An Ultra Fast Semantic Segmentation Model for AMR's Path Planning

Hoai-Linh Tran [1], Thai-Viet Dang [2*]

[1] School of Electrical and Electronics Engineering, Hanoi University of Science and Technology, Hanoi, Vietnam
[2] School of Mechanical Engineering, Hanoi University of Science and Technology, Hanoi, Vietnam
Email: [1] linh.tranhoai@hust.edu.vn, [2] viet.dangthai@hust.edu.vn
*Corresponding Author

Abstract—**Computer vision plays a significant role in mobile robot navigation due to the abundance of information extracted from digital images. On the basis of the captured images, mobile robots determine their location and proceed to the desired destination. Obstacle avoidance still requires a complex sensor system with a high computational efficiency requirement due to the complexity of the environment. This research provides a real-time solution to the issue of extracting corridor scenes from a single image. Using an ultra-fast semantic segmentation model to reduce the number of training parameters and the cost of computation. In addition, the mean Intersection over Union (mIoU) is 89%, and the high accuracy is 95%. To demonstrate the viability of the prosed method, the simulation results are contrasted to those of contemporary techniques. Finally, the authors employ the segmented image to construct the frontal view of the mobile robot in order to determine the available free areas for mobile robot path planning tasks.**

Keywords—*Artificial Intelligence; Computer Vision; Sematic Segmentation; Mobile Robot; Path Planning.*

I. INTRODUCTION

Mobile robots safely navigate their surroundings by identifying obstacles and moving objects in real time [1]. The sensor system in conjunction with the obstacle detection navigation algorithm consists of a laser scanner, sensor, and camera [2]. The use of lidar or digital cameras to provide information about a moving environment has become widespread in recent years [3, 4]. Despite the high cost and large number of computational steps [5, 6], it is still possible to perform the computations. Lidars provide depth information in all directions, preserving a world of perfect approximation. In addition, the camera provides inexpensive scene data for subject detection [7-14]. Due to the widespread availability of inexpensive, high-precision single cameras, the disadvantages that previously existed have been eradicated. Various forms of real-time image segmentation modeling and automatic navigation have been implemented successfully based on computer vision environment recognition techniques.

Vision-based indoor mobile robot navigation has gained popularity as a sensing method for autonomous navigation due to its ability to provide detailed information about the environment [15-18] that may not be obtainable via combinations of other types of sensors. Semantic segmentation is a computer vision technique that involves the partitioning of an image into different regions or segments based on their semantic meaning. These partitions would help understanding the area around to improve the movement

planning purposes. Semantic segmentation via deep learning (DL) is now a crucial task in computer vision, with applications including scene understanding, robotic perception, and image compression [19-26]. Semantic segmentation will precisely define semantic classifications such as buildings, transportation infrastructure, trees, and low vegetation as imaging technology advances [10-14]. Minae et al., [27] examined the interrelationships, advantages, and difficulties of these DL-based segmentation models. By contrasting the datasets, Li et al., [28] provided the essential methods of semantic segmentation for the fundamental datasets used in various structures. To remedy the lack of standard datasets for evaluating object segmentation. The 2D semantic labeling competition is recommended as a solution to this issue in [29]. In [30], Fusic et al. presented a DL algorithm for scene terrain categorization based on visual sensors. Even though the process obtains a high processing speed, its precision remains quite low when environmental factors change. Shelhamer et al., [31] converted modern classifier networks (AlexNet, VGG, and GoogLeNet) into fully convolutional networks (FCNs) to improve segmentation model performance and accuracy. Wang et al., used VGG-FCN models such as [30] or Unet [32] to address the image segmentation problem yields accurate results, but is unsuitable for infrastructure deployment due to their high computational cost. Rusli et al., [33] navigated mobile robot using the Canny edge algorithm and Hough line transform. There, boundary markers detected road markings and obstructions.

Multiple binary masks could be combined as a result of segmentation to segment the input image into distinct classes. In addition, multi-class semantic segmentation achieved remarkable results for mobile robot path planning in environments with numerous obstacles and complex topologies [30-33]. Analyzed image datasets, the proposed classification algorithm distinguishes between terrain and obstacles [30]. However, scene terrain classification-based mobile robot navigation had not been demonstrated. In general, the semantic segmentation and the path planning are done with high-power computing devices, which limit the application to big robots or to purely computer simulations. In this paper, to conserve memory resources and guarantee processing speed, we segmented available and unavailable regions using binary semantics, which make the segmentation model's architecture is now more streamlined and rapid, suitable for small computer boards, with quality assured. The smaller size of computing unit would also help

to reduce the overall dimensions of the robot, or to leave more loading capability for other tasks given to the robot.

Therefore, the authors devise a segmentation model using FCN as the decoder and MobilenetV2 as the encoder in order to solve the semantic segmentation problem [34, 35] and achieve efficient scene comprehension for autonomous driving. This combination will allow to perform the required tasks using only embedded computers while still achieving real-time performance quality. For the encoder, the authors employ a previously trained model from precursor networks. The authors then generate fractional predictions in the decoder block using multiscale fusion. The photography method captures authentic images of an indoor environment. Using published image data to compare with published segmentation methods, the proposed semantic segmentation model obtains an overall precision of 95%. In addition, the new image dataset from the Ducktown maintains its efficacy. The authors then effectively apply perspective correction to the segmented image to construct a frontal view of the general area, which detects the areas available for real-time movement. On the basis of the segmentation model's output, the authors can determine the areas that will serve as input for the autonomous mobile robot navigation system.

