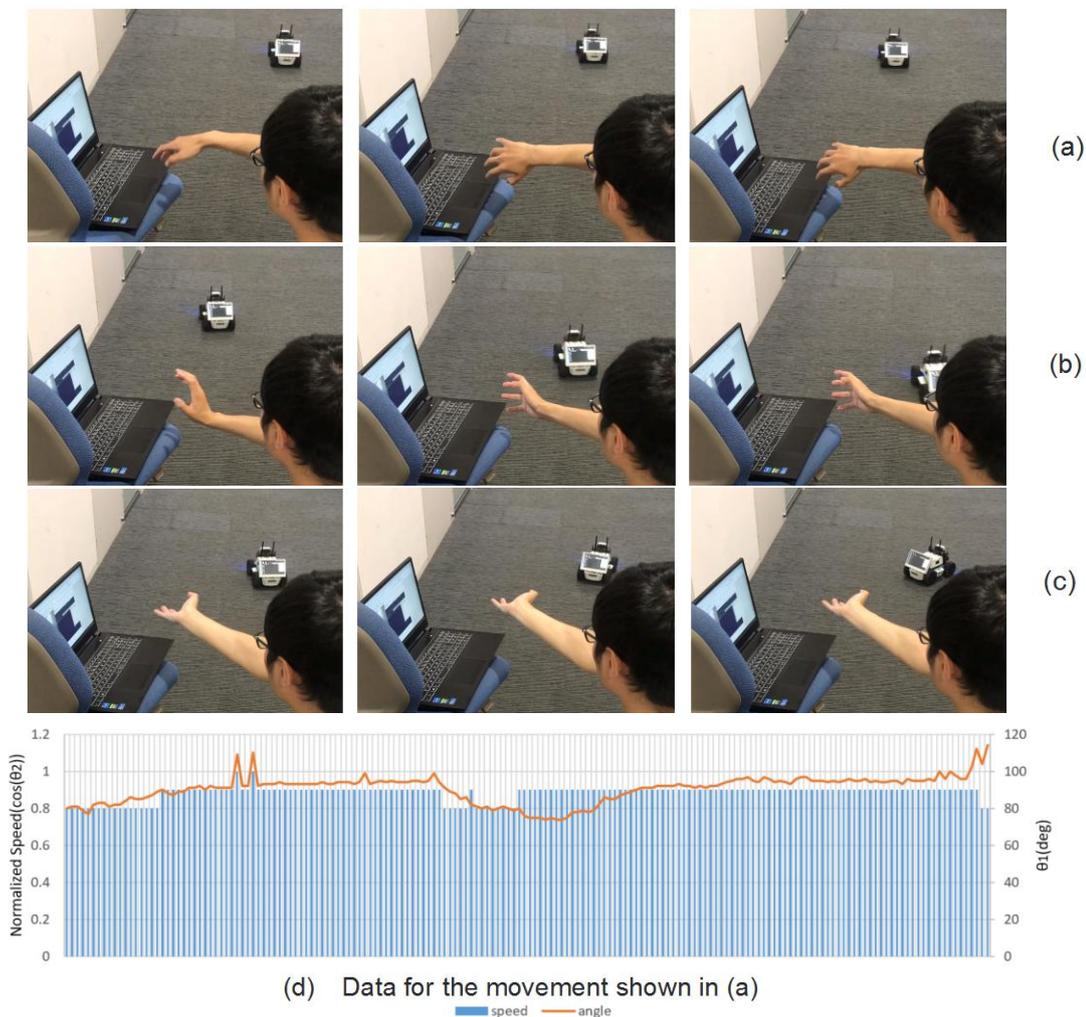
smaller value of  $\theta_2$  results in a higher absolute speed of the vehicle. If the z-value of the normal vector is negative, it means that the car will rotate in place.

Given the challenges of detecting the palm when it's perpendicular to the camera, achieving  $\theta_2$  of 90 degrees to stop the vehicle becomes problematic. To address this, we introduced a strategy: the vehicle halts its movement when a fist gesture is detected by the system. This detection primarily relies on the angles between the joints of the fingers when forming a fist.

## Experiment

In this experiment, we utilize a robotic car named LIMO [4] and set the robot LIMO to Mecanum wheel mode. By capturing images using OpenCV, processing them with MediaPipe Hands, and then transmitting motion commands to the robotic car via a TCP socket, we successfully achieved the ability to control the car's movement through the pitch and rotation of the hand. We managed the speed and direction of the vehicle through a camera and its corresponding code.

Fig.4 (a) illustrates the hand gesture and resultant motion of the car moving laterally from right to left. Fig.4 (b) displays the gesture required for moving the car diagonally backward to the right. Fig.4 (c) shows the rotation of the car. Fig.4 (d) shows the variations of speed and angle with respect to time in Fig.4 (a). The complete implementation and demo video can be accessed at: <https://github.com/include-yy/simple-gesture-control-small-car>.



**Fig.4** Actual gesture-controlled movement

## Conclusion

This paper implements a novel gesture-based control method and applies it to robotic cars, making human-computer interaction simpler and more natural. Future research could focus on applying the gesture-based control method to control robot systems composed of multiple units and improving the algorithms to make them more suitable for specific application environments, such as well-lit areas or dark rooms. Additionally, integrating other human-machine interaction methods, such as voice interaction, could lead to a comprehensive control system.

## References

- [1] Long, S. K., & Bliss, J. P. The Effect of Control Device on Performance in a Robotic Arm Task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 60(1), 795–799.(2016)
- [2] Zhang, Fan, et al. "Mediapipe hands: On-device real-time hand tracking." *arXiv preprint arXiv:2006.10214* (2020).
- [3] Kanstantsin Sokal. MediaPipe 3D Face Transform. Google for Developers. (2020).  
<https://developers.googleblog.com/2020/09/mediapipe-3d-face-transform/>.
- [4] Limo - Agilex Robotics. <https://global.agilex.ai/products/limo/>.

# A Dynamic Human Following Method Based on Ultra-Wideband System for Mobile Robots: Application and Strategies in Librarian Assistance

Yun Kang, Taehoon Kim, Chengyan Gu, Jinchul Roh and Joono Cheong

Department of Control and Instrumentation Engineering

Korea University, Republic of Korea

terry771@korea.ac.kr, kim\_taecheon@korea.ac.kr, guchengyan@korea.ac.kr, jcroh97@korea.ac.kr, jncheong@korea.ac.kr

## Abstract

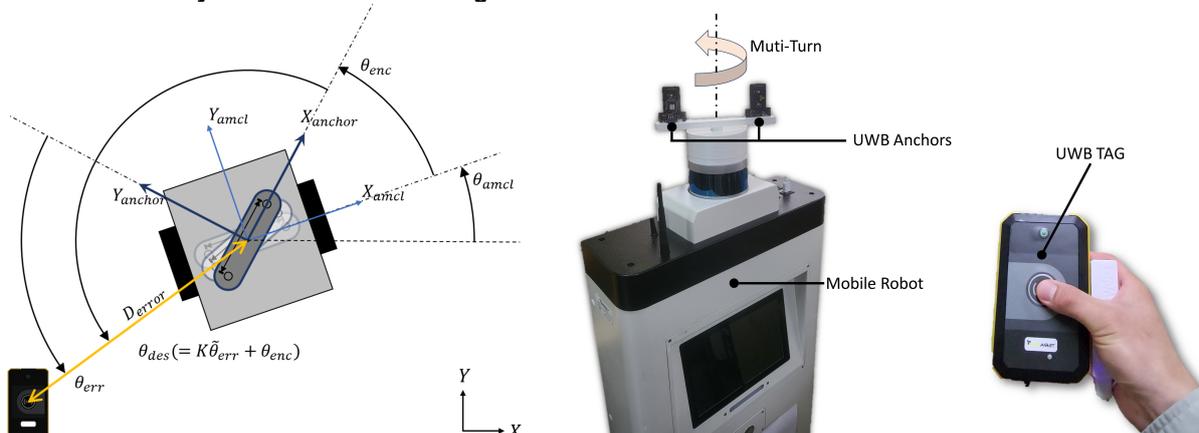
This study introduces a human following method for mobile robots using ultra-wideband (UWB) system and its application in indoor navigation for librarian assistance. In the UWB system, it comprises a set of two UWB anchors (receivers) and a UWB tag (transmitter). A set of anchors was installed on top of a mobile robot to track the heading angle and distance of a tag. To achieve an unrestricted tracking range, the anchors were integrated with a motorized dynamic motion. The tag will function as a remote carried by an individual and regarded as a continuous goal pose for the robot. Our human-following capability spans a comprehensive range from  $-180^\circ$  to  $180^\circ$ , covering a maximum distance of 60m. This following system has been implemented for librarian assistance in a local library (Sejong city), utilizing the TEB (Timed Elastic Band) algorithm for conducting obstacle avoidance and human following simultaneously. Beyond librarian assistance, autonomous navigation within the library has been achieved through simultaneous localization and mapping.

**Topics:** Robotics, Mechatronics, Sensors, Signal processing, Wireless communication

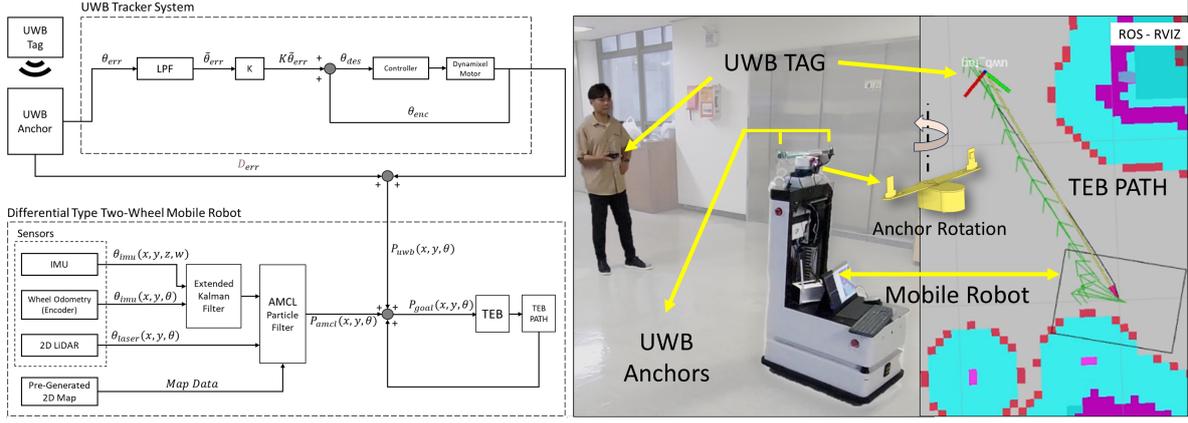
## 1. Introduction

Due to the advances in indoor mobile robot services, the capability of human following for mobile robots is becoming an indispensable element.[1] A typical example is autonomous carts in retail environments. These systems enhance the shopping experience by reducing the physical burden on customers and pave the way for natural indoor human-robot interactions[2][3][4] As human-robot interaction (HRI) becomes increasingly evolving in daily activities[5][6], ensuring smooth, intuitive, and safe interactions is needed. In light of this, we identify several areas of improvement for current human-following methods as follow. First, cameras, which are conventionally used in human-following, present limitations in tracking range and distance of the target. Second, the primary sensors used for human-following, such as cameras and LiDAR, are susceptible to disturbances caused by reflections or changes in light intensity. Third, for RGB cameras and LiDAR sensors in human-following, unexpected intervention of obstacles between human and robot can lead to potential system disruptions and suddenly fail to capture significant data. Through our UWB based dynamic human following method, we aim to address these challenges and make a significant contribution to enhancing human-robot interactions.

## 2. UWB-Based Dynamic Human Following Method



**Fig.1:** Left: Schematic diagram of UWB based dynamic human following method. Right: Image of the robot and a set of UWB based dynamic human following system.



**Fig.2:** Left: Block diagram of UWB based human following method. Right: Captured image while TEB path planning while conducting UWB based dynamic human following.

