Estimating Traffic Density in Uncertain Environment: A Case Study of Danang, Vietnam

Doan Phuoc Mien*, Tran The Vu, Ngo Van Si

Abstract—The road traffic monitoring system in Vietnam faces significant challenges due to the impact of uncertain factors. The problems of blurred lanes, many types of vehicles in traffic, the view from remote surveillance cameras, etc, directly affect the intelligent traffic management system. We propose an algorithm to estimate the traffic density in an uncertain environment to identify and solve the above problems. Through empirical evaluation of the data set from cameras collected on the website http://0511.vn, it has been proved that this proposed method has high performance in real-time and has an equivalent accuracy to other high-complexity methods.

Index Terms—estimation, density, the traffic flow, the uncertain environment.

1. Introduction

According to the publication in [1], it is estimated to 2050 that the number of vehicles participating in traffic in the developing countries will increase 12 times compared to developed countries. Therefore, the traffic situation will be more and more complicated and difficult to manage and regulate the traffic flow. Currently, there is a significant amount of research focused on estimating traffic density. In these researches, we had an understanding of the research situation at domestic and foreign countries to prove that our research direction is appropriate.

Nowadays, the traffic problem has been widely researched and has had many practical applications in life. Since 1994, there have been many published researches such as [2] that have created a model to estimate the traffic density in the Southern California area. Most of the vehicle detection methods participating in traffic focus on analyzing the useful information for the traffic management such as the real-time traffic density and the number of vehicle types on the roads. This paper presents the vehicle classification methods by SVM method and calculate the traffic density in case of the uncertain environment such as: the abnormal change about weather, many various vehicle types participating in traffic. In recent years, the number of surveillance cameras has been increasing over the roads in inner cities. Therefore, this data source is abundant to enable the traffic safety stations to use an effective way.

In fact, there are very few applications to be applied to the real life because of some reasons: low camera quality, inaccurate results when the environment is changed due to affecting of light, rain, wind and so on. In addition, the cultural elements of traffic participants is also one of the reasons why it has not been able to deploy applications for intelligent traffic systems.

In the research [3] selected Mixture of Gaussian (MoG), based on the adaptive learning algorithm. To build a model containing the background pixels, MoG uses K Gaussian components including the weight, the average value and the deviation. To track changes in the new frame, the Gaussian components of both frames are compared. The Gaussian components of pixel are updated based on the learned coefficient from new pixels. If there are no any Gaussian component fits the new pixel value, this pixel is considered foreground.

In the research [4], it is proposed to use an IoT model by using sensors to identify and count the number of cars, drivers, walkers, trams, subways and ferries. The proposed model will be applied in real-time monitoring of traffic flow based on the collected data from sensors through pre-processing. And then, it will being made a decisions based on the collected and analyzed data. The proposed model with purpose reduces traffic jams on highways and crowded streets.

In the above researches, it is focused on clearly demarcated lanes, highway sections, so it has not been mentioned about the uncertain elements of traffic in Vietnam such as many missed lane markings, many various vehicles participating in traffic, lane encroachment situation, passing through red-light and so on. In this paper, we will estimate the traffic density in the uncertain environment like the roads in Da Nang, Vietnam.

Doan Phuoc Mien is with Tra Vinh university, Tra Vinh, Vietnam.

Tran The Vu is with the University of Danang - VN-UK Institute for Research and Executive Education, Vietnam.

Ngo Van Si is with Vietnam Research Institute of Electronics, Informatics and Automation Danang, Vietnam.

**Corresponding author: Doan Phuoc Mien (e-mail: phuocmien@tvu.edu.vn)*
Manuscript received February 13, 2023; revised April 10, 2023; accepted May 04, 2023.
Digital Object Identifier 10.31130/jst-ud.2023.045E

TABLE 1: Some traffic density estimation methods

Research scope	Related work
Domestic research	In the research [5] estimated traffic density through images from video. The input image will be processed through stages such as lane detection, lane area calculation. At the same time, they use YOLOV3 model to detect vehicles and calculate the total of area of vehicles. The final step is that give a conclusion about the density based on the calculated ratio of the area.
Research from other countries	In the research [6], [7] used the HOG and LBP feature and the learning method based on the vector machine (SVM) to estimate the traffic density. Some researchers [8], [9] use the reduced Pyramid CNN model. After that, it will use CNN networks to estimate the context at the different levels in order to limit the counting error and the density map with the better quality. These two solutions achieved high performance and both of them use architecture based on multi-column model (MCNN) and density level classification. However, there are some disadvantages in the methods as: (1) Multi-column CNNs are difficult to train according to the training method described in the research [10]. Bulky network structure requires more time to train. (2) Multi-column CNNs have a prophylactic structure, so the different columns seem to do the same and have no the discernible difference. (3) Both of two solutions require density level classification before sending images in MCNN. (4) These jobs devote a large part of the parameters to density level classification to label the input regions instead of allocating parameters for the final density map generation.