II. PROPOSED METHOD

A. Multi-class Semantic Segmentation based on MobilenetV2 Network

The preceding architectural design has inspired numerous variations. Using models such as LeNet and AlexNet [36], convolutional neutral networks (CNNs) have been shown to produce the most advanced results for image classification problems. Successive enhancements, such as VGGNet, GoogleNet, and ResNet [36, 37], have increased efficiency and efficacy. Eventually, the convolutional neural networks were developed into a full convolutional network. The authors create a network based on a full convolutional neural network (FCN) [38-43] to accomplish instantaneous pixel-by-pixel labeling while maintaining reliable segmentation results. VGG is preferred to AlexNet because the former model is more well-known but makes less accurate predictions. Semantic segmentation is a computer vision and image processing algorithm whose objective is to classify and segment each pixel in an image into classes corresponding to distinct semantic content. The authors implement a segmentation network for instant pixel-specific labeling that is primarily based on the FCN model's concept. The FCN network model for semantic segmentation is constructed as depicted in Fig. 1.

B. MobilenetV2

MobileNetV2's input channels are extended using 1 x 1 point convolution, as shown in Fig. 2. Then, use depth convolution for input linear feature extraction and linear convolutional integration to combine output features while shrinking the size of the network. It replaces Relu6 with a linear function after size reduction so that the output channel size matches the input.

In Fig. 3, In addition to Depthwise Separable Convolutions, Linear bottlenecks and Inverted Residual Block (shortcut links between bottlenecks) are suggested for usage

in MobileNetV2 [34]. Since the input and output of a block in a conventional residual architecture typically have more channels than the intermediary layers, MobileNet v2's residual block is the inverse of this design. To reduce the number of model parameters, the authors employ a depth-separated convolution transform and an inverted residual block between the layers. The method allows for the MobileNet model to be simplified while maintaining its functionality.
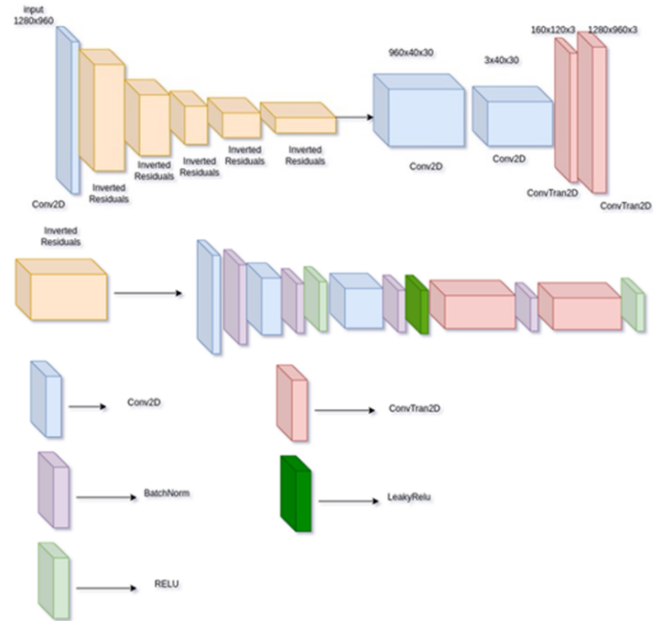

Fig. 1. The architecture of Multi-class Semantic Segmentation based on MobilenetV2

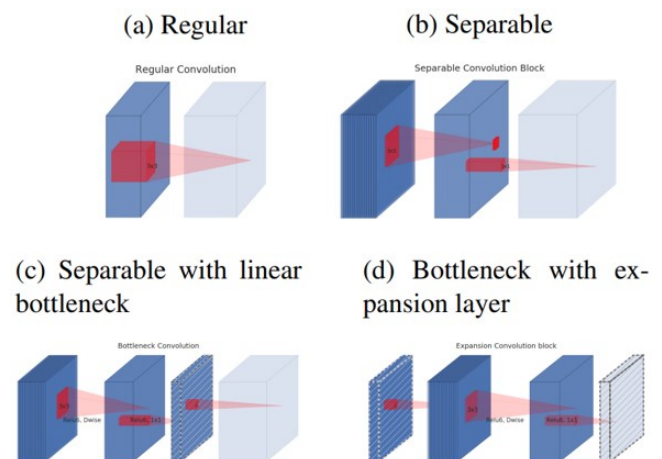

Fig. 2. The architecture of MobilenetV2

Fig. 3. Model of separable convolutional blocks in MobilenetV2

C. Model Training

The study's experiments were conducted on a server with the following specifications using Python 3.11.0 with Tensorflow 1.4 framework: CoreI7 11th CPU 2.50 GHz - 4.90 GHz, Nvidia 2080TI 12GB VRAM, 32 GB RAM, 64-bit OS, and English Windows 10 Home. The training specific parameters [44, 45] are set as follows: batch size is 4, learning rate is 0.001 and epochs is 100. The training is done by

optimizing the *Balanced Cross Entropy* (BCE) loss function [45, 46] expressed in Equation (1).