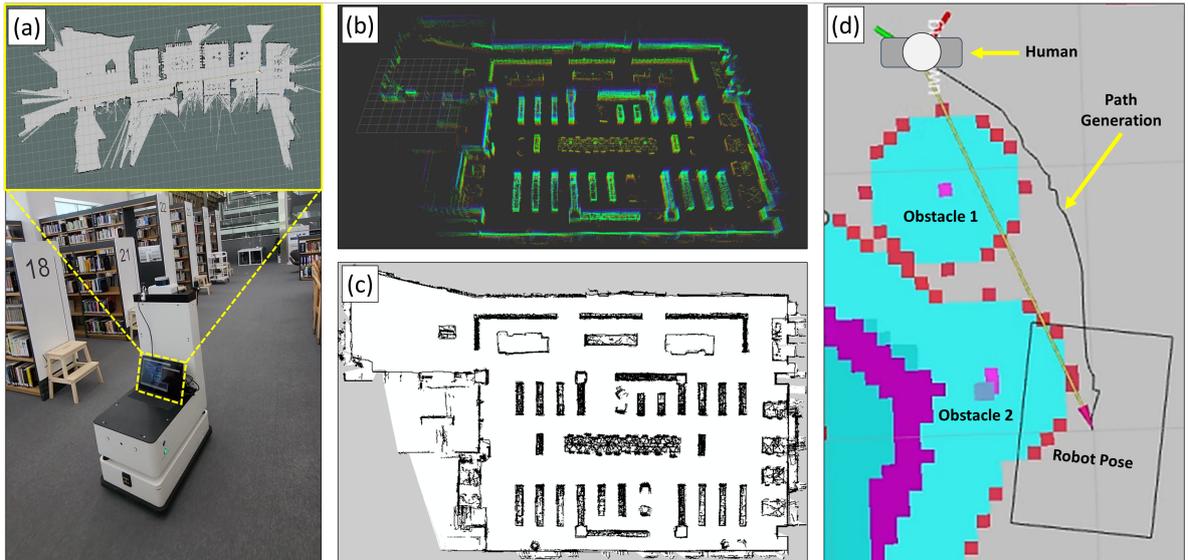
In this section, we introduce the UWB-based dynamic human-following method. As illustrated in figures 1 and 2, two UWB anchors (Terabee, CH) are securely affixed to the rotational axis of the motor (Dynamixel XH-430-210-R, Robotis, KR) on the top of the robot (Nargo60, Twinny, KR). As the target UWB tag moves, the anchor follows the tag through the rotation of the motor. The anchors provide the tag's heading angle ( $\theta_{err}$ ) to the proposed UWB tracker system, resulting in a motor input position as expressed in the following.

$$\theta_{des} = K\bar{\theta}_{err} + \theta_{enc} \quad (1)$$

Simultaneously, the motor's encoder ( $\theta_{enc}$ ) and the distance ( $D_{err}$ ) data from the UWB anchor are employed in the subsequent process. Initially, as delineated in the block diagram of Figure 2, the mobile robot's IMU (Inertial Measurement Unit) and wheel odometry (encoder) data are processed through an EKF (Extended Kalman Filter). In conjunction with LiDAR, map, and with previous EKF filtered data, it undergoes an AMCL (Adaptive Monte Carlo Localization) filtering process. From the resultant map, the robot's anticipated current position ( $P_{amcl}(x, y, \theta)$ ) combined with  $\theta_{enc}$  and  $D_{err}$  generates the goal position ( $P_{goal}(x, y, \theta)$ ). This goal position then guides local planning to the target point via the TEB (Timed Elastic Band) algorithm[7].

### 3. SLAM, Navigation, and Human Following Application in Library Environment

In this section, we discuss the applications and strategies for librarian assistance, incorporating not only the UWB-based dynamic human-following system but also the in-door navigation through construction of 2D/3D maps local library via SLAM (fig.3(a)).



**Fig.3:** (a) Mobile robot for 2D SLAM and navigation. (b) 3D map generated by 3D LiDAR manually. (c) 2D map generated by processing 3D map. (d) Obstacle avoidance while human following.

These maps (Fig.3(a), (b)) facilitate the indoor navigation, allowing the robot to transport books along predefined paths in the library during routine operations. Upon request from a librarian (triggered by pressing a button on the UWB tag), the robot is to move the requested position by global planning with real-time collision avoidance (fig.3(d)). Once set, as the librarian starts to move, the robot then follows him/her. In a *follow mode*,

the robot follows the librarian maintaining a safe distance of 1.8 meters. In a *wait mode*, the robot stands still until the librarian signals to follow. This mode is used for the librarian to collect books from shelves and to stack them up on the rack of the robot. The transition between modes was facilitated through an integrated mobile application in conjunction with a smart Wi-Fi switch module (Smart Wi-Fi RF relay switch, eMylo, CN), allowing librarians to seamlessly initiate the change via a single tactile interaction on their smartphone device. When a user interacts with a smartphone interface via tactile input, the embedded system transmits the corresponding data to a designated server. This data is subsequently relayed to a mobile robot, facilitating seamless mode transitions from any location at any time. The whole task software is being enhanced for finer and more seamless librarian helping services in a broader scale.

Consequently, it was ascertained that using the UWB-based dynamic human-following method, the robot could continuously and real-time track a human within a 360 degree range and up to a distance of 60 meters. To elucidate this validation further, we provide a demonstration video[8] of the proposed UWB-based dynamic human-following method.

## Conclusion

Through this study, we gained insights summarized in following.

(i) The proposed human following method enables a comprehensive range of following (from  $-180^\circ$  to  $180^\circ$ , within 60m) by integrating a UWB sensor system with a motorized tracking platform. Therefore, it effectively mitigates restrictions related to the tracking range and distance of the target.

(ii) Because of the characteristic of high bandwidth UWB communication, the system is largely impervious to interferences from reflections or variances in environmental light condition.

(iii) Due to the inherent extremely narrow pulses and a very broad frequency bandwidth of UWB signals, the system allows the signals to penetrate most objects with ease. This leads to the simultaneous human following, and the real-time local planning can give a dependable navigation in cluttered environments like public libraries or open spaces.

We believe our method proved its potential in indoor autonomous navigation, human following, and a broader range of human-robot interaction.

## References

- [1] Aine Ilina Tarmizi, Ahmad Zaki Shukor, Nur Maisarah Mohd Sobran and Muhammad Herman Jamaluddin, Latest Trend in Person Following Robot Control Algorithm: A Review, *J. of telecommunication, electronics and computer engineering.*, vol. 9, no. 3, p. 169—174, 2017.
- [2] Yen Leng Ng, Cheng Siong Lim, Kumeresan A. Danapalasingam, Michael Loong Peng Tan and Chee Wei Tan, Automatic human guided shopping trolley with smart shopping system, *J. of Teknologi.*, vol. 73, no. 3, p. 49—56, 2019.
- [3] Soh Nishimura, Hiroshi Takemura and Hiroshi Mizoguchi, Development of attachable modules for robotizing daily items -Person following shopping cart robot-, *Int. Conf. on IEEE International Conference on Robotics and Biomimetics (ROBIO).*, p. 1506—1511, 2007.
- [4] Ryumin, Dmitry, Ildar Kagirov, Alexandr Axyonov, Nikita Pavlyuk, Anton Saveliev, Irina Kipyatkova, Milos Zelezny, Iosif Mporas and Alexey Karpov, A Multimodal User Interface for an Assistive Robotic Shopping Cart, *Electronics.*, vol. 9, no.12, p. 2093, 2020.
- [5] Diego Rodríguez-Guerra, Gorka Sorrosal, Itziar Cabanes, Carlos Calleja, A Multimodal User Interface for an Assistive Robotic Shopping Cart, *IEEE Access.*, vol. 9, p. 108557-108578, 2021.
- [6] Yun Kang and Joono Cheong, Smart Wrist Band Considering Wrist Skin Curvature Variation for Real-Time Hand Gesture Recognition, *J. of Korea Robotics Society.*, vol. 18, no. 1, p. 18-28, 2023.
- [7] Rösmann, Christoph, Frank Hoffmann and Torsten Bertram, Integrated online trajectory planning and optimization in distinctive topologies, *Robotics and Autonomous Systems.*, vol. 88, p. 142-153, 2017.
- [8] Yun Kang and Joono Cheong, Demonstration: A Dynamic Human Following Method Based on Ultra-Wideband System for Mobile Robots, [Online], <https://youtu.be/SpNzBmk8w6M>, Accessed: Sept. 23, 2023.



# An Advanced Motor-less Walking Assist Harnesses Power by Integrated Dual-Slider Lever Mechanism

Xiuyuan Wu, Keisuke Osawa, Yuntian Wang and Eiichiro Tanaka

Graduate School of Information, Production and Systems / Faculty of Science and Engineering / Tanaka Laboratory  
Waseda University / Japan  
tanakae@waseda.jp

## Abstract

This study focuses on designing a mechanical walking assist system to address mobility issues faced by individuals with walking disorders. This system aims to facilitate foot flexion and extension while promoting a healthier walking pattern. It employs an offset slider crank mechanism to lift the foot and achieve dorsiflexion (DF), along with a four-bar linkage that assists foot kicking, plantar flexion (PF), by utilizing the action-reaction force from lever kicks against the ground. The system incorporates forefoot mechanical pressure and heel sensors for triggering and releasing the lever, respectively. Experimental results demonstrate successful assistance for both PF, and DF, leading to reduced muscle activity and promoting safer and healthier walking. This innovative apparatus holds the potential to significantly improve mobility and quality of life for individuals with walking disorders and impairments. Moreover, it could find applications in other domains requiring mechanical assistance, thereby contributing to cost-effective walking rehabilitation and serving as a valuable reference for experimental configurations and future research endeavors.

**Topics:** Machine design

## Introduction and background

The global concern over aging issues has escalated in countries similar to Japan with declining birthrates [1]. Recognizing the significance of moderate daily exercise in reducing life-related diseases, walking has been established as a pivotal workout associated with mortality rates [1]. However, muscle deterioration around the age of 65 mostly leads to a decline in walking pace and posture irregularities, amplifying the risk of hypertension, cerebral infraction, and Parkinson's disease [1]. An indispensable strategy for enhancing the quality of life for both senior individuals and those with partial muscle functions is facilitating both PF and DF of the foot, which is critical for restoring a healthy gait cycle and reducing mortality.

Existing solutions to address foot flexion often involve motorized motion and tailored to individual's body conditions [2,3,4]. Yet, these solutions encounter challenges related to limited force output and weight concern. To overcome these limitations, we have developed a prototype motor-less walking assist harnessing user's bodyweight by using ratchet-pawl mechanism integrating bilateral assistances [6,7,8], yet it met few defects on utilizing the energy, concerning the compact and effectiveness, an advanced motor-less assist with integrated dual-slider lever mechanism to assist the PF and DF was proposed, thus assisting walking.

## Goals, proposed method, novelty

The main aim of this study is to create a motor-free walking aid for the lower limbs. It operates using a combination of sliders and a linkage system to facilitate foot movement during walking (Fig. 1). Springs are employed to drive the linkage system, aiding in kicking the ground and raising the foot (Fig. 2). The concept employs the force generated by springs and levers to propel the foot forward during push-off when the heel leaves the ground. Additionally, an offset slider crank mechanism, coupled with a compression spring, is utilized to ensure a natural walking posture by actively lifting the foot.

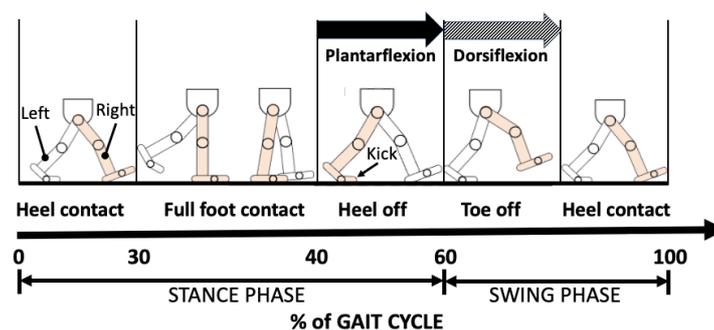


Fig. 1 Gait cycle decomposition

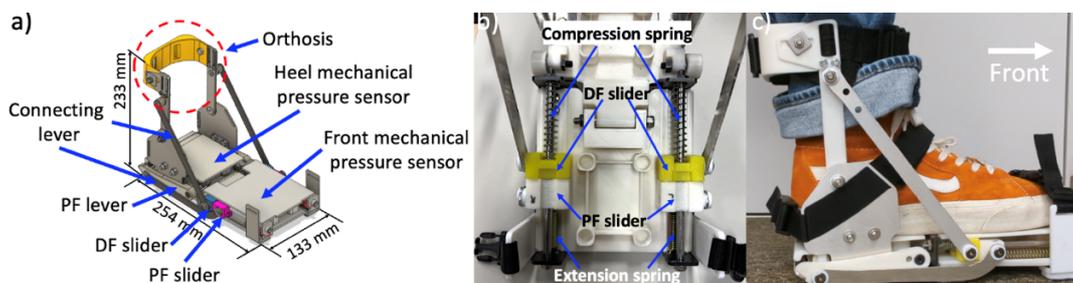


Fig. 2 Motor-less walking assist overview

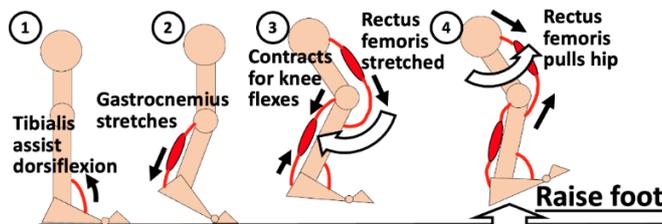


Fig. 3 Stretch reflex concept for advanced walking assistance

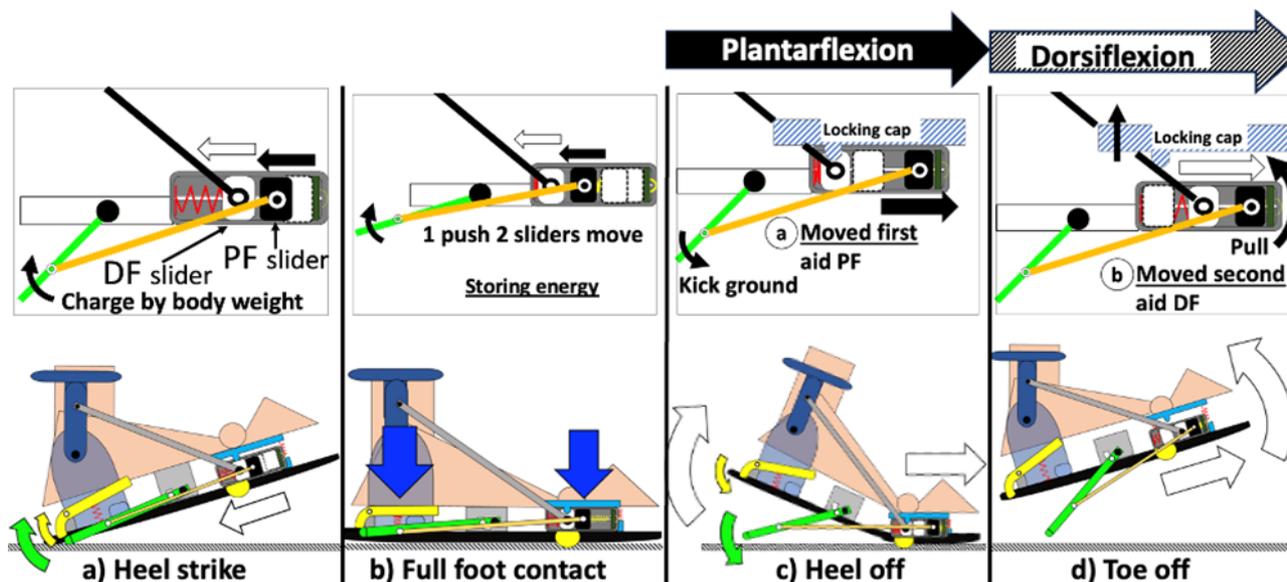


Fig. 4 Mechanism of the assistance

The device allows for customization of the springs to achieve optimal assistance with varying spring stiffness and deformation length. This novel motor-less walking assist integrates both foot raising and kicking assistance into one and focuses on assisting only ankle joint for gait training purpose with the concept of stretch reflex (Fig.3) [5]. Integrating sliders and levers under the forefoot resulting in compact, higher efficiency via adjusting springs' deformation length, and lighter weight by 3D printing with polylactic acid (PLA) and aluminum A5052 for higher structural stiffness. Figure 4 shows the mechanism how to assist the ankle motion.