2. Implementation Methodologies

The data is the traffic videos that we have collected. It is analyzed to find uncertain elements such as the lane compliance elements when participating in traffic, the quietly small vehicles in the frame... so we propose a generalized processing model as Fig. 1, it will be presented details in section 3. The main research of this paper is the traffic density estimation through those image frames. The input image will be processed through the following stages: determining the lane, calculating the lane area. Besides that, we will use YOLO to detect the vehicles and calculate the total area of the vehicles. We will calculate the ratio and make a conclusion. Our program follows the overall model 1 for processing.

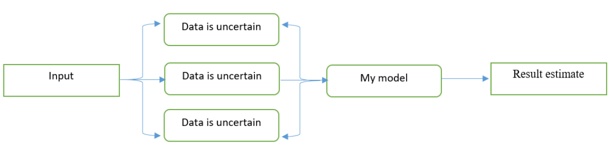

Fig. 1: The overal model

Some uncertain factors in traffic are shown in Fig. 2 including environmental elements, infrastructures, traffic participants and vehicles participating in traffic. In this paper, we will focus on analyzing uncertain elements such as lane encroachment situation, lane occupancy for parking to estimate the density accurately.

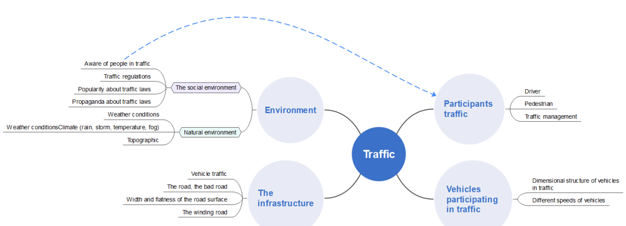

Fig. 2: The environmental elements

2.1. Proposed model

Our proposed model is shown in Fig. 3. The input data will be the extracted frames from the video. Then,

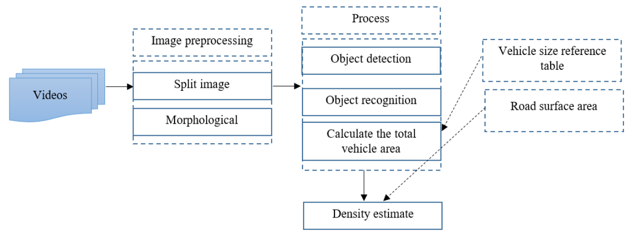

Fig. 3: Proposed model

the objects are extracted and recognized. Next, depending on the type of object, we will match that object with the reference table to calculate the size of the object. Besides that, the road surface area is also determined. From there, we will estimate the density accurately.

2.2. Data preprocessing

2.2.1. Data Collection

In general, there are many types of vehicles participating in traffic on the streets in Vietnam. Particularly in Da Nang, there are several types of vehicles in circulation on the street such as: cycle rickshaw, bus, garbage truck and car with 5-seat to 7-seat and 40-seat and 80-seat. The dimensions of types of vehicles are presented in Table 2.

TABLE 2: Vehicle size reference table (Unit: mm)

Vehicle	Lenght	Width	Height
Motobike	1900	600	1050
Motocycle	2050	725	1102
Car	3700	1550	1450
Garbage truck	5430	1780	2100

2.2.2. Splitting image frames and processing images

The input is a video recording of the vehicle on the road, this video can be formatted by any type of video file. In this step, we proceed to extract each frame from the video, transform this image into a digital image object for input in the next step. OpenCV provides the

Mat library class to support this. Mat can see as a class with two parts of data:

- Matrix header contains information such as: matrix size, storage method, storage address
- A cursor to a matrix containing pixel values (using multidimensional space based on storage method). The header size of the matrix is unchanged. However, the size of the matrix itself can change from one image to another and is usually larger by order of intensity. To calculate the length and width of the road surface, we determine based on surveillance camera. However, each camera has its own direction of view, so we evaluate the camera deviation to change the appropriate weight for determining the length and width of each observation road area such as Nguyen Hue School Gate, the intersection in Tay Rong Bridge. Each intersection will be identified in the area outlined in red. Because the camera keeps the same direction of rotation, we assign each camera to an own coordinate. Specifically at intersections, the determination of the area will be shown as Fig. 4.