$$L_{Balanced-CE}(y, \hat{y}) = \{\beta \times y \log(\hat{y}) + (1 - \beta) \times (1 - y) \log(1 - \hat{y})\}$$

$$(1)$$

where $\hat{y}$ is the class SoftMax probability and y is the ground truth of the corresponding prediction, $\beta = 1 - \frac{y}{H \times W}$, and $H \times W$ presents the total of pixel in the image. The BCE loss function addresses the issue of class imbalance, where the number of samples from one class significantly outweighs the other class, leading to biased learning and poor performance on the minority class. Because of that, β is used for adjusting in the number of false negatives and false positives such as follows: reducing the number of false negatives when $\beta > 1$ or reducing the number of false positives when $\beta < 1$.

Finally, network parameters are optimized using Adam's optimum function [47-51] with a learning rate of 0.001 and 100 epochs in order to maximize the balanced cross-entropy [52-54] described by Equation (1). Gaussian blur [55, 56] and Gaussian noise [57, 58] are applied to the data set as pre-processing to assure the quality of the raw images before they are fed into the suggested segmentation model (see Fig. 4). Using the aforementioned approaches to generate more generic datasets improves the quality of the segmentation model, but at the expense of image quality.
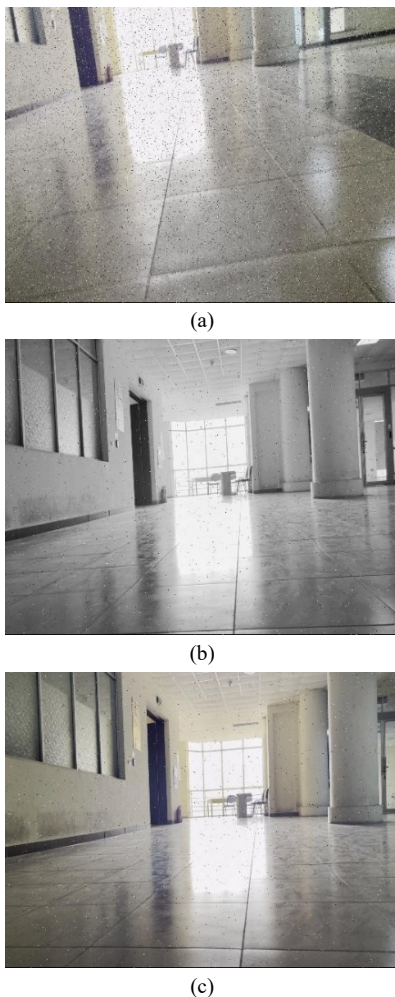

(a)

(b)

(c)

Fig. 4. The suitable steps from (a) to (c) of pre-processed real images

III. PROPOSED SYSEM

A. Hardware

The Jetson Nano 4GB developer AI includes the following components of the Artificial Intelligence Development Board: It has a 128-core Maxwell graphics processing unit, a quad-core ARM A57 at 1.43 GHz, 4GB of 64-bit LPDDR4 at 25.6 GB/s of memory, and 16GB of EMMC for storage. A version of NVIDIA Jetson Nano 77-Degree 1080p CSI Camera Module with 8-Megapixel IMX219 Camera is shown in Fig. 5.

Fig. 5. Jetson nano with 8MP Sony IMX219 77-degree CSI camera

B. Mobile Robot

The authors assess the transformation's efficacy and verify semantic segmentation's important function in building the frontal perspective of the floor. Then, the mobile robot's ideal path planning can be created. The experimental results support the effectiveness of the collision-free zone detection method. When the forward perspective is standardized, an all-encompassing plan for navigation and obstacle avoidance can be developed. Experiments depicted in Fig. 6 were done to evaluate our suggested sematic segmentation for use with four-wheel mobile robot navigation.
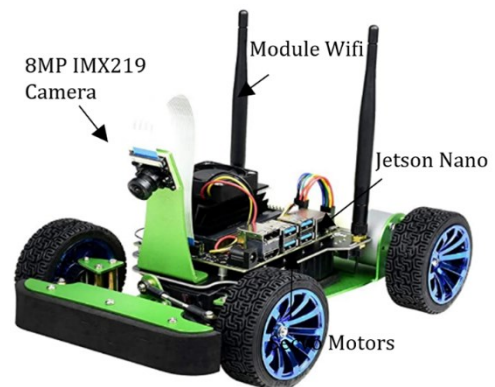