### Experiments, results, analysis of the results

Overall, 3 participants with close body conditions, height, weight, and gender, were chosen for this experiment. The content included muscle activity variations of tibialis anterior (TA) and gastrocnemius lateral head (GLH) measured by electromyograph (EMG) and ankle angle variation by motion capture. All participants had been instructed to walk on the treadmill at the given speed (0.8 [km/h]) for 20 seconds and rest in 3 minutes between each session. Both data were acquired at the same time during the experiment with participant self-determined beat per minute from without wearing the device for retaining a solid experimental baseline. For participants, they had been instructed to walk on the treadmill sequentially from softer spring to stiffer spring installed to the device, by practicing the experimental protocols, the changing in muscle activity and ankle angle variations between springs and without device could be observed for efficiency verification.

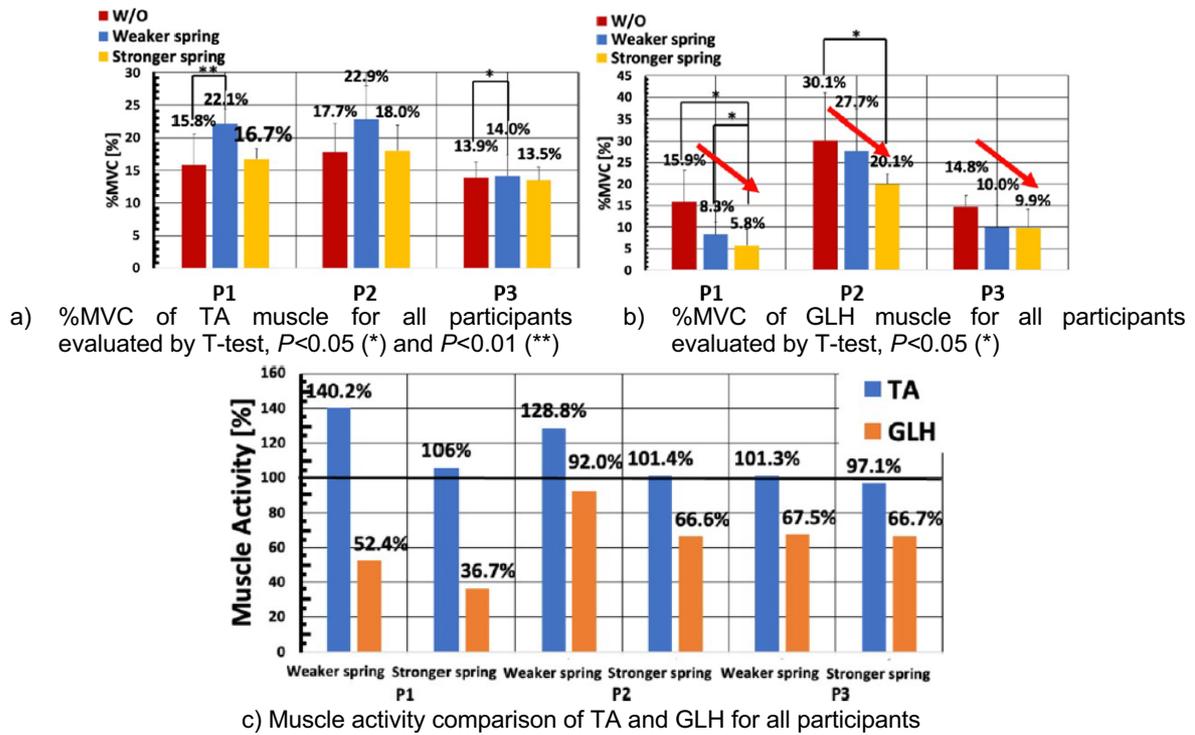


Fig. 5 %MVC (maximum voluntary contraction) and muscle activity comparison of TA and GLH for all participants

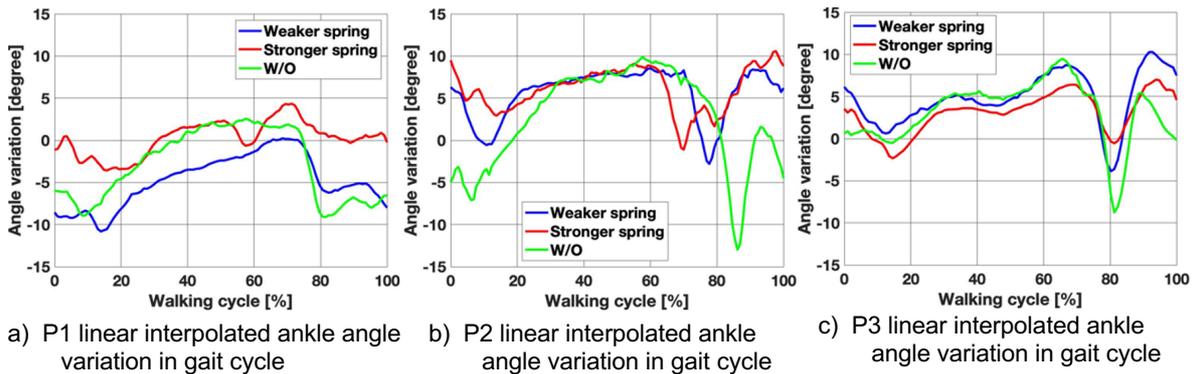


Fig. 6 Ankle angle variations comparison for all participants, trajectories indicate in increasing angle in both PF (40-60%) and DF (60-75%)

Figure 5 shows the muscle activity and ankle angle variations for all participants [9]. Both participants received full assistance to the GLH muscle under weaker and stronger springs, leading to a significant drop-in activity with the stronger spring. The TA muscle, however, received only subtle assistance due to installed steel parts collecting feedback. It also showed reduced activity with the stronger spring due to increased inertia from plantarflexion assistance [9]. The apparatus effectively improved the gait cycle and posture, enabling users to walk with less energy expenditure, safer posture, and appropriate stride length. The T-test confirmed the reproducibility of the EMG results in this study. Figure 6 displays the measured angle variation results for each participant. Dorsiflexion was consistently larger with both springs [10]. The stronger spring provided adequate force, making walking easier and faster. Both springs guided gait posture. Phase transformation from stance to swing was restricted (60% to 80%, mid-swing 73% to 87%) [10]. The weaker spring guided posture, while the stronger one fully assisted motion, increasing ankle force and walking speed. Dorsiflexion was assisted regardless of spring, thanks to the apparatus's mechanism. The weaker spring suits able-bodied individuals, guiding posture, while the stronger one benefits seniors or patients with less strength, providing full assistance and improving gait cycle, resulting in a safer posture and larger stride length.

### Conclusion

A motorless walking assistance device was developed, using an integrated slider and linkage mechanism to promote a safer and healthier walking posture. It assists both able-bodied individuals and those with deformities, improving the gait cycle and potentially enhancing walking speed and posture. Customizable assistive moments on the ankle can be achieved by replacing the springs. Comprehensive motion capture and electromyography experiments confirm improvements. The design minimizes stumbling risk on uneven terrains for individuals with

limited muscle strength. Future efforts focus on streamlining the design and enhancing flexion assistance with a user-friendly approach.

## References

- [1] National Health Promotion Campaign, Japan: Ministry of Health, Labour and Welfare, 2022, pp. 1–26.
- [2] Yong-Lae Park, Bor-rong Chen, Néstor O Pérez-Arancibia, Diana Young, Leia Stirling, Robert J Wood, Eugene C Goldfield, Radhika Nagpal, Design and control of a bio-inspired soft wearable robotic device for ankle-foot rehabilitation, *Bioinspiration & Biomimetics*, vol. 9, no. 1, p. 016007, 2014.
- [3] Marta Moltedo, Tomislav Baček, Ben Serrien, Kevin Langlois, Bram Vanderborght, Dirk Lefeber, Carlos Rodriguez-Guerrero, Walking with a powered ankle-foot orthosis: the effects of actuation timing and stiffness level on healthy users, *Journal of NeuroEngineering and Rehabilitation*, vol. 17, no. 1, Jul. 2020.
- [4] Kei Nakagawa, Naoto Kadono, Tomohiro Shimoda, Takafumi Mitsuhashi, Eiichiro Tanaka, Louis Yuge, Intramuscular coherence of the lower flexor muscles during robotic ankle-assisted gait, *Journal of Motor Behavior*, Vol. 54, Issue 3, pp. 344-353, 2021.
- [5] Eiichiro Tanaka, Ryosuke Niwa, Kazuki Osawa, Keyaki Nakajima, Keiichi Muramatsu, Keiichi Watanuki, Shozo Saegusa, Louis Yuge, Motion assistance apparatus enabled for neuro-rehabilitation of patients and for the promotion of exercise for the elderly, *IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM2015)*, pp. 937 - 942
- [6] E.Tanaka, A. Ide, J. Zhou, J. Zhuang, K. Osawa, Development of motor-less and small type of a walking assistive device, *Robomech2022*, 2A1-E10, (2022).(In Japanese)
- [7] X.Wu, A. Ide, K. Osawa, E. Tanaka, Full mechanical walking assistive device aims to offer normal gait postures due to unhealthy gait, *MIPE2022*, A1-1-02, (2022).
- [8] Xiuyuan WU, Asagi IDE, Keisuke OSAWA, Kei NAKAGAWA, Louis YUGE, Eiichiro TANAKA, Development of a motorless walking assistive device for foot flexions with instant torque output in gait cycle, *Journal of Advanced Mechanical Design, Systems, and Manufacturing*, Vol. 17, No. 4, Paper No.23-00175, (2023).
- [9] Xiuyuan Wu, Keisuke Osawa, Eiichiro Tanaka, A Motorless Walking Assistance Powered by Integrated Dual Slider-Lever Mechanism, *Robomech2023*, 1A2-E12, (2023).
- [10] Xiuyuan Wu, Yuntian Wang, Keisuke Osawa, Eiichiro Tanaka, Development of a comprehensive motorless walking assistance and evaluation of muscle activity, *ASME, IDETC-CIE 2023, MR-6-1: Medical and Rehabilitation Robotics, IDETC2023-111375*, (2023).

## Development of a Social Coexistent Two-Wheeled Inverted Pendulum Balancing Robot

Chengyan Gu, Taehoon Kim and Joono Cheong

Department of Control and Instrumentation Engineering,

Korea University, South Korea

[guchengyan@korea.com](mailto:guchengyan@korea.com) , [kim.taehoon@korea.ac.kr](mailto:kim.taehoon@korea.ac.kr) , [jncheong@korea.ac.kr](mailto:jncheong@korea.ac.kr)

### Abstract

Research on social coexistence mobile robots capable of harmonious coexistence with humans has gained importance as they are being pervasive in the public places. In order to realize the goal of friendly coexistence with humans, we development a two-wheeled inverted pendulum balancing robot known as KUSBOT (Korea University Social Robot). The robot is designed for achieving stable navigation in complex road conditions and engaging in safe human-robot interaction within social settings. Based on the dynamic analysis of the KUSBOT, the feasibility of the robot is empirically validated.