Fig. 4: The determination of the area

2.3. Detecting and identifying objects participating in traffic

To identify objects participating in traffic, we experiment by deep learning method with distributed dataset as Fig. 5.

The number of vehicles in each lane during the week is shown in Figures 6, 7, and 8.

2.3.1. Training

To train the model shown in Fig. 9, we use video from 0511.vn (file Camer-aQT.MP4, CameraCR.MP4...) to generate the input images. After we got about 3000 images, of which 2100 images are used for training and 900 images are used for testing. We started using Yolo to assign the automatic label for the objects such as: motorcycles, cars, garbage trucks and other vehicles. We do not keep the vehicle dimensions but we will refer to the dimensions in Table 2.

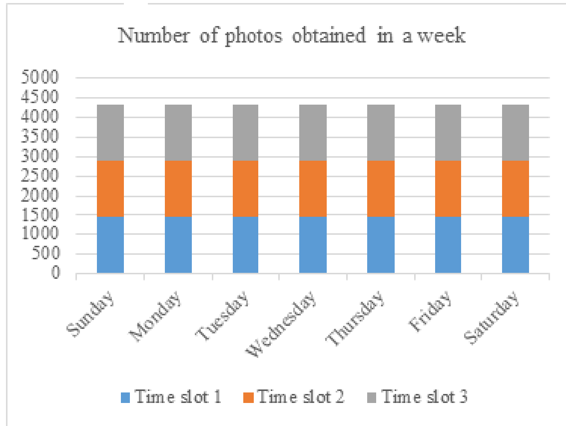

Fig. 5: Number of photos obtained in a week

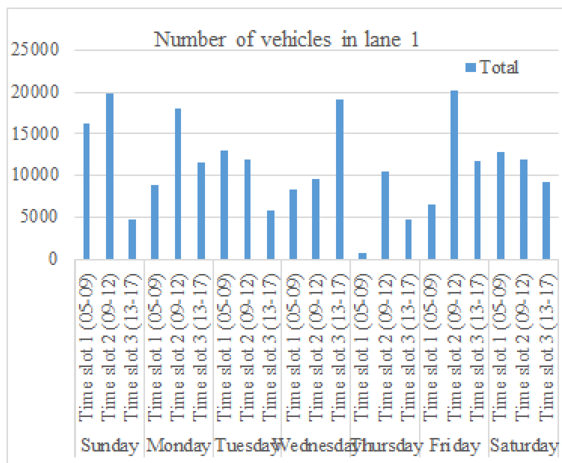

Fig. 6: Number of vehicles in lane 1

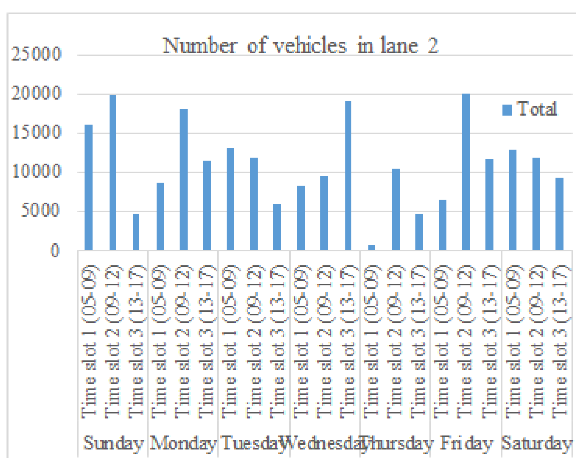

Fig. 7: Number of vehicles in lane 2

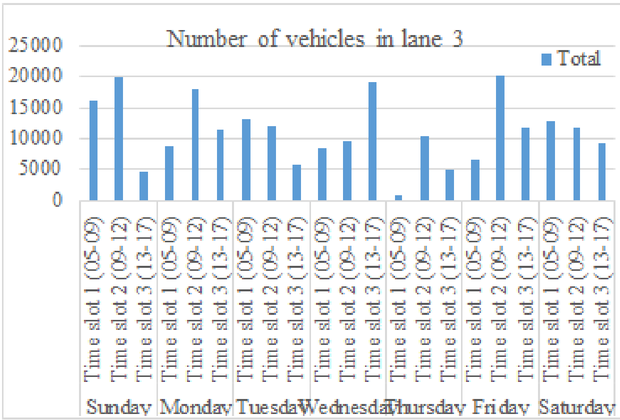

Fig. 8: Number of vehicles in lane 3

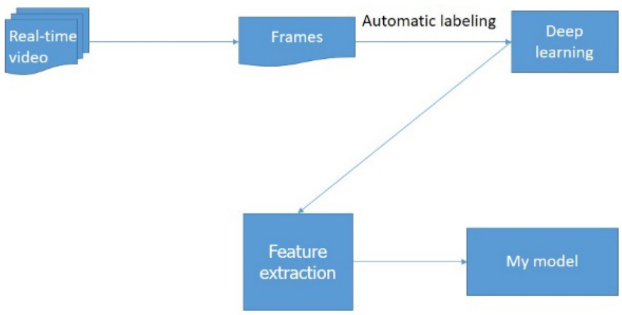

Fig. 9: Training model

Fig. 10: Training dataset

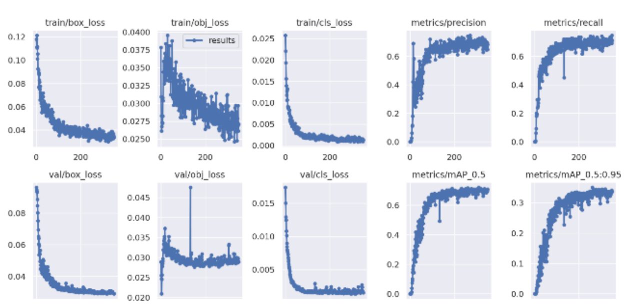

Fig. 11: Modeling training results

2.3.2. Identification.