Fig. 6. Wheeled mobile robot

IV. EXPERIMENTAL RESULTS

A. Practical Face Recognition

The dataset used for training the semantic segmentation model is 1200 images from the library of Ducktown data set simulation software [59, 60]. In the Ducktown data set simulation software, there are full objects such as houses, cars, people, objects and roads, etc. Some images represent the environment moving around the robot as shown in Fig. 7.
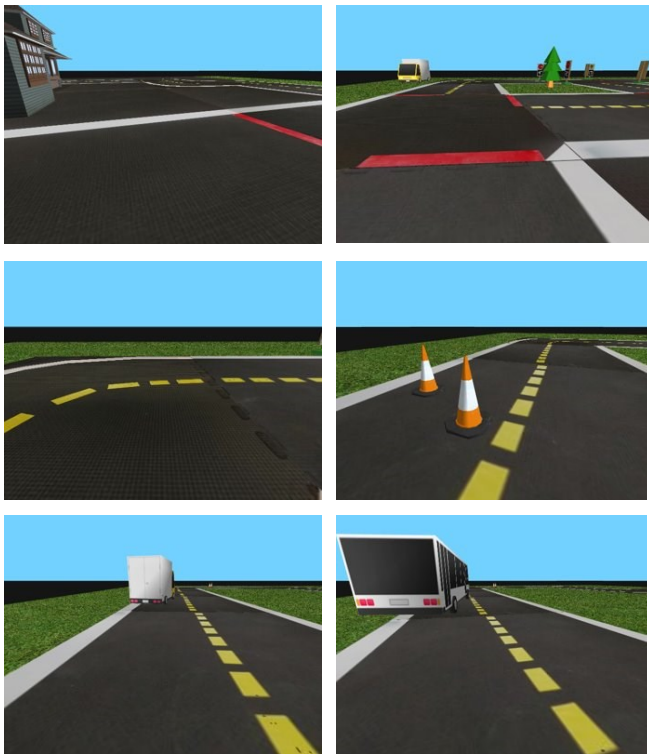
Fig. 7. The road, static obstacles and dynamic obstacles in Ducktown data set [59]

With the superfast structure of the proposed semantic segmentation model, the model training parameters are shown as shown in Fig. 8.

From the successful training of the semantic segmentation model, the robot vision system will generate the output after segmenting the acquired image to create two distinct regions, the area that allows movement (yellow) and the obstacle area (purple) as shown in Fig. 9.

Applying the proposed method of scaling from the size of the images of the environment to the actual size of the moving robot environment, the authors can apply the control model. Integrated PID control with tracing control moves from the navigation plan generated from the image processing results area of the semantic segmentation model.

The steering control plan for the control the mobile robot to drive two wheels at the same time, the remaining two wheels are responsible for balancing the mobile robot as shown in Fig. 10. With $vr(t)$: linear velocity of the right wheel; $vl(t)$: linear velocity of left wheel; $r(t)$: angular velocity of the right wheel; $1(t)$: angular velocity of right wheel; r: radius of each wheel; L: distance between two rudders; R: instantaneous radius of curvature of the robot trajectory, relative to the mobile robot's center axis; ICC: instantaneous center of the orbital curve; $R - L/2$: radius of curvature of the orbit described by the left wheel; $R + L/2$: radius of curvature of the orbit described by the right wheel.
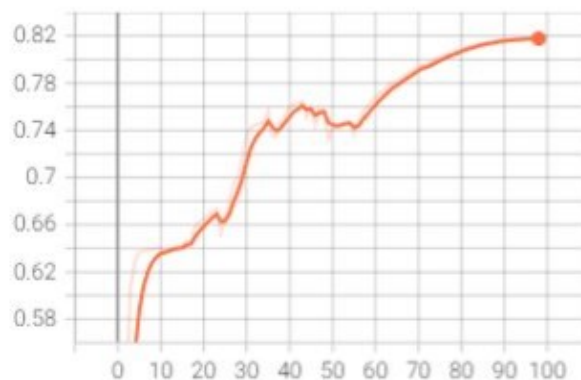

(a)

(b)
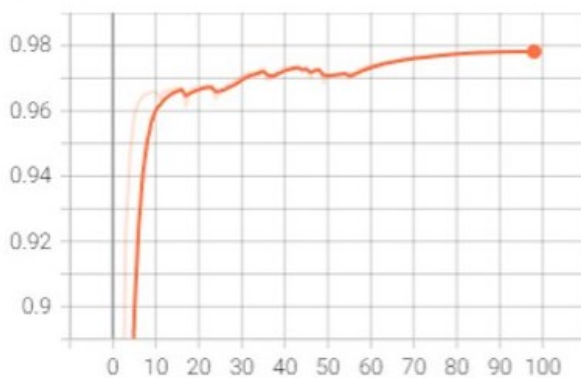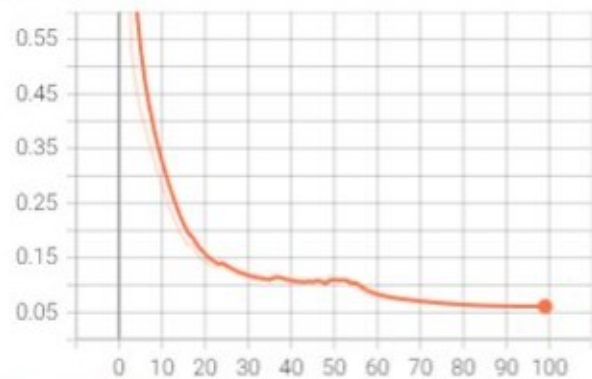
(c)