**Topics:** Robotics, mechatronics, machine design.

### I. Introduction

With the gradual maturation of autonomous driving technology in the current society, intelligent mobile robots are undergoing remarkable development at an astonishing pace. Especially in places with high levels of social interaction, such as airports, banks, and restaurants, there is a significant emergence of guide service-oriented mobile robots. Currently, the prevailing mobile robots in use are mostly four-wheeled or multi-wheeled robots. However, due to the characteristics of the external shape and movement mode of multi-wheeled mobile robots, the robot does not seem compliant against the intentional or unintentional forces applied by humans [1]. As a result, when there is a risk of contacts during their movement in public places, there might be potential harms toward humans and/or the robot. And people may experience negative emotions, such as fear, when encountering such multi-wheeled mobile robots, hindering the achievement of harmonious coexistence between humans and machines. On recognition of this phenomenon, this paper introduces a novel two-wheeled inverted pendulum balancing robot named KUSBOT (Korea University Social Robot) as a good substitute for conventional multi-wheeled mobile robots, as shown in Fig. 1.



Fig. 1. KUSBOT robot

The development of inverted pendulum robots has remained a prominent topic in the field of robot automatic control [2][3]. The two-wheeled inverted pendulum balancing robot system is essentially a walking inverted pendulum system. Its movement tendencies change based on shifts in its center of mass. When subjected to external forces, the front part of the robot will move backward, the motion direction of the robot itself will be compliant with the direction of external forces. Consequently, during interactions between the robot and humans, collisions can be effectively managed by the two-wheeled inverted pendulum balancing robot, preventing potential harm to life or property.

The developed KUSBOT features an exterior with a soft shell composed of polyethylene foam and sponge. Through this soft shell, the safety of the robot's internal components is ensured, as well as the safety of humans during interactions. Furthermore, the KUSBOT robot is equipped with high-precision gyroscopes and accelerometers, allowing real-time determination of the robot's own state through sensors.

### II. Design Concept

The overall model diagram of the KUSBOT is illustrated in Fig 2. The overall design concept of the KUSBOT primarily emphasizes safety and reducing human resistance. The KUSBOT has a cylindrical form factor, enabling

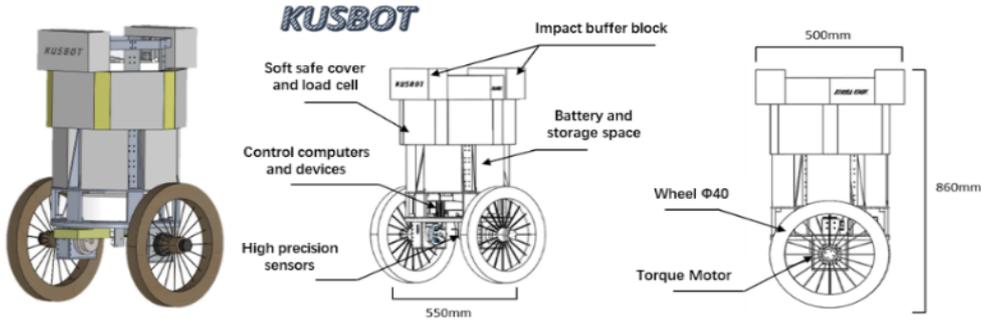


Fig. 2. Overall shape of the KUSBOT

it to exhibit excellent maneuverability even in narrow spaces. Combining the characteristics of the two-wheeled inverted pendulum robot, the KUSBOT is capable of achieving superior soft rotations. The KUSBOT's height is 860mm, which corresponds to the typical height of a person's waist area. This height facilitates interaction between the robot and humans by allowing convenient hand interaction. The upper part of the robot comprises the battery compartment and storage space, covered with a soft material. The robot employs a large 2.8kWh battery weighing 30kg, thus the upper part of the robot is designed in a hexagonal honeycomb structure, providing both stability and storage capacity. As the upper part is the primary interaction area between humans and the robot, it is to be covered not only with a soft cover made of polyethylene foam and sponge but also equipped with six pressure sensors beneath the covering material.

The two-wheeled inverted pendulum is highly sensitive to external action in the forward and backward directions due to its inherent structure, while sensitivity to action in the lateral directions is lower as they rotate around the rotation axis of the wheel. Thus, the pressure sensors are employed to monitor external forces in six directions: left, right, front-left, front-right, rear-left, and rear-right. By observing the forces detected by these sensors, appropriate interactive actions can be performed accordingly. Based on the characteristics of the inverted pendulum type robot described above, we are committed to deploying this type of robot in places with high human traffic, such as hospitals and community centers. The goal is to enhance coexistence between humans and robots by utilizing this type of robot, thereby reducing people's fear towards robots and facilitating a realization of an intelligent lifestyle.

The lower section of the KUSBOT consists of the control and drive components. For the control, we utilize the real-time control system from Beckhoff, which consists of a main PC and terminals, so that control loop runs at 1ms correctly. By continuously reading data from the gyroscope and accelerometer located at the lower end of the robot, we can perform real-time monitoring of the robot's state, enabling prompt balancing. We utilized high-precision accelerometers and gyroscopes in KUSBOT. The accelerometer has a dynamic range of -2.5g to 2.5g, with a resolution of 375mV/g. The gyroscope sensor used has a dynamic range of -100deg/s to 100deg/s, with a resolution of 20mV/deg/s and a Bias Instability of 3.5deg/h. The driving part of the robot employs two harmonic geared torque motors with a maximum torque of 127Nm each. These motors are directly connected to 16-inch wheels. Due to the characteristics of the harmonic geared, the motor backlash is negligible.

### III. Controller Design

In order to achieve balance control for the KUSBOT, dynamic modeling of the robot is essential. As illustrated in Fig. 3(a), we use the virtual joint method [4] to define the intuitive generalized coordinates for representing the robot's state and establishing its dynamic relationships. Ultimately, by the Lagrangian equation [5][6], we can derive the dynamic model of the KUSBOT. (For the space limit of this paper, the detailed derivation of the equation is omitted here. Please refer to [7] for the full derivation of the dynamic equation.)

To accurately comprehend the pitch rotation angle and angular velocity of the robot, we employ high-precision gyroscopes and accelerometers to measure the yaw and pitch angular velocities as well as accelerations along the x, y, and z axes. By applying filters such as low pass filter, complementary filters, and Kalman filters to the

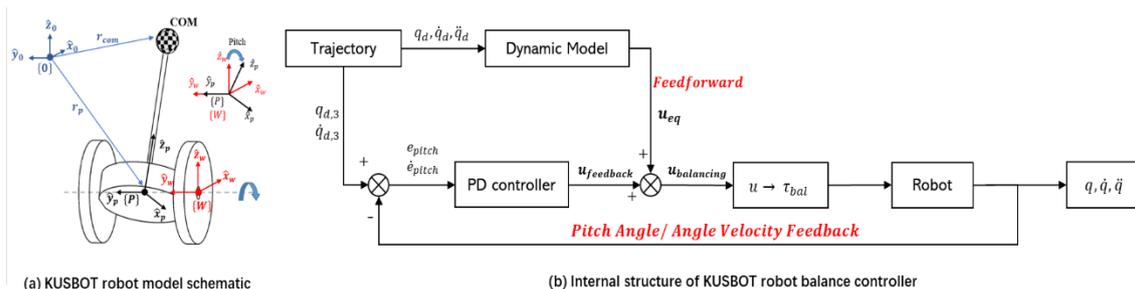
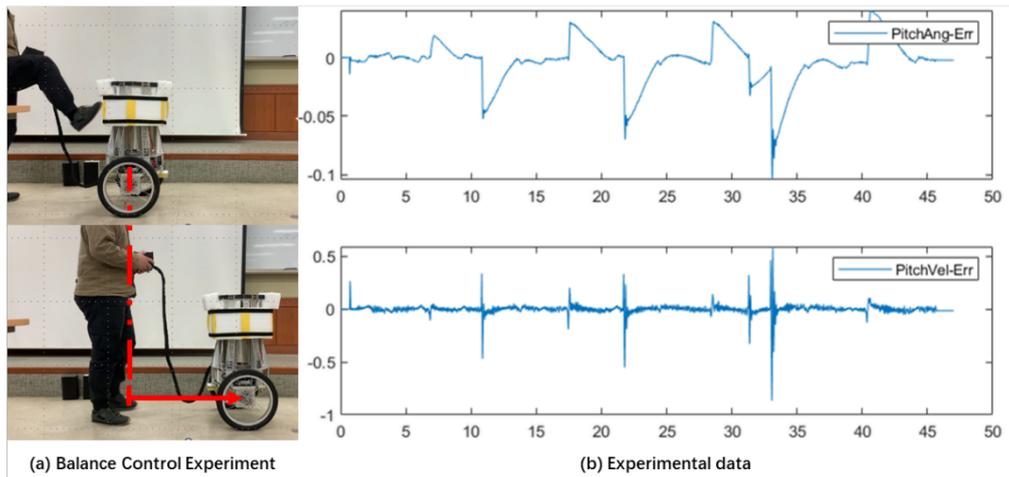


Fig. 3. Control system of KUSBOT with PD control.



**Fig. 4.** KUSBOT robot balance control experiment

sensor signals, we can obtain a more precise depiction of the KUSBOT's motion state. By calculating the error between the current state and the desired state of the robot, the PD controller is employed to achieve a faithful balancing control of the robot. Simultaneously, by leveraging the robot's dynamic model, a feedforward control is achieved with a prior knowledge on the trajectories, as shown in Fig. 3(b).

Based on the designed balancing controller, we conducted experiments by using the KUSBOT. In the experiments, the operator applied a certain opposing force to the front of the KUSBOT while it was in a static balancing state. Results showed that the robot could maintain stable balancing both with a complaint response toward the external force by stepping slightly back. The experimental scenes are displayed in Fig. 4(a). With a sufficiently strong external force applied by the operator, the KUSBOT moved backward for a certain distance and still managed to maintain a stable balanced state. By conducting the experiment three times, the experimental data for Fig. 4(b) were obtained, indicates that the KUSBOT's pitch angle error and angular velocity error converge towards zero when subjected to external forces. This convergence signifies that the KUSBOT is rapidly recovering to its balanced state.

#### IV. Conclusion

In this paper we have introduced a socially interactive two-wheeled inverted pendulum balancing robot and conducts physical development and experimental validation. This robot is capable of various interactions with humans in a social environment and possesses a high level of safety due to its materials and form factor. In the future, we aim to achieve safe navigation of the KUSBOT robot on complicated road conditions by upgrading its tires and navigation sensors. Furthermore, we are working toward on integrating unmanned driving technology with the two-wheeled inverted pendulum balancing robot, aiming to enhance KUSBOT's social interaction capabilities.

#### References

- [1] G. Campion, G. Bastin, and B. Dandrea-Novet, Structural properties and classification of kinematic and dynamic models of wheeled mobile robots, *IEEE Transactions on Robotics and Automation*, vol. 12, no. 1, p.47-62, 1996
- [2] Ha, Y-S. and Yuta, S., Trajectory tracking control for navigation of the inverse pendulum type self-contained mobile robot, *Robotics and Autonomous Systems*, vol. 17, no. 1-2, p. 65-80, 1996.
- [3] Grasser, F., D'Arrigo, A., Colombi, S., and Rufer, A., Joe: A mobile, inverted pendulum, *IEEE Transactions on Industrial Electronics*, vol. 49, no. 1, p. 107, no. 114, 2002.
- [4] H. Jeong, H. Kim, J. Cheong, and W. Kim, "Virtual joint method for kinematic modeling of wheeled mobile manipulators," *International Journal of Control, Automation and Systems*, vol. 12, no. 5, p. 1059–1069, 2014.
- [5] S. Jeong and T. Takahashi, Wheeled inverted pendulum type assistant robot: design concept and mobile control, *Intel Serv Robotics*, vol. 1, no. 4, p 313-320, 2008.
- [6] H. Jin, J. Hwang, and J. Lee, A Balancing Control Strategy for a One-Wheel Pendulum Robot Based on Dynamic Model Decomposition: Simulations and Experiments, *IEEE/ASME Transactions on Mechatronics*, vol. 16, no. 4, p 763–768, 2011.

*17th International collaboration Symposium on Information, Production and Systems (ISIPS 2023), and  
Workshop on AI application strategic industries: MASR 2023 (Mobility, AI, Semiconductor, and Robot)*

[7] M. Kim, J. Park, D. Park, C. Park and J. Cheong, Dynamic Inversion Based Real-Time Trajectory Planning Method for Wheeled Inverted Pendulum using Asymptotic Expansion Technique, IEEE Access, (in press).

## Development of a Stepper Motor Powered Lower Limb Exoskeleton with 3-DOF Kinematic-Biological Matched Hip Joint Structure

Yuntian Wang, Yifan Fang, Xiuyuan Wu, Keisuke Osawa and Eiichiro Tanaka  
Graduate School of Information, Production and Systems  
Waseda University, Japan  
tanakae@waseda.jp

### Abstract

Due to the challenges of aging, there's a pressing need for rehabilitation and walking assistive. Elderly individuals often suffer from joint pain, muscle weakness, and balance issues, impacting their quality of life. To address this, a versatile walking assistive device that use stepper motor is introduced, aiding various movements and detecting user intentions. However, existing devices has a kinematic and biological mismatched hip structure, which may lead to secondary damage. A novel kinematic-biological matched hip joint structure is proposed, enhancing mechanical harmony. The study presents this innovative structure and control method, contributing to improved mobility and independence for the elderly population.

**Topics:** robotics

### Introduction and background

A high demand of rehabilitation and walking assistive device is raised due to the aging problem. A considerable portion of the elderly population is affected by joint pain, muscle weakness, and poor balance. Without external assistance, these individuals may struggle to walk and more likely to fall, which will cause negative effects on their quality of life (QOL). To provide elderly assistance both in their daily life and in rehabilitation progress, a walking assistive device is proposed and built. This device can assist users in the motions of standing up, sitting down, and going up or down stairs as well as supporting people of different heights walking in appropriate gaits. It can also automatically detect users' intentions using only pressure sensors and potentiometers, and it can be used outdoors for extended periods of time.