Fig. 12 shows the procedure applied to carry out the traffic vehicle identification process. We use video from 0511.vn (CameraQT.MP4, CameraCR.MP4, Tay Rong Bridge...). Each frame in the video will be identified with the built labels from the training model: empty, car, motorcycle, rudimentary car, bus, ambulance, police car.

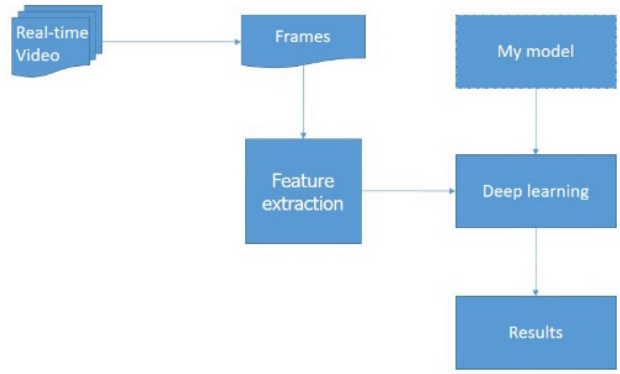

Fig. 12: Object recognition model

In the traffic environment in Vietnam, especially in the areas of roads that have not been upgraded or repaired, the lane identification is not possible. Sometimes, there are cases of vehicles encroaching on other lanes, rudimentary car participating in the traffic, so applying our proposed model to estimate density results are shown in Fig. 13 and Fig. 14.

Fig. 13: Medium density

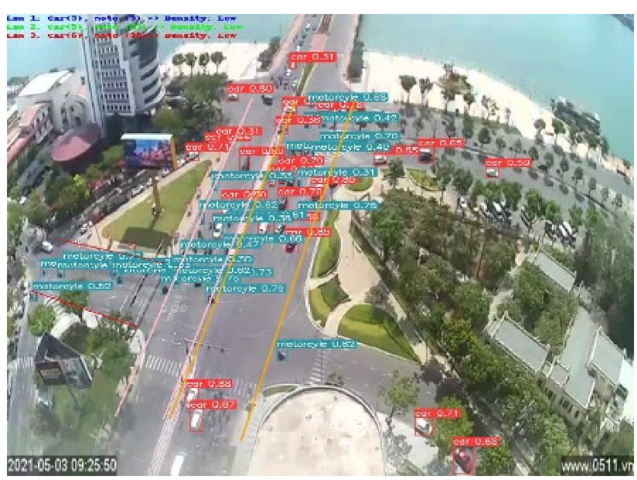

Fig. 14: Low density

2.4. Calculating total of vehicle area

The total of vehicle area determined on the road (TL) and the road surface area (DT) will be calculated to know the level of vehicle moving on the whole road. The formula to calculate TL is as:

$$TL = \sum_{i}^{n} (X_i * SL_i) \tag{1}$$

With i = motocycle (m), car (o)

$$X_m = H_m * W_m \tag{2}$$

$$X_o = H_o * W_o \tag{3}$$

By H is the height of the vehicle and W is the width of the vehicle; SL_m is the number of motobikes and SL_o is the number of the car.