Fig. 8. The diagrams of training process based on semantic segmentation model with (a): trained accuracy, trained loss and trained mIoU
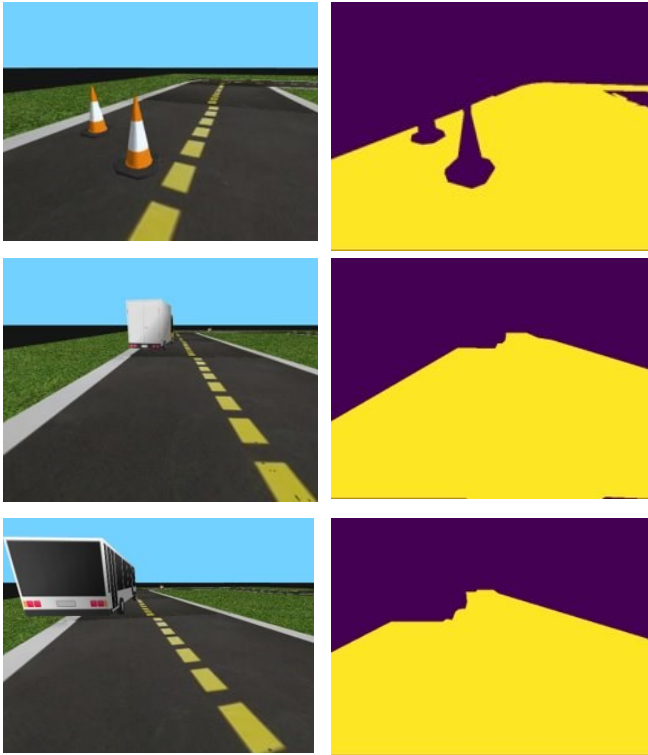
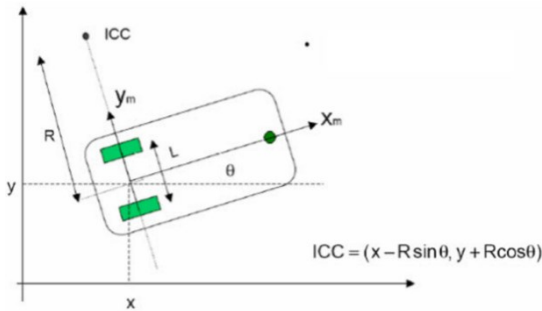Fig. 9. The output image of the semantic segmentation model from the image obtained from the frontal view

Fig. 10. The kinematics of mobile robot

Construct a linear state equation (2) describing the position of the vehicle.

$$\begin{bmatrix} v_x(t) \\ v_y(t) \\ \dot{\theta}(t) \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 \\ \sin\theta & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v(t) \\ \omega(t) \end{bmatrix}$$

$$= \begin{bmatrix} v(t)\cos\theta \\ v(t)\sin\theta \\ \omega(t) \end{bmatrix} \begin{bmatrix} v(t) \\ \omega(t) \end{bmatrix} \begin{bmatrix} \frac{1}{2}(v_r + v_l)\cos\theta \\ \frac{1}{2}(v_r + v_l)\sin\theta \\ (v_r - v_l)/L \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{2}\cos\theta & \frac{1}{2}\cos\theta \\ \frac{1}{2}\sin\theta & \frac{1}{2}\sin\theta \\ -1/L & 1/L \end{bmatrix}$$

(2)

Fig. 11 shows the PID closed-loop control structure for a self-propelled robot. When input θe is specified, the PID controller will give a control signal to ensure that the self-

propelled robot can follow the path in the navigation strategy. The control kinematics then calculate the velocity values for the wheels. In the next stage, the deflection angle, θe, will be feedback and compared with the set value, the deviation will be adjusted by the PID controller to ensure zero approach when the time t goes to infinity.
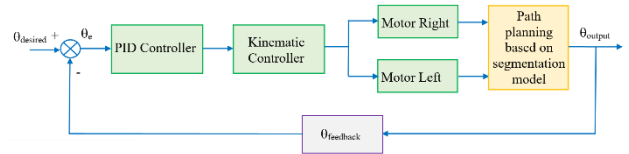

Fig. 11. PID control diagram for mobile robot tracking the path

B. Control Management Procedure

Fig. 12 depicts the results of the comparison between the PID controller and P with the motion of the vehicle. The self-propelled vehicle starts from the origin xOy to ensure that the motion follows the moving object. In Fig. 12, the changing of deflection angle corresponding to the PID controller has been really smaller than the P controller. With less than 0.12 rad, mobile robot has completely and smoothly tracked the path.
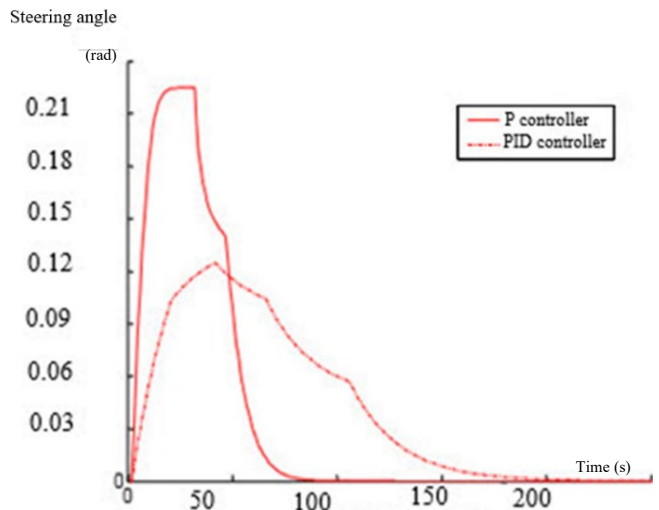

