Urban Road Marker Classification Using HOG and LBP Features and Artificial Intelligence

Zamani Md Sani[1], Liang Jin Chuan[2], Hadhrami Abd Ghani[3]

[1,2] Universiti Teknikal Malaysia Melaka
[3] Universiti Malaysia Kelantan
zamanisani@utem.edu.my

Abstract. Road markers provide road information to driver to ensure road and their safety. Different types of markers indicate different kinds of information. Accidents may occur if the drivers do not follow the rules associated with the road markers or the road markers are not seen clearly by the drivers. This paper proposed a vision-based system to classify three types of markers using image processing and artificial neural network (ANN). The length of the feature vector on HOG and LBP are the features extracted and use in training neural network pattern recognition tool. The result shows an accuracy of 99.4% with HOG and LBP features as input vectors.

Keywords: Road marker, image processing, feature extraction, artificial neural network and classification.

1 Introduction

In Malaysia, road accident happens every year and has become a concerning issue. It is one from the top five principal causes of death in Malaysia between 2017 and 2018 [1] as shown as in Figure 1. Based on the road accident data in Malaysia, there is a relation between the population and the death rate. The amount of road deaths increases from year to year due to the growth population in Malaysia [2].

Road marker used as indicator on a road surface to give information. They are widely placed with the road marking machines. Besides, they are applied in other facilities used by vehicles to mark parking spaces or designate areas for other purposes. It is also used to show regulation for parking and stopping. It plays important role on urban roads because they ensure the road safety and make the flow of travel paths smooth. They are also used on roadways to provide guidance to the road users such as drivers and pedestrians. Different types of marker indicate different kind of information to minimize the confusion and uncertainty about their meaning for example double solid line road marker indicates that overtaking of vehicles is not allowed. There are several causes leading to the road accidents. For example, human error, condition of the road and vehicle problem [3][4]. The main reason of road accidents was caused due to people driving recklessly and ignoring the traffic rules. Even though there is law enforcement and camera installation on the road to punish and fine those who break the rules, this incident still often happens. To improve the situation, researchers have been conducting research towards the autonomous car driving. In this paper, three types of markers are being classified which are the double dashed, dashed and double solid by using image processing and artificial neural network (ANN).

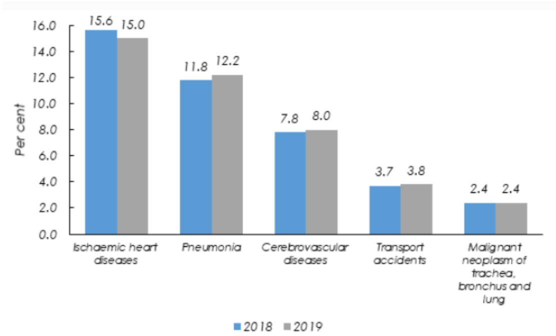


Fig. 1 Top Five Principal Causes of Death in Malaysia Between 2018 and 2019

1.1 Previous Research

Paula et al. [5] presented an automatic classification technique to classify five types of road markers. Paula's approach for lane marker detection used between three to five features extracted from the image, and later applied the Bayesian classifier for dashed, single dashed and double solid markers, while the second stage differentiated between dashed-solid and solid-dashed lines. However, the results of Paula's classification were found to be prone to abrupt changes in each frame causing inconsistent results while driving. Furthermore, zooming into the confusion matrix table results indicated that the classes for solid-dashed and dashed-solid had lower accuracy compared to the rest of the classes.

Two types of lane marker had been performed by Toan et al. [6] which are the dashed and solid lanes under various environmental condition. Using perspective camera model work, the adaptive threshold is applied to measure the distance in the static images and use it to detect the final line segments of the left and right boundaries of the road lane. In the work from Zhang [7], a robust lane detection of using two Stage Feature Extraction with YOLO able to detect and the lane markers but the classification markers types were not performed.

Interestingly in the works of Tang et al. [8] and Ali et al [9] worked on 6 and 5 types of lane marker classification which are the most like Paula's work and based on the numbers of markers being classified. In Tang et al. the semantic information of the markers is extracted, and the decision trees model is used. Whereas in Ali's the classification method is using the seed fill algorithm to classify it markers. Nevertheless, both works did not comprehensively compare their marker classification methods with Paula in terms of intensive accuracy using the same videos provided from Paula.

In the continuous research, Zamani et. al [10][11][12] worked on road marker classification, initially with only two types of markers using features and Artificial Neural Networks, and later expanding to five types of marker classification using the features extracted from customised Region of Interest. Nevertheless, for a real-time classification process, the classification algorithm needs to know the temporal values prior to classification and requires many images of the same types of marker for validation purposes to confirm transitions. The need for so many images would delay the classification process and potentially cause lags in the driver alerts.

Through the review, it is hard to find a standard to