However, the majority of present exoskeletons as well as proposed walking assistive device have a problem of the mismatch between the kinematic and biological hip (as is shown in Fig. 1), which can lead to unwanted stress on the hip. To prevent secondary damage caused by the misalignment of exoskeleton's mechanical and biological hip, a novel hip joint structure for improving the current walking assistive device is purposed.

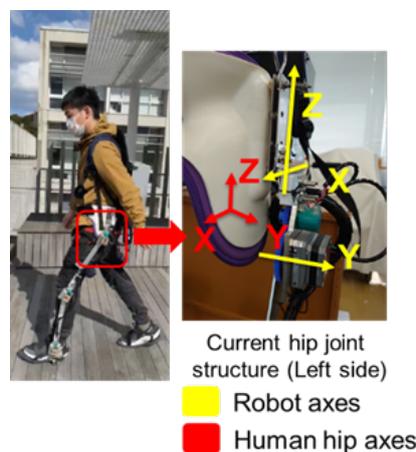


Fig. 1 Current issue of hip joint

### Goals, proposed method, novelty

The goal of this research is to purposed a lower limb exoskeleton with a kinematic-biological matched hip joint structure, which can help elderly to prevent dangerous accidents and improve life quality.

In this paper, we developed a lower limb movement assistive device that assists the hip, knee, and ankle joints of both legs using stepping motors [1], and we report on its assistive effects. Usually the developers of lower limb exoskeleton usually don't use the stepper motor because of its' slipping problem, but we consider it is safer to be slipped in accidents.

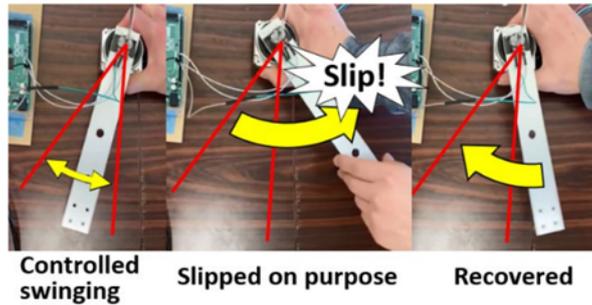


Fig. 2 Intentional slip and recover of the stepper motor

Stepping motors can easily control angles and angular velocities through pulse signals. Although they may slip due to overload, it can recover to the original position to restore synchronization through control by measuring the angle during desynchronization (as is shown in Fig. 2). The motor is controlled based on the toe's trajectory. The trajectory point is used for the calculation of inverse kinematics, then the motor can be controlled based on the calculated target value.

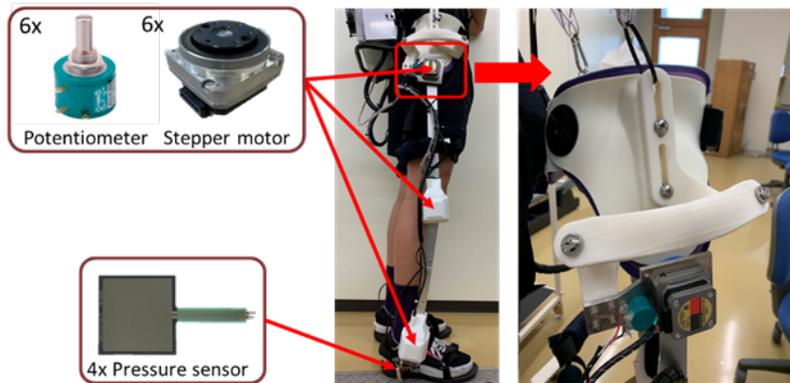


Fig. 3 Structure of the proposed lower limb exoskeleton

Figure. 3 shows the current device and its hip structure. Each joint of the device equipped with a set of potentiometer and stepper motor. The potentiometer is used for measuring the current angle of the joint. Additionally, there are 2 pressure sensors attached on each foot, one in the front and one under the heel. By combining the read value of the potentiometer and pressure sensor, the user's intention including walking, going up/down stairs and standing up/sitting down can be obtained, and the device can support multiple motions.

The newly proposed hip joint's structure has three 1-DOF rotational joints (robot joints) and one 3-DOF ball joint (human hip joint), and by including these joints the total DOF of the joint is 6, which provide the mechanism 3 DOFs (Roll, Pitch and Yaw). Three robot joints are locating on a sphere which center is coincident with human hip joint's center, and all three joints' axes are coincident at the center of the sphere. By connecting joints using three curved bars and human thigh with a straight bar, a closed loop four bar mechanism can be constructed while eliminating kinematic and biological mismatch [2].

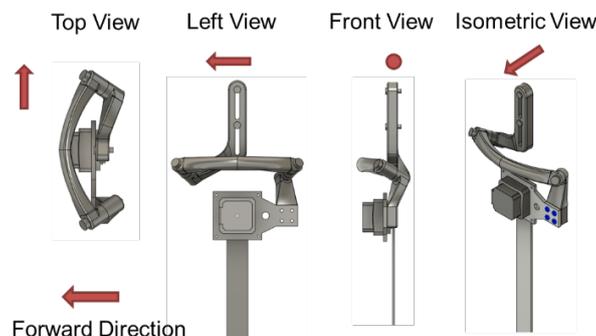


Fig. 4 The 3D model of the hip joint structure (left side)

Figure. 4 shows the 3D model of designed hip joint structure. The novelty is that it uses tilt axes to eliminate the mismatch between the kinematic and biological structure while avoiding the interference between the

exoskeleton and human body. Since each axis of the joint of the hip structure is coincident at the human's hip joint center, the robot's hip joint structure can behave like a ball joint as human hip. And since the human's hip joint can only provide a limited abduction/adduction angle and external/internal rotation angle, the chance of singularity of the mechanism is not needed to be considered.

### Experiments, results, analysis of the results

This exoskeleton assists full leg motion mainly for walking and similar activities. There are 4 muscles play the main roles in support those activities. They are RF (Rectus femoris), BF (Biceps femoris), TA (tibialis anterior), and GAS (gastrocnemius).

There are 3 able bodied 20's participants been experimented the subjects are required to perform the motion of the assistant functions for 3 to 5 times in a roll and compare their %MVC values of wearing the device and without wearing the device of performing the same motion. The %MVC comparison bar graphs of each motion are listed below in Fig. 5.

For normal walking, all muscles activity decreased, it proved that this device is effective to assist user's walking. During motions, agonist and antagonist muscles are simultaneously used, BF and GAS are used against the weight of user. By using this device, the equivalent meaning of the gravity compensation can be realized, then their muscle activities can be decreased.

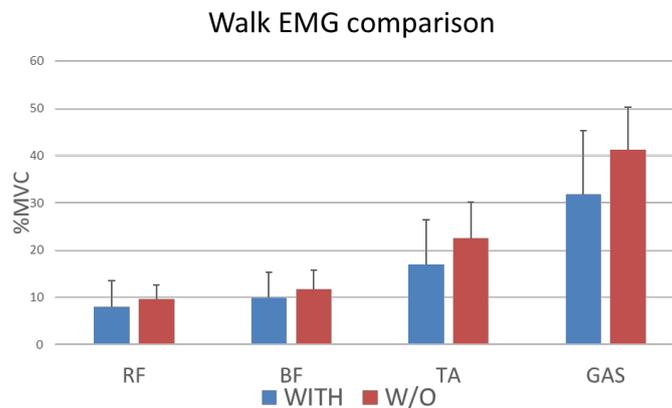


Fig. 5 Experiment result of EMG test

### Conclusion

As a conclusion, a lower limb exoskeleton with a kinematic-biological matched hip joint structure has been developed, and through experiments its effectiveness has been proved. The device can be used outdoor, and can satisfy the daily use of elderly. It can not only assist walking motion, but also assist going up/down stairs and standing up/sitting down. With the newly designed hip joint structure, the kinematic and biological mismatch can be eliminated.

### References

- [1] Yifan Fang, Bingkai Hou, Xiuyuan Wu, Yuntian Wang, Keisuke Osawa, Eiichiro Tanaka, A Stepper Motor Powered Lower Limb Exoskeleton with Multiply Assistance Functions for the Elderly Daily Use, Journal of Robotics and Mechatronics, Vol. 35, No. 3, pp. 601-611, (2023).
- [2] Yuntian Wang, Yifan Fang, Xiuyuan Wu, Keisuke Osawa, Eiichiro Tanaka, Development of a 3-DOF Kinematic-Biological Matched Hip Joint Structure for Lower Limb Exoskeleton, ICMDT2023, ThE1-1, p. 65, (2023).



## Computer vision based surface defect inspection of injection molded products

Hee Min Noh, Pyeon MinJi, Chang Hyun Lee, \*Hun Kee Kim

AI system laboratory

Inha university, Republic of Korea

[ksi064146161@inha.edu](mailto:ksi064146161@inha.edu), [pyun1999@inha.edu](mailto:pyun1999@inha.edu), [95222304@inha.edu](mailto:95222304@inha.edu), [hunkeekim@inha.ac.kr](mailto:hunkeekim@inha.ac.kr)

### Abstract

This study presents a new way to utilize computer vision to detect surface defects such as cracks and shrinkage in injection molded products and distinguish between good and bad products. Previously, surface defects of injection products were mainly detected visually, but the detection process was inconsistent due to the operator's subjectivity, and the accuracy of the detection was poor due to human error in manual inspection. To solve the above-mentioned problems and develop a new method for automatically detecting defective products, computer vision and various image processing techniques were introduced to implement an optimal algorithm for the product to be inspected. In the existing detection system using computer vision, only one product is inspected, but in this study, we propose a method to detect three products simultaneously. Through this process, we achieved a high accuracy of about 95% and a recall of 96% in detecting defective injection products.

**Topics:** image processing, molded product, surface defect

### Introduction and background

In manufacturing, product quality control is a key factor in increasing cost-effectiveness, reliability, and customer satisfaction. In the case of injection molded products, surface defects can have a serious impact on the performance and appearance of the product, so early detection of defects is critical. Traditionally, surface defect inspection of injection molded products has been done manually. This is a labor-intensive and error-prone process that relies on the human eye. In addition, as injection products become more diverse and complex, and production speeds increase, quality inspection methods have not kept pace.

In response to these challenges, computer vision technology is gaining traction. Computer vision is a branch of artificial intelligence that extracts information from images or videos and has been applied to automatic quality inspection, manufacturing automation, and faster and more accurate defect detection [1].

In this paper, we explore how computer vision technology can be applied to solve the problem of surface defect detection in injection molded products. The main goal of this research is to overcome the limitations of existing methods and develop an efficient detection system [2].

The thesis is organized as follows: The first chapter introduces the basics of computer vision techniques. The second chapter describes the computer vision-based surface defect detection algorithm we developed. The third chapter presents the experimental results, and the final chapter discusses the conclusions of this work and future research directions.

### Goals, proposed method, novelty

Computer vision is a technology that allows machines to recognize and process information from images or videos. Computer vision is used in many different fields, and there are several key concepts involved.

First, image processing is the process of converting or analyzing digital images. The main techniques include histogram smoothing, which is a method of evenly adjusting the distribution of brightness values across the entire range to optimize the contrast of an image; filtering, which is a technique of applying various filters to an image to highlight certain patterns or remove noise; and image transformation, which involves performing transformations such as scaling, rotating, and warping an image.

Second, feature extraction is the process of identifying and summarizing important information or patterns in data, which is used to succinctly and effectively represent the information in an image. Examples include SIFT and SURF, algorithms that extract scale- and rotation-invariant features from images; HOG, which extracts features from the distribution of local gradient directions in an image; and color histograms, which represent the distribution of colors in an image.

Finally, pattern recognition aims to find structural patterns or rules in given data to classify or predict that data, which is based on machine learning, statistics, and algorithms. Examples include K-NN, an algorithm that classifies or predicts a given data point by referring to the 'k' closest data points; SVM, an algorithm that classifies data by finding the widest line separating two groups of data; and decision trees, which generate decision paths based on data features and are suitable for classifying or predicting complex data [3].

In this paper, we will focus on image processing.

## Computer Vision based Surface Defect Detection Algorithm

### 1. image acquisition environment building

When detecting a single product, the amount of light illuminating the subject is not a significant limitation. However, if three products are being photographed simultaneously, the amount of light received varies depending on the position of the product, which can lead to poor detection of defective products. To overcome this problem, LEDs were installed at both ends of the darkening box to create an environment that allows the subject to receive light smoothly (Fig.1). The speed at which the product travels when inspecting for surface defects should be the same as in the field, so the speed of the conveyor belt was set at 30 cm/s in consultation with the user company. Market research showed that most companies using goods transportation equipment also use this figure.