Fig. 12. The comparison between PID controller and P controller about the steering angle of mobile robot

V. CONCLUSIONS

This paper proposes an ultra-fast semantic segmentation model to reduce the number of training parameters and the cost of computation. In addition, the mean Intersection over Union (mIoU) is 89%, and the high accuracy is 95%. Based on segmentation results, the mobile robot's path planning is constructed successfully. The proposed models outperform state-of-the-art methods, which necessitate larger datasets for training, while using fewer resources in the training model.

CONFLICT OF INTEREST

The author declares no conflict of interest.

REFERENCES

[1] T. V. Dang and N. T. Bui, "Multi-scale Fully Convolutional Network based Semantic Segmentation for Mobile Robot Navigation," *Electronics*, vol. 12, no. 3, p. 533, 2022.

[2] V. Pagire and S. Mate, "Autonomous Vehicle using Computer Vision and LiDAR", *I-manager's Journal on Embedded Systems*, vol. 9, no. 2, pp. 7-14, 2021.

[3] Q. Luo, H. Wang, Y. Zheng, and J. He, "Research on path planning of mobile robot based on improved ant colony algorithm," *Neural Computing and Applications*, vol. 32, pp. 1555-1566, 2020.

[4] T. V. Dang and N. T. Bui, "Obstacle Avoidance Strategy for Mobile Robot based on Monocular Camera," *Electronics*, vol. 12, no. 8, p. 1932, 2023.

[5] A. Gianibelli, I. Carlucho, M. D. Paula, and G. G. Acosta, "An obstacle avoidance system for mobile robotics based on the virtual force field method," *2018 IEEE Biennial Congress of Argentina (ARGENCON)*, pp. 1-8, 2018, doi: 10.1109/ARGENCON.2018.8646065.

[6] P.G. Luan and N.T. Thinh, "Real-Time Hybrid Navigation System-Based Path Planning and Obstacle Avoidance for Mobile Robots," *Appl. Sci.,* vol. 10, no. 10, p. 3355, 2020.

[7] I. Maulana, A. Rusdinar, and R.A. Priramadhi, "Lidar applications for Mapping and Robot Navigation on Closed Environment," *J. Meas. Electron. Commun. Syst.*, vol. 4, no. 1, pp. 20-26, 2018.

[8] X. Pan, L. Gao, A. Marinoni, B. Zhang, F. Yang, and P. Gamba, "Semantic Labeling of High Resolution Aerial Imagery and LiDAR Data with Fine Segmentation Network," *Remote Sens.*, vol. 10, no. 5, p. 743, 2018.

[9] W. Xu, L. Yang, and S. Cao, "A Review of Semantic Segmentation Based on Context Information," *2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC)*, pp. 494-498, 2018, doi: 10.1109/ITOEC.2018.8740714.

[10] C. Peng, Y. Li, L. Jiao, Y. Chen, and R. Shang, "Densely Based Multi-Scale and Multi-Modal Fully Convolutional Networks for High-Resolution Remote-Sensing Image Semantic Segmentation," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 8, pp. 2612-2626, Aug. 2019, doi: 10.1109/JSTARS.2019.2906387.

[11] Y. Wang, Z. Sun, and W. Zhao, "Encoder- and Decoder-Based Networks Using Multiscale Feature Fusion and Nonlocal Block for Remote Sensing Image Semantic Segmentation," in *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 7, pp. 1159-1163, July 2021, doi: 10.1109/LGRS.2020.2998680.

[12] M. Pastorino, G. Moser, S. B. Serpico, and J. Zerubia, "Semantic Segmentation of Remote-Sensing Images Through Fully Convolutional Neural Networks and Hierarchical Probabilistic Graphical Models," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-16, 2022, doi: 10.1109/TGRS.2022.3141996.

[13] M. Oršic' and S. Šegvic, "Efficient semantic segmentation with pyramidal fusion," *Pattern Recognition*, vol. 110, p. 107611, 2021.

[14] A. Gupta, A. Anpalagan, L. Guan, and A. S. Khwaja, "Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues," *Array*, vol. 10, p. 100057, 2021.

[15] R. Leonard, N. Brilly, and R. Rusman, "Vision-based vanishing point detection of autonomous navigation of mobile robot for outdoor applications," *J. Mechatron. Electr. Power Veh. Technol.,* vol. 12, pp. 117-125, 2021.

[16] P. Wan, S. Gao, L. Li, B. Sun, and S. Cheng, "Obstacle Avoidance Path Planning Design for Autonomous Driving Vehicles Based on an Improved Artificial Potential Field Algorithm," *Energy*, vol. 12, p. 2342, 2019.