The foumula for calculating DT is as follows: DT = Ls * Ws (4), where Ls is the restricted length of the observation path and Ws is the restricted width of the observation line.

The density ratio is calculated as follows: R = TL/DT * 100

Use this R ratio to divide traffic flow into 3 levels: low (empty roads – or few cars), medium (high vehicle) and high (congested roads). Levels are presented in Table 3. At this level, we base on experiments to create the optimal range for the classification.

TABLE 3: Density reference table

Num	Density	Ratio (R)
1	Low	R<30%
2	Medium	<=30 R <80%
3	High	R>=80%

3. Traffic density estimation

The density estimation model is shown in Fig. 15 The input data is the images extracted from the video and features to serve the classification of the objects. From the training model is applied to give the results of vehicle identification and direct reference to the corresponding size table. Based on the total of vehicle size and the monitored road surface area, the application will give a density ratio including 1 in 3 states such as low, medium and high (It is shown in Table 3).

4. Experimental results

4.1. Configuration settings

We tested on the machine with CPU configuration: Intel core i9 10900x 3.7g, mainboard: Asus ws x299 pro workstation, Ram: gskill trident z rgb 64g/3000 (4x16g), Hard drive: Samsung 970evo 1tb nvme m.2 pcie, vga: nvidia quadro rtx 5000 16g.

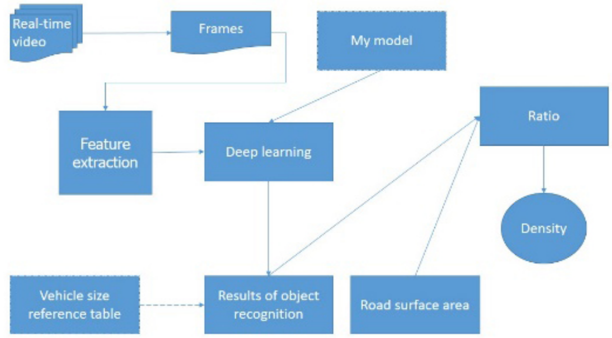

Fig. 15: Calculate traffic density

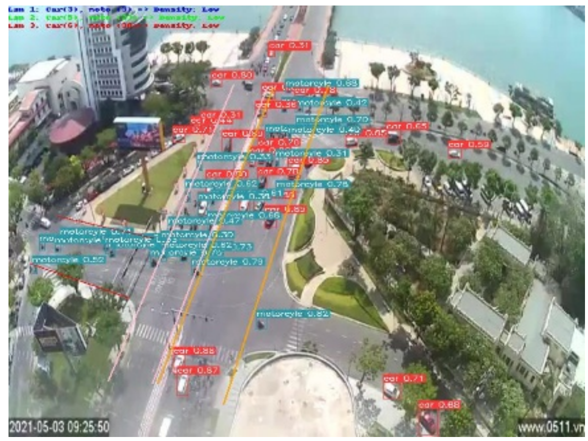

Fig. 16: Result of vehicle indentification by camera at West Dragon Bridge

4.2. Identification results

Fig. 16 shows that we split into 3 lanes. With the identification results in each lane that the corresponding density will be calculated.

4.3. Comparison

YOLOv5 and SVM are two popular algorithms in the field of small object detection. YOLOv5, or You Only Look Once version 5, is a real-time object detection algorithm that uses a convolutional neural network (CNN) to perform object detection. On the other hand, Support Vector Machine (SVM) is a classical machine learning algorithm that is widely used for binary classification.

In terms of performance, YOLOv5 generally outperforms SVM in terms of accuracy and speed. YOLOv5 uses a deep learning approach which allows it to learn complex patterns and relationships between objects in an image. This results in a higher accuracy in detecting small objects compared to SVM, which is limited by its linear decision boundary. Additionally, YOLOv5 is able to process images in real-time, making it suitable for real-time applications.

However, SVM has its own advantages over YOLOv5. SVM is a simpler algorithm, which makes it easier to interpret and understand the decision-making process. Additionally, SVM is less computationally expensive compared to YOLOv5, which requires a large amount of computing resources and time to train the CNN.