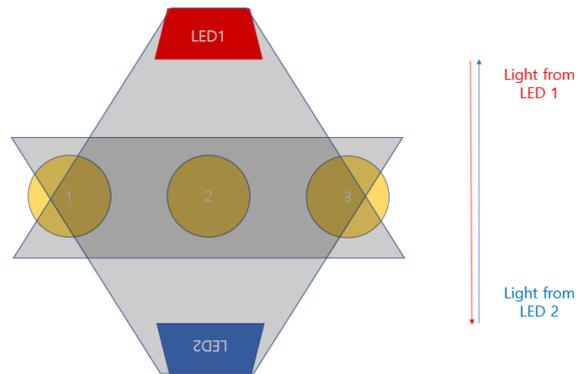


Fig.1: Lighting environment schematic

In order to detect the surface defects of injection products as they move on the conveyor belt, high computing speed is required and the features must be accurately detected. First, the three-channel RGB image was converted to a one-channel grayscale image to speed up the computation (Fig.2-1) [4]. In addition, the convolution was performed using a 5x5 Gaussian filter, a technique for smoothing images to eliminate noise and preserve edges in gray images of injection products (Fig.2-2) [5]. The background image data other than the surface of the injection part is unnecessary and can be removed to speed up the computation. Here, thresholding is applied to the image in Fig 2-2(Fig.2-3).

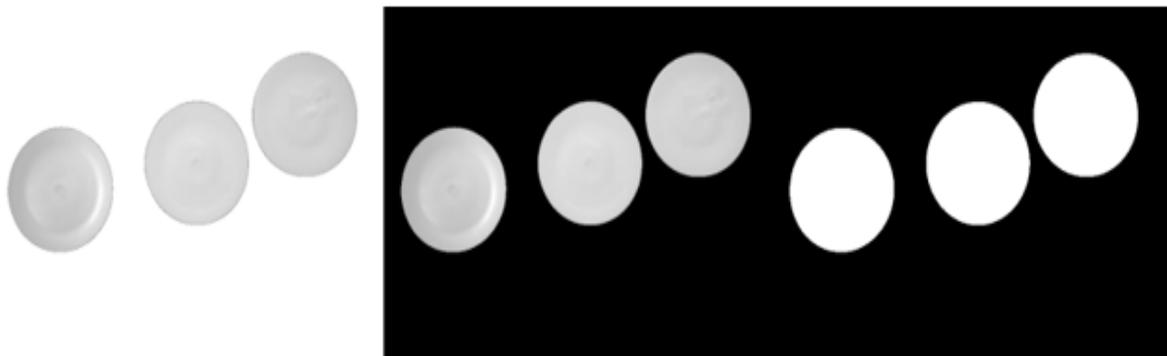


Fig.2: pre-processed images 1

Histograms were constructed and smoothed to maximize the contrast in the images with the various filters applied. This allowed the surface defects, shrinkage and cracks, to be clearly distinguished from the surface color, and the boundaries of the surface defects to be outlined (Fig.3-1) [6]. To locate the borders of the injection parts, the edge detection technique Canny was used. The image pixels were scanned as a whole, and the parts of the data with high gradient in the vertical and horizontal directions were considered as edges, and hysteresis thresholding was used to distinguish between real and fake edges to display only real edges (Fig.3-2) [7]. To improve visibility, we inverted the image above and applied bitwise OR as shown in Figure 2-3(Fig.3-3).

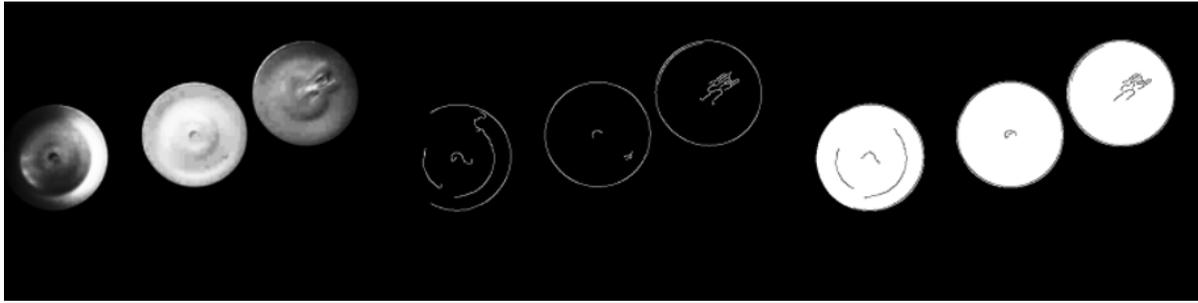


Fig.3: pre-processed images 2

### Experiments, results, analysis of the results

We found the size of the black pixel inside the edge and distinguished whether the product is defective and the type of defect based on the value. In this case, if only one product is inspected, the total size of the black pixel determines whether the product is defective or not, but since we are inspecting three products at the same time, we need to find the size of each product's black pixel separately. Therefore, we designed the algorithm to detect three products simultaneously by finding the size of the black pixel in the closed curve corresponding to the outline of the product using the contour method. Normal products usually have pixel values below 800, while defects such as shrinkage and cracks have values between 1000 and 1500.

Then, using the contour technique, the color of the injection product was marked differently according to the type of defect, and the text was marked whether it was a defective product or a normal product. (Fig.3). In addition, since the angle of the camera may misrecognize whether a product is defective or not, data from when the injection product was first captured by the camera as it moved through the conveyor belt to when it was released was collected and averaged to determine whether the product was ultimately defective or not.

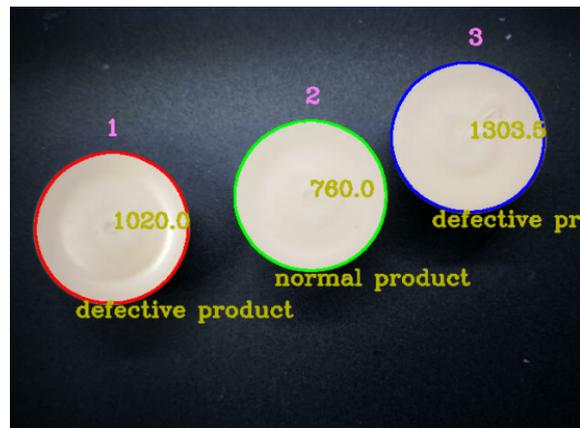


Fig.4: Image of injection product defect judgement

We tested the algorithm constructed in Chapter 2 and obtained the inspection accuracy and recall. The experiment was conducted with 100 products, including 50 good products, 25 cracked products, and 25 shrinkage products, and the accuracy and reproducibility were calculated using the confusion matrix.

Accuracy was calculated as the Number of samples correctly judged as defective or good  $(TN + TP) / \text{total number of samples } (TN+FP+FN+TP) * 100\%$ , and recall was calculated as the number of samples judged as good  $(TP) / \text{number of samples that were actually good } (FN+TP) * 100\%$  [8]. The results of the experiment were as follows:

		Preditive Values	
		Positive (1)	Negative (0)
Actual Values	Positive (1)	48	2
	Negative (0)	3	47

Table 1: Confusion matrix of experimental results

The number of times a good product was judged as good was 48, the number of times a good product was judged as bad was 2, the number of times a bad product was judged as good was 3, and the number of times a bad product was judged as bad was 47. Using these values, we can calculate the accuracy and recall as follow:

$$\text{Accuracy} = (48+47) / (48+2+3+47) * 100\% = 95\%$$

$$\text{Recall} = 48 / (48+2) = 96\%$$

## Conclusion

The application of computer vision technology to detect surface defects in injection products is one of the promising methods for quality control. Various image preprocessing techniques were effective in refining the data, which allowed the algorithm to effectively distinguish between good and bad products. In addition, by developing an algorithm that can inspect three products simultaneously, the speed of inspection is increased, making the quality control process more productive.

The validation of the algorithm was also successful. The 95% accuracy and 96% recall rate emphasize the robustness and reliability of the system, reducing the likelihood of defective products and increasing customer satisfaction. If implemented in an automated line, the algorithm can significantly reduce the dependence on manual work, contributing to a reduction in labor costs and greater consistency of inspection results.

In the future, we plan to add techniques to image preprocessing to detect smaller defects and optimize the algorithm for automated lines. They also plan to combine it with machine learning or deep learning techniques to increase the accuracy and speed of detection.

## References

- [1] Guo, Meng-Hao, et al., Attention mechanisms in computer vision: A survey, *Computational visual media* 8.3 p.331-368, 2022.
- [2] Luo, Qiwu, et al., Automated visual defect detection for flat steel surface: A survey, *IEEE Transactions on Instrumentation and Measurement* 69.3., p 626-644, 2020.
- [3] Mishra, Alok, Pallavi Asthana, and Pooja Khanna. The quality identification of fruits in image processing using Matlab, *International Journal of Research in Engineering and Technology.*, p. 92—95, 2014.
- [4] Ahmad, Ijaz, Inkyu Moon, and Seok Joo Shin., Color-to-grayscale algorithms effect on edge detection—A comparative study, *2018 International Conference on Electronics, Information, and Communication (ICEIC). IEEE*, p.1~4, 2018.
- [5] Liao, Dahai, et al., Surface defect detection of Si3N4 ceramic bearing ball based on improved homomorphic filter-Gaussian filter coupling algorithm, *AIP Advances.*, p.2~4, 2022.
- [6] Liu, Yepeng, et al., Image smoothing based on histogram equalized content-aware patches and direction-constrained sparse gradients, *Signal Processing* 183., p., 2021
- [7] Sekehravani, Ehsan Akbari, Eduard Babulak, and Mehdi Masoodi, Implementing canny edge detection algorithm for noisy image, *Bulletin of Electrical Engineering and Informatics.*, p.1404-1410, 2020.
- [8] Luque, Amalia, et al., The impact of class imbalance in classification performance metrics based on the binary confusion matrix, *Pattern Recognition* 91., p.216-231, 2019.

## Graphene nanostructures grown on stainless steel by plasma chemical vapor deposition

Siyu Jia, Jun Kameoka and Kenji Ueda

Graduate School of Information, Production, and Systems

Waseda University, Japan

jiasiyu@toki.waseda.jp, jkameoka@waseda.jp and k-ueda@waseda.jp

### Abstract

In this study, we focus on the growth of graphene on stainless steel (SUS 304), which can be used as electrodes in electrochemical systems. The effect of the growth temperature on the morphological and structural characteristics of graphene nanostructures is illustrated. Graphene nanostructures on the SUS substrate are controlled by the growth temperature and the nanostructures transform from the amorphous carbon layer with graphite like layer at 600 °C to few-layer graphene at 700 °C, and ultimately to vertical graphene nanowalls at 900 °C. This structure transformation can be explained by the equilibrium between carbon diffusion into the SUS substrate, carbon etching effects by H<sub>2</sub> and the supply rate of carbon sources.

**Topics:** materials engineering, production

### Introduction and background

Graphene, hexagonal arrangement of carbon atoms forming one-atom thick planar sheet, is a promising material for future electronic applications due to their high electrical conductivity and mobility as well as chemical and physical stability [1]. Planar graphene films with respect to the substrate have been synthesized using various methods including mechanical exfoliation, epitaxial growth of SiC and chemical vapor deposition (CVD) [2-4]. Plasma-enhanced CVD (PECVD) is developed on the basis of CVD and used to synthesize planar graphene films at low temperatures or vertically standing few-layer graphene [5]. In addition, PECVD is becoming one of the most promising techniques to produce carbon materials, including diamond, aligned carbon nanotube films, and carbon nanowalls (CNWs), owing to its feasibility and potentiality for large-area production with reasonable growth rates at relatively low temperatures.

Thanks to its chemical stability and high electrical conductivity, SUS is widely used as electrodes in electrochemical systems (batteries, supercapacitors, catalysts). Various research groups have studied the growth of graphene nanostructures on SUS substrates, focusing on improving its conductivity, corrosion resistance and mechanical properties (surface hardness) and morphological characteristics for applications to energy, catalysis and sensors [6-7]. However, the growth of carbon nanostructures on SUS substrates at high temperatures is a complex process, since the composition of SUS (mainly based on Fe, Cr, and Ni) on the substrate's surface can vary considerably [8]. Under high temperatures, a deposition reaction occurs between the carbon atoms and SUS forming chromium iron carbides (Cr, Fe)<sub>x</sub>C<sub>y</sub> [9]. After the pyrolysis of the methane, most of the available carbon atoms preferentially reacted with iron in SUS to form iron carbide, whereas only an extremely small quantity of carbon to form graphene layer [10]. Furthermore, temperature exerts an influence on the solubility of carbon and the precipitation of carbides, making an impact on the growth of graphene nanostructures on SUS substrates. PECVD is used to expand the attainable temperature for the graphene nanostructures growth on SUS substrates through the advantageous influence of plasma enhancement accelerating methane decomposition and concurrently lowering the activation energy required for graphene growth. Nevertheless, limited research has been conducted on the impact of carbon dissolution and deposition on the morphology during the growth process. Thus, in this study, we discussed the effect of growth temperature on graphene nanostructures grown on SUS substrates by PECVD in combination with the growth mechanism.

### Experimental details and growth mechanism

Graphene nanostructures were synthesized using microwave plasma enhanced CVD (PECVD) employing a CH<sub>4</sub>/H<sub>2</sub> mixture. Substrates used for growth experiments were SUS 304 with the thickness of 50-100 μm. Fig. 1 shows the schematic of reactor used for the growth of graphene nanostructures.

Growth experiments were carried out at a substrate temperature of 600~1000 °C, total gas pressure of 2.6 KPa (20 Torr), deposition time of 30 min (after 2 min of H<sub>2</sub> pretreatment) and CH<sub>4</sub> and H<sub>2</sub> flow rates of 10 and 6.5 sccm, respectively. Carbon deposits were evaluated by scanning electron microscopy (SEM), X-ray diffraction (XRD) and Raman spectroscopy.