[17] S. Feraco, S. Luciani, A. Bonfitto, N. Amati, and A. Tonoli, "A local trajectory planning and control method for autonomous vehicles based on the RRT algorithm," *2020 AEIT International Conference of Electrical and Electronic Technologies for Automotive (AEIT AUTOMOTIVE)*, pp. 1-6, 2020, doi: 10.23919/AEITAUTOMOTIVE50086.2020.9307439.

[18] D. Qin, "Path Planning Algorithm Based on Visual Image Feature Extraction for Mobile Robots," *Mob. Inf. Syst.*, vol. 2022, 2022.

[19] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, "Greedy layer-wise training of deep networks," *Advances in Neural Information Processing Systems*, vol. 19 pp. 153-160, 2007.

[20] D. Erhan, P.A. Manzagol, Y. Bengio, S. Bengio, and P. Vincent, "The difficulty of training deep architectures and the effect of unsupervised pretraining," *Proceedings of the 12th IC on Artificial Intelligence and Statistics (AISTATS'09)*, vol. 5, pp. 153-160, 2009.

[21] N. A. Othman, M. U. Salur, M. Karakose and I. Aydin, "An Embedded Real-Time Object Detection and Measurement of its Size," *2018 International Conference on Artificial Intelligence and Data Processing (IDAP)*, pp. 1-4, 2018, doi: 10.1109/IDAP.2018.8620812.

[22] Y. Xiao, Z.Q. Tian, J.C. Yu, Y.S. Zhang, S. Liu, S. Du, and X. Lan, "A review of object detection based on deep learning," *Multimedia Tools and Applications*, vol. 79, pp. 23729-23791, 2020.

[23] A. Ammar, A. Koubaa, M. Ahmed, and A. Saad, "Aerial Images Processing for Car Detection using Convolutional Neural Networks: Comparison between Faster R-CNN and YoloV3," *Preprints.org, arXiv: 2019100195*, 2019.

[24] D. Michael, *U.S. Patent No. 5,872,870*. Washington, DC: U.S. Patent and Trademark Office. 1999.

[25] Z. Ren, F. Fang, N. Yan, and Y. Wu, "State of the art in defect detection based on machine vision," *International Journal of Precision Engineering and Manufacturing-Green Technology*, vol. 9, no. 2, pp. 661-691, 2022.

[26] Y. Li *et al*., "A Deep Learning-Based Hybrid Framework for Object Detection and Recognition in Autonomous Driving," in *IEEE Access*, vol. 8, pp. 194228-194239, 2020, doi: 10.1109/ACCESS.2020.3033289.

[27] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image Segmentation Using Deep Learning: A Survey," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 7, pp. 3523-3542, 2022, doi: 10.1109/TPAMI.2021.3059968.

[28] B. Li, Y. Shi, Z. Qi, and Z. Chen, "A Survey on Semantic Segmentation," *2018 IEEE International Conference on Data Mining Workshops (ICDMW)*, pp. 1233-1240, 2018, doi: 10.1109/ICDMW.2018.00176.

[29] ISPRS Working group III/4. Isprs 2D Semantic Labeling Contest. Available online: http://www2.isprs.org/commissions/comm3/wg4/ (accessed on 1 May 2023).

[30] S.J. Fusic, K. Hariharan, R. Sitharthan, and S. Karthikeyan, "Scene terrain classification for autonomous vehicle navigation based on semantic segmentation method," *Trans. Inst. Meas. Control.,* vol. 44, pp. 2574-2589, 2022.

[31] W. Sun and R. Wang, "Fully Convolutional Networks for Semantic Segmentation of Very High Resolution Remotely Sensed Images Combined With DSM," in *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 3, pp. 474-478, March 2018, doi: 10.1109/LGRS.2018.2795531.

[32] W. Xiaolei, H. Zirong, H. Shouhai, H. Mei, X. Le, and Z. Xiang, "A deep learning method for optimizing semantic segmentation accuracy of remote sensing images based on improved UNet," *Scientific Report,* vol. 13, no. 1, p. 7600, 2023.

[33] L. Rusli, B. Nurhalim, and R. Rusyadi, "Vision-based vanishing point detection of autonomous navigation of mobile robot for outdoor applications," *Journal of Mechatronics, Electrical Power, and Vehicular Technology*, vol. 12, pp. 117-125, 2021.

[34] T. V. Dang, "Smart Attendance System based on Improved Facial Recognition," *Journal of Robotics and Control*, vol. 4, no.1, pp. 46-53, 2023.

[35] T. V. Dang, "Smart home Management System with Face Recognition based on ArcFace model in Deep Convolutional Neural Network," *Journal of Robotics and Control*, vol. 3, no. 6, pp. 754-761, 2022.

[36] B. Michalski and M. P-. Wójcik, "Comparison of LeNet-5, AlexNet and GoogLeNet models in handwriting recognition," *Journal of Computer Sciences Institute*, vol. 23, pp. 145-151, 2022.

[37] E. Suherman, B. Rahman, D. Hindarto, and H. Santoso, "Implementation of ResNet-50 on End-to-End Object Detection (DETR) on Objects," *Sinkron: jurnal dan penelitian teknik informatika*, vol. 8, no. 2, pp. 1085-1096, 2023

[38] A. A. Süzen, B. Duman, and B. Şen, "Benchmark Analysis of Jetson TX2, Jetson Nano and Raspberry PI using Deep-CNN," *2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, pp. 1-5, 2020, doi: 10.1109/HORA49412.2020.9152915.