In conclusion, both YOLOv5 and SVM have their strengths and weaknesses in small object detection. The choice between the two algorithms ultimately depends on the specific requirements of the task and the resources available.

The results of using SVM for wrong lane detection are promising. The implementation of the SVM algorithm has shown to effectively classify the different lane patterns and accurately identify instances of vehicles driving in the wrong lane. This helps to improve the overall traffic management and safety on the roads. In addition, SVM has the ability to handle high-dimensional data, making it well-suited for this application where multiple factors such as vehicle type, speed, and road conditions need to be considered. However, further research and testing are necessary to validate the performance and robustness of the SVM algorithm in real-world conditions.

5. Conclusion

In this paper, we provide a solution to estimate traffic density from determining the road surface area, vehicle area based on the proposed method to identify vehicles participating in traffic on the specified road surface. The proposed method is suitable in uncertain traffic conditions such as: having many types of vehicles on the road, many blurred lanes, the encroachment on the lane of vehicle. The application of the media size reference helps the for the increasing density estimation with the correct ratio. The solution is implemented for real-time videos.

References

[1] K. Abdulrahim and R. A. Salam, "Traffic surveillance: A review of vision based vehicle detection, recognition and tracking," *International journal of applied engineering research*, vol. 11, no. 1, pp. 713–726, 2016.

[2] L. Muñoz, X. Sun, R. Horowitz, and L. Alvarez, "Traffic density estimation with the cell transmission model," in *Proceedings of the 2003 American Control Conference, 2003.*, vol. 5. IEEE, 2003, pp. 3750–3755.

[3] M. F. Chowdhury, M. R. A. Biplob, and J. Uddin, "Real time traffic density measurement using computer vision and dynamic traffic control," in *2018 Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision & Pattern Recognition (icIVPR)*. IEEE, 2018, pp. 353–356.

[4] P. Agarwal, P. Matta, and S. Sharma, "Analysis based traffic flow control decision using iot sensors," *Materials Today: Proceedings*, vol. 46, pp. 10 707–10 711, 2021.

[5] L. Q. P. Vu, M. L. Le, and N. M. T. Tran, "Uoc luong mat do giao thong dua vao hoc sau va dien tich," *The conference on information technology and its applications*, pp. 164–171, 2019.

[6] H. Lee, J. Lee, and Y. Chung, "Traffic density estimation using vehicle sensor data," *Journal of Intelligent Transportation Systems*, vol. 26, no. 6, pp. 675–689, 2021.

[7] D. Prasad, K. Kapadni, A. Gadpal, M. Visave, and K. Sultanpure, "Hog, lbp and svm based traffic density estimation at intersection," in *2019 IEEE Pune Section International Conference (PuneCon)*. IEEE, 2019, pp. 1–5.

[8] D. B. Sam, S. Surya, and R. V. Babu, "Switching convolutional neural network for crowd counting," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2017, pp. 4031–4039.

[9] V. A. Sindagi and V. M. Patel, "Generating high-quality crowd density maps using contextual pyramid cnns," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 1861–1870.

[10] Y. Zhang, D. Zhou, S. Chen, S. Gao, and Y. Ma, "Single-image crowd counting via multi-column convolutional neural network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 589–597.

Doan Phuoc Mien is lecturer in the Department of Information Technology at Tra Vinh University. He is currently a postgraduate student at the University of Danang. His research areas include Image Processing, Machine Learning and Internet of Things.

Tran The Vu Vice Dean of Academic Affairs Office, the University of Danang (UD) - VN-UK Institute for Research and Executive Education (VN-UK IREE), 158A Le Loi St., Danang, Vietnam. Tran The Vu achieved his Ph.D. degree in Information systems and technology at University of Wollongong in 2014. He specialised in digital ecosystem for smart transportation. Dr. Tran The Vu experts on context awareness system, machine learning and Internet of things in general and in specific in smart transportation. Further, he has a solid relationship network with researchers and specialists in several institutes and industries working well on multi-discipline projects.

Ngo Van Si was born on July 16th, 1961 in Duy Chau, Duy Xuyen, Quang Nam. He obtained a Bachelor's degree in Radio Physics in 1983 from Hue University of Sciences, a Master's degree in Electronics and Telecommunications in 1991 from Hanoi University of Technology, and a PhD in Electronic Engineering in 1996 from Hanoi University of Technology. He retired in 2022.