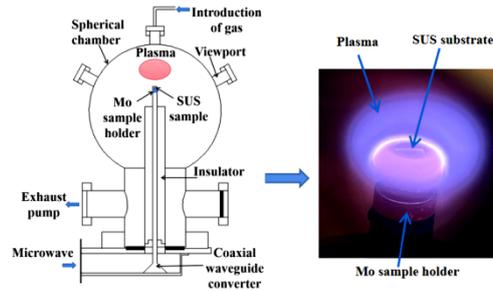
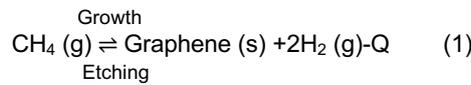


Fig. 1: Schematic of the reactor used for the growth of graphene nanostructures.

The growth mechanism of graphene nanostructures for PECVD is illustrated in Fig. 2. The interaction between high-energy electrons in plasma and gas molecules ionizes and reacts, forming various free radicals through dissociation reactions. In PECVD,  $\text{CH}_4$  is dissociated as  $\text{CH}_x$ ,  $\text{C}_2$  and  $\text{H}_2$  radicals. The radicals migrate to form nucleation sites under ion irradiation and graphene nanosheets coalesce at the nucleation site to form nuclei and grow, which called nucleation and coalescence mechanism [11]. In addition, a reversible equilibrium between  $\text{H}_2$  plasma etching and  $\text{CH}_4$  plasma growth introduces the etching and growth mechanism [12]. The balanced competition is represented by Equation 1:



, where Q is reaction heat. In general, the morphology of graphene nanostructures are determined by these two mechanisms. Within transition metals characterized by elevated carbon solubility, such as nickel (Ni), the process of plasma-enhanced graphene growth appears as a typical dissolution and precipitation mechanism. Through the regulation of temperature, the quantity of carbon dissolved within the Ni substrate is managed, resulting in the acquisition of different layers of graphene. In the metal substrates characterized by limited carbon solubility (like Cu), the primary formation process of graphene entails the catalyzed dissociation of the carbon source on the surface of the copper substrate which called surface-catalyzed dissociation processes. The growth process of graphene on the SUS substrate involves numerous mechanisms for carbon incorporation due to the solubility of carbon in iron. Carbon diffusion initiates from the initial stage of carbon deposition, and in the presence of plasma, carbon undergoes nucleation on the SUS substrate, resulting in the formation of diverse graphene nanostructures.

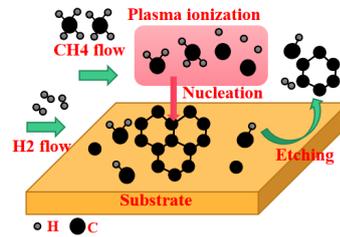
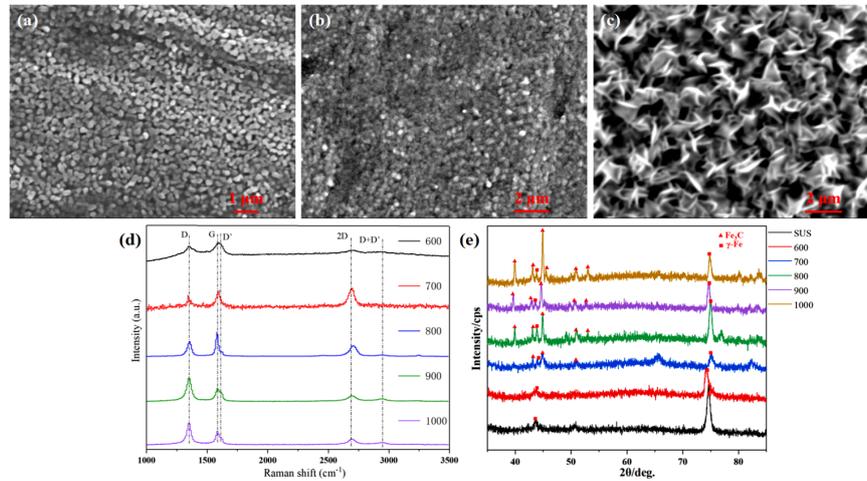


Fig. 2: Schematic illustration of the growth mechanism by PECVD.

## Results, analysis of the results

Fig. 3(a-c) show the SEM images of graphene nanostructures deposited at 600, 700 and 900 °C. The surface at 700 °C was smoother than that at 600 °C, and vertically grown graphene-like nanosheets could only be seen more than 900 °C. The vertically grown graphene-like nanosheets, also called graphene nanoflakes, is three-dimensional graphene, which can be described as graphite sheet nanostructures with edges that are consisted of stacks of planar graphene sheets standing almost vertically on the substrate [13]. Raman spectrum gives information about the hybridization and the defect density of carbon materials. The pristine Raman spectrum of graphene predominantly comprises two distinct bands of significance: the G band, situated around  $1580 \text{ cm}^{-1}$ , associated with vibrational symmetry and crystalline quality, and the 2D band, positioned near  $2700 \text{ cm}^{-1}$ , attributable to two-mode non-elastic phonon scattering. However, due to the defects and disorder of graphene films, the Raman spectrum may exhibit D band ( $1350 \text{ cm}^{-1}$ ) and D' band ( $1620 \text{ cm}^{-1}$ ). The breadth of characteristic bands in Raman spectrum is typically indicative of structural attributes, lattice imperfections, molecular oscillations, and interactions of materials. Amorphous carbon generally comprises various carbon configurations with irregular structures and a large number of lattice defects, giving rise to the large broad overlapping characteristic bands as shown in Fig. 3d (black line at 600 °C). In the case of samples grown at 700~1000 °C, the Raman spectrum exhibited a D band, G band, and 2D band. The appearance of G and 2D bands indicated the presence of graphene on the sample. The  $I_{2D}/I_G$  ratio provides information related to the number of graphene layers.  $I_{2D}/I_G$  ratio reached to the highest of 1.18 at 700 °C, signifying few-layer graphene formed on the surface and this ratio gradually decreased with increasing temperature, indicating the formation of multilayer graphene or graphite like layers. The G band is accompanied by a shoulder D' band, as well as a strong D band and D+D' band appears. These characteristics are associated with the presence of finite-size graphite crystals and graphene edges, which are prevalent features of vertical graphene nanostructures [14]. When the temperature rose to 900 °C, vertical graphene nanostructures were formed on the surface. According to the XRD results in Fig. 3e,  $\text{Fe}_3\text{C}$  diffraction peak (obtained from the comparison of standard PDF cards) began to appear at 700 °C,

indicating that carbon atoms diffused into the substrate in the configuration of interstitial carbons and part of them reacted with iron to form iron carbide. The type of diffraction peak at 600 °C was the same as the SUS substrate, and the displacement of the  $\gamma$ -Fe diffraction peak towards a smaller angle indicated the diffusion of carbon atoms into the  $\gamma$ -Fe lattice, resulting in the enlargement of the lattice constant. The rise of temperature was accompanied by the formation of carbides in the crystal, which reduced the supersaturation of carbon in the lattice and decreased the shift of diffraction peak.

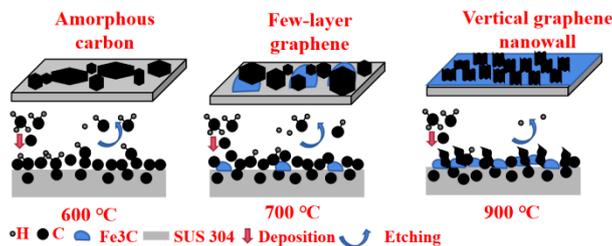


**Fig. 3:** Surface SEM images of graphene nanostructures at (a) 600, (b) 700 and (c) 900 °C, (d) Raman spectra of samples (600~1000 °C), (e) XRD patterns of samples (600~1000 °C).

Based on the above results and growth mechanism, we proposed a model for graphene grown on the SUS 304 substrate as shown in Fig. 4. The radicals ( $\text{CH}_x$  and  $\text{C}_2$ ) for growth of graphene nanostructures, some diffused into the SUS 304 substrate, and some adsorbed on the SUS 304 surface and migrated to form nucleation sites under the ion irradiation. The diffusion coefficient of carbon in iron is calculated from Equation 2 [15]. Here,  $D_0$  is diffusion vibration factor,  $Q$  is activating energy,  $R$  is gas constant and  $T$  is temperature:

$$D = D_0 e^{-\frac{Q}{RT}} \quad (2)$$

It could be inferred that at 600 °C, carbon infiltrated into the SUS substrate to form a supersaturated austenite structure. Nevertheless, the diffusion coefficient of carbon remained smaller and less carbon infiltrated, causing the relatively high proportion of  $\text{CH}_4$  (high supply rate of carbon sources) on the SUS substrate. According to Equation 1, the reversible reaction moved to the growth side, leading to the nucleation and a substantial accumulation of carbon on the surface to form an amorphous carbon layer with graphite like layer. Increasing growth temperature, the diffusion of carbon led to the formation of carbides on the surface. In contrast to 600 °C, significant increase of diffusion coefficient (large carbon diffusion into the SUS substrate), decreasing the  $\text{CH}_4$  proportion to move Equation 1 to the etching side. Furthermore, heat effects moved Equation 1 to the growth side. The etching effect reduced the accumulation of disordered carbon on the surface and the growth effect facilitated the nucleation and growth of graphene nanostructures. Equation 1 achieved a balance between growth and etching, forming few-layer graphene at 700 °C. Further increasing temperature, the diffusion coefficient at 900 °C was 17 times larger than that at 700 °C, leading to the rapid formation of iron carbide, and accelerating the nucleation density of graphene. Concurrently, small and disordered graphene nanosheets grow at granular sites, giving rise to vertical graphene nanostructures.



**Fig. 4:** Schematic illustration of the growth mechanism of graphene on the SUS 304 substrate.

## Conclusion

Graphene nanostructures were prepared on SUS 304 substrates by PECVD. Temperature affected the growth mechanism of graphene nanostructures to modify their structures and morphologies. At 600 °C, high supply rate of carbon sources resulted in more accumulation of carbon and nucleations on the surface to form the amorphous carbon layer with graphite like layer. Subsequently, at 700 °C, a delicate balance was achieved among etching, penetration, and growth, leading to the formation of few-layer graphene with  $I_{2D}/I_G$  ratio of 1.18. At 900 °C, the rapid formation of iron carbide and the high nucleation density gave rise to the formation of vertical graphene nanowalls.

## References

- [1] Geim Andre K., and Konstantin S. Novoselov., The rise of graphene. *Nature materials*, 6.3(2007), 183-191.
- [2] Lotya Mustafa, et al., Liquid phase production of graphene by exfoliation of graphite in surfactant/water solutions. *Journal of the American Chemical Society*, 131.10(2009), 3611-3620.
- [3] Mohsin Ali, et al., Synthesis of millimeter-size hexagon-shaped graphene single crystals on resolidified copper. *Acs Nano*, 7.10(2013), 8924-8931.
- [4] Huang Han, et al., Bottom-up growth of epitaxial graphene on 6H-SiC (0001). *ACS nano*, 2.12(2008), 2513-2518.
- [5] Hiramatsu Mineo, et al., Nucleation control of carbon nanowalls using inductively coupled plasma-enhanced chemical vapor deposition. *Japanese Journal of Applied Physics*, 52.1S(2013), 01AK05.
- [6] Giese André, et al., Synthesis of carbon nanowalls from a single-source metal-organic precursor. *Beilstein journal of nanotechnology*, 9.1(2018), 1895-1905.
- [7] Ruammaitree Akkawat, et al., Surface hardening of stainless steel by coating graphene using thermal chemical vapor deposition. *Solid State Phenomena*, 283(2018), 173-178.
- [8] Bertran-Serra Enric, et al., Temperature-modulated synthesis of vertically oriented atomic bilayer graphene nanowalls grown on stainless steel by inductively coupled plasma chemical vapour deposition. *Applied Surface Science*, 610(2023), 155530.
- [9] Wang Z. B., et al., Chromizing behaviors of a low carbon steel processed by means of surface mechanical attrition treatment. *Acta Materialia*, 53.7(2005), 2081-2089.
- [10] Nen-Wen Pu, et al., Graphene grown on stainless steel as a high-performance and ecofriendly anti-corrosion coating for polymer electrolyte membrane fuel cell bipolar plates. *Journal of Power Sources*, 282(2015), 248-256.
- [11] Sun Jian, et al., Large-scale nanoelectromechanical switches based on directly deposited nanocrystalline graphene on insulating substrates. *Nanoscale*, 8.12(2016), 6659-6665.
- [12] Wei Dacheng, et al., Critical crystal growth of graphene on dielectric substrates at low temperature for electronic devices. *Angewandte Chemie*, 125.52(2013), 14371-14376.
- [13] Zhang Huan, et al., Efficient and controllable growth of vertically oriented graphene nanosheets by mesoplasma chemical vapor deposition. *Carbon*, 147 (2019), 341-347.
- [14] Hiramatsu Mineo, et al., Graphene nanowalls. *New Progress on Graphene Research*, 10(2013), 51528.
- [15] Homan Clarke G., Diffusion of carbon in alpha iron. *Acta Metallurgica*, 12.9(1964), 1071-1079.

## A study on the high-aspect-ratio oxide etching characteristics using a 2 MHz bias power ICP system

Kwang-Ho Kwon

Department of Control and Instrumentation Engineering, Korea University, Sejong, 30019, Republic of Korea  
[kwonkh@korea.ac.kr](mailto:kwonkh@korea.ac.kr),

### Abstract

In this study, we attempted to investigate individual effects of various plasma-related factors on high aspect ratio (HAR) etching characteristics of SiO<sub>2</sub> with using the amorphous carbon layer (ACL) mask. Experiments were carried out in an inductively coupled plasma etching system in CF<sub>4</sub>/C<sub>4</sub>F<sub>8</sub>/He and C<sub>2</sub>F<sub>6</sub>/C<sub>4</sub>F<sub>8</sub>/He gas systems with low-frequency (2 MHz) bias power sources. For HAR oxide etching, a Hexafluoro isobutylene (C<sub>4</sub>H<sub>2</sub>F<sub>6</sub>) gas with a low global warming potential of 18 was also examined, as an alternative gas for C<sub>4</sub>F<sub>8</sub> gas. Based on these results, the possibility for HAR oxide etching using C<sub>4</sub>H<sub>2</sub>F<sub>6</sub> gas chemistry was suggested in this work.

**Topics:** Give between 1 and 5 topics. Use the topics on the [site of the symposium](#). Indicate the thematic session topic 'Thematic session topic' if it is the case.

### Introduction and background

In recent years, conventional etching techniques used with capacitively coupled plasma reactive ion etching (CCP-RIE) systems often face challenges, such as contact hole distortion, clogging, tilting and non-opening.[1] To address these issues, researchers have been exploring alternative gas chemistries and process parameters to improve oxide etch properties. This study focuses on investigating the properties of HAR oxide etching in an inductively coupled plasma (ICP) system using low-frequency (2 MHz) bias power. In our previous work, we first suggested the applicability of ICP etching system with a low-frequency (2MHz) bias power to the HAR oxide etching with a pattern of 100nm or less.[2] Obviously, evident benefits in respect to etching selectivity and etching profile may be obtained from the individual control of physical and chemical etching pathways.[3] Thus, in this work, we tried individually to control the physical and chemical effect on the oxide etching characteristics. We also intend to contribute to the development of eco-friendly and technologically advanced next-generation HAR oxide etching technology by evaluating oxide films and ACL materials etching processes using C<sub>4</sub>H<sub>2</sub>F<sub>6</sub> gas chemistry with a low global warming potential.

### Experiments, results, analysis of the results

Both etching and plasma diagnostics experiments were performed in an ICP etcher described in previous publications.[4, 5] Experiments were carried out with CF<sub>4</sub>/C<sub>4</sub>F<sub>8</sub>/He and C<sub>2</sub>F<sub>6</sub>/C<sub>4</sub>F<sub>8</sub>/He gas mixtures, to control the individual control of radical flux. On the other hand, C<sub>4</sub>H<sub>2</sub>F<sub>6</sub> which has low GWP potential of 18, has been evaluated as the alternative gas to solve environmental problems. Accordingly, an experiment was also performed using CF<sub>4</sub>/C<sub>4</sub>H<sub>2</sub>F<sub>6</sub> (or C<sub>4</sub>F<sub>8</sub>)/He as a mixed gas composition for the etching process, to confirm the C<sub>4</sub>H<sub>2</sub>F<sub>6</sub> etching characteristics as an alternative of C<sub>4</sub>F<sub>8</sub> gas. Both the etching characteristics of CF<sub>4</sub>/C<sub>4</sub>F<sub>8</sub>/He chemistry and CF<sub>4</sub>/C<sub>4</sub>H<sub>2</sub>F<sub>6</sub>/He are compared and evaluated.

- Individual control of chemical effect on the oxide etching characteristics.

Figure 1 shows the changes in the etched oxide profile according to the two gas mixtures. The gas mixing ratio of C<sub>4</sub>F<sub>8</sub> was decreased from 100% to 50% from left to right. As shown in this figure, it can be confirmed that in the etched ACL profile of the sample using CF<sub>4</sub>, a slightly more pronounced ACL facet phenomenon was observed compared to C<sub>2</sub>F<sub>6</sub>. This is because the selectivity is determined by the CF<sub>x</sub>/F ratio. If an enough amount of polymer is not formed on the sample surface, the ACL thin film can be significantly affected by ions accelerated by a large V<sub>DC</sub> (self-DC bias voltage). Meanwhile, it can be observed that to improve the etch selectivity of the ACL thin film, it is necessary to form a thick polymer on the surface of the ACL thin film.

From this figure, we confirmed that the high aspect ratio oxide etching process is possible by using the ICP system to which a 2 MHz bias power is applied. In addition, it was also confirmed that each plasma parameter such as the ion density, ion energy, and radicals could be individually changed in the plasma etching process. This method will be practical and useful for examining the etching characteristics.

Figure 2 shows the C 1s spectra detected from the surface of the ACL thin film after partial etching of the ACL thin film using two mixed gases. The atomic % was extracted under two mixed gas conditions. For the case of CF<sub>4</sub>/C<sub>4</sub>F<sub>8</sub> gas mixture, C 1s and F 1s were 62.74 % and 37.26 %, respectively. Under the condition of the C<sub>2</sub>F<sub>6</sub>/C<sub>4</sub>F<sub>8</sub> gas mixture, C 1s and F 1s were 67.65 % and 32.35 %, respectively. In particular, when C<sub>2</sub>F<sub>6</sub> gas was used, the atomic % of C 1s was greater. As shown in this figure, when C<sub>2</sub>F<sub>6</sub>/C<sub>4</sub>F<sub>8</sub> gas is used, the amount of polymer

formed on the surface of the ACL thin film increases. This was confirmed to be due to an increase in the C-F bond.

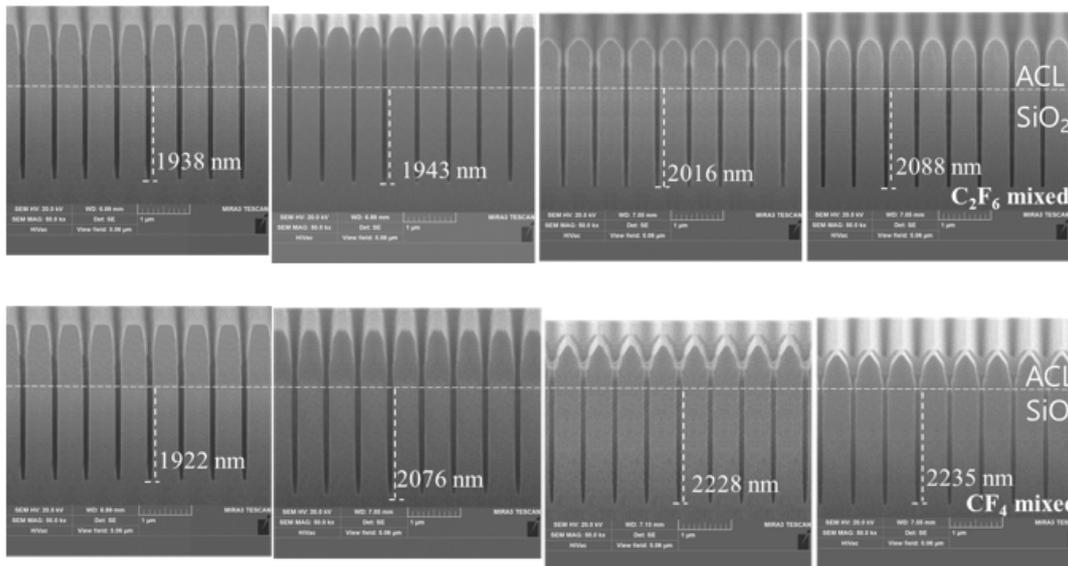


Fig.1: The comparison between  $\text{CF}_4$  and  $\text{C}_2\text{F}_6$  etching profiles for individual control of radical flux.

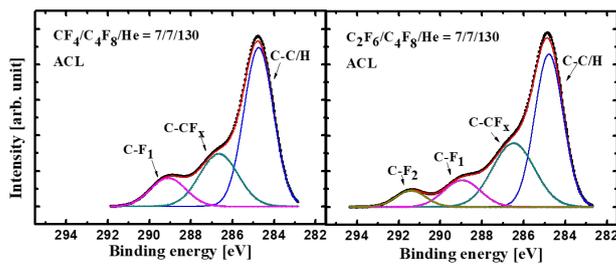


Fig. 2 XPS C 1s deconvolution of ACL material etched at the  $\text{CF}_4/\text{C}_4\text{F}_8/\text{He}$  plasma.

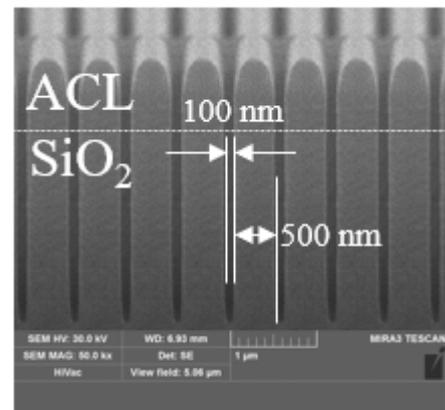


Fig3: The etched oxide profile using  $\text{C}_4\text{H}_2\text{F}_6$  gas

- The oxide etching characteristics of  $\text{C}_4\text{H}_2\text{F}_6$  gas as an alternative of  $\text{C}_4\text{F}_8$  gas.

Figure3 shows the etched profile of HAR oxide film. As shown in this figure, the anisotropic etched profile was obtained. This figure definitely indicates that the  $\text{C}_4\text{H}_2\text{F}_6$  gas can be applied to the HAR oxide etching process as an alternative of the  $\text{C}_4\text{F}_8$  gas.

## Conclusion

In this work, it has been suggested that high aspect ratio oxide film etching process with a size of 100 nm or less, which has been processed only by the RIE system so far, is surely possible by using the ICP system to which a 2 MHz bias power is applied.

It was also confirmed that each plasma parameter such as the ion density, ion energy, and radicals could be individually changed in the plasma etching process. This method will be practical and useful for examining the etching mechanism.

In addition, this study investigated the etching characteristics of HAR oxide using the eco-friendly gas  $\text{C}_4\text{H}_2\text{F}_6$  as a replacement for  $\text{C}_4\text{F}_8$  in an ICP system. Overall, the findings suggest that using  $\text{C}_4\text{H}_2\text{F}_6$  gas as an alternative to  $\text{C}_4\text{F}_8$  gas in HAR oxide etching processes is feasible. It offers the advantages of reduced global warming potential and higher selectivity to ACL materials

## References

- [1] M. Wang, Mark J. Kushner, High energy electron fluxes in dc-augmented capacitively coupled plasmas. II. Effects on twisting in high aspect ratio etching of dielectrics, *J. Appl. Phys.* 107, p.023309, 2010.
- [2] J. Kim, G. Choi, K. -H. Kwon, High-aspect-ratio oxide etching using  $CF_4/C_6F_{12}O$  plasma in an inductively coupled plasma etching system with low-frequency bias power, *J. Vac. Sci. Technol. A* 31, p.050825, 2022.
- [3] Y. Oh, A. Efremov, J. Lee, J. Lee, Y. Choi, K. -H. Kwon, Etching kinetics and dielectric properties of SiOC films exposed to Ar and  $CF_4$  plasmas, *Thin Solid Films.*, p.139185, 2022.
- [4] A. Efremov, N.-K. Min, B.-G. Choi, K.-H. Baek, K.-H. Kwon, Model-Based Analysis of Plasma Parameters and Active Species Kinetics in  $Cl_2 / X$  ( $X = Ar, He, N_2$ ) Inductively Coupled Plasmas, *J. Electrochem. Soc.*, 155, p. D777, 2008.
- [5] A. Efremov, Y. Kim, H.-W. Lee, K.-H. Kwon, A Comparative Study of HBr-Ar and HBr- $Cl_2$  Plasma Chemistries for Dry Etch Applications, *Plasma Chem. Plasma Proc.*, 31, p. 259, 2011.



## **PROGRAMME CHAIRS**

Choong-ho CHO, Korea University, Korea

Hee-hyol LEE, Waseda University, Japan

Jae-hyun PARK, Inha University, Korea

## **PROGRAMME COMMITTEE**

Dae-kug LEE, Korea University, Korea

Eiichiro TANAKA, Waseda University, Japan

Handdeut CHANG, Incheon National University, Korea

Hun-kee KIM, Inha University, Korea

Joono CEONG, Korea University, Korea

Kenji HASHIMOTO, Waseda University, Japan

Kenji UEDA, Waseda University, Japan

Ki-jung PARK, Incheon National University, Korea

Kwang-ho KWON, Korea University, Korea

Masao ARAKAWA, Waseda University, Japan

Min-ho JO, Korea University, Korea

Myung-sup KIM, Korea University, Korea

Soong-keun HYUN, Inha University, Korea

Yeong-jin KIM, Incheon National University, Korea

Young-bum HUR, Inha University, Korea

Young-eun SONG, Korea University, Korea



早稲田大学 情報生産システム研究科

Graduate School of Information, Production and Systems, Waseda University



TOP GLOBAL  
UNIVERSITY JAPAN



公益財団法人 北九州産業学術推進機構

Kitakyushu Foundation for the Advancement of Industry Science and Technology

**ISIPS2023**

*17th International collaboration Symposium on Information, Production and Systems  
and Workshop on Next-generation Technology for AI Application Strategic Industries:  
MASIR 2023 (Mobility, AI, Semiconductor, and Robot)  
16<sup>th</sup> – 17<sup>th</sup> November 2023, Kitakyushu, Fukuoka, JAPAN*