[39] Y. Nan, J. Ju, Q. Hua, H. Zhang, and B. Wang, "A-MobileNet: An approach of facial expression recognition," *Alexandria Engineering Journal*, vol. 61, no. 6, pp. 4435-4444, 2021.

[40] Z. Hong, Z. Zhi-hui, L. Ying, C. Xiao-feng, L. Xiaonan, Z. Jie, S. Kwong, Jun, and Z. Jun, "Local Binary Pattern-Based Adaptive Differential Evolution for Multimodal Optimization Problems", *IEEE Transaction on Cybernetics*, vol. 50, no. 7, pp. 3343-3357, 2020.

[41] S. Biswas, S. Kundu, and S. Das, "Inducing Niching Behavior in Differential Evolution Through Local Information Sharing," in *IEEE Transactions on Evolutionary Computation*, vol. 19, no. 2, pp. 246-263, 2015, doi: 10.1109/TEVC.2014.2313659.

[42] F. Caraffini, A. V. Kononova, D. Corne, "Infeasibility and structural bias in differential evolution," *Infomation Science*, vol. 496, pp. 161–179, 2019.

[43] Z. -J. Wang *et al.*, "Dual-Strategy Differential Evolution With Affinity Propagation Clustering for Multimodal Optimization Problems," in *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 6, pp. 894-908, Dec. 2018, doi: 10.1109/TEVC.2017.2769108.

[44] X. Li, D. Chang, T. Tian and J. Cao, "Large-Margin Regularized Softmax Cross-Entropy Loss," in *IEEE Access*, vol. 7, pp. 19572-19578, 2019, doi: 10.1109/ACCESS.2019.2897692.

[45] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley Interscience, 2006.

[46] J. Brownlee, *Probability for Machine Learning: Discover How To Harness Uncertainty With Python*. Machine Learning Mystery, 2019.

[47] Q. Tang, M. Lécuyer, "DP-Adam: Correcting DP Bias in Adam's Second Moment Estimation," *Preprints.org, arXiv:2304.11208*, pp. 1-6, 2023.

[48] J.R. Sashank, K. Satyen, K. Sanjiv, "On the Convergence of Adam and Beyond," *Preprints.org, arXiv:1904.09237v1*, pp. 1-23, 2018.

[49] G. Chen, C. K. Qu, and P. Gong, "Anomalous diffusion dynamics of learning in deep neural networks," *Neural Networks*, vol. 149, pp. 18-28, 2022.

[50] D. P. Kingma and J. B. Adam, "A method for stochastic optimization," *Preprints.org, arXiv:1412.6980v9*, 2014.

[51] W. Liu, Y. Wen, Z. Yu, and M. Yang, "Large-Margin Softmax Loss for Convolutional Neural Networks," *Proceedings of the 33rd International Conference on Machine Learning*, vol. 48, pp. 507-516, 2016.

[52] P. Reverdy and N. E. Leonard, "Parameter Estimation in Softmax Decision-Making Models With Linear Objective Functions," in *IEEE Transactions on Automation Science and Engineering*, vol. 13, no. 1, pp. 54-67, Jan. 2016, doi: 10.1109/TASE.2015.2499244.

[53] I. Goodfellow, Y. Bengio, and A. Courville, "Softmax Units for Multinoulli Output Distributions," *Deep Learning*, pp. 180–184, 2016.

[54] B. Gao and L. Pavel, "On the Properties of the Softmax Function with Application in Game Theory and Reinforcement Learning," *Preprints.org, arXiv:1704.00805*, pp. 1-10, 2017.

[55] T. Lindeberg, "A time-causal and time-recursive scale-covariant scale-space representation of temporal signals and past time," *Biological Cybernetics*, vol. 117, pp. 21–59, 2023.

[56] A. C. Bergstrom, D. Conran, and D. W. Messinger, "Gaussian Blur and Relative Edge Response," *Preprints.org, arXiv:2301.00856v1*, pp. 1-12, 2023.

[57] R. Legin, A. Adam, Y. Hezaveh, and L. P. Levasseur, "Beyond Gaussian Noise: A Generalized Approach to Likelihood Analysis with Non-Gaussian Noise," *The Astrophysical Journal Letters*, vol. 949, no. 2, 2023.

[58] I. M. Bright, M. L. R. Meister, N. A. Cruzado, Z. Tiganj, E. A. Buffalo, and M.W. Howard, "A temporal record of the past with a spectrum of time constants in the monkey entorhinal cortex," *Proceedings of the National Academy of Sciences*, vol. 117, no. 33, pp. 20274–20283, 2020.

[59] C.B. Maxime, G. Floria, C. Yanjun, M. Bhairav, P. Liam, "Duckietown Environments for OpenAI Gym," 2018. https://github.com/duckietown/gym-duckietown (accessed on 1 May 2023).

[60] Z. Lorincz, *Imitation Learning in The Duckietown environment*., PhD thesis, Budapest University of Technology and Economics, Faculty of Electrical Engineering and Informatics, 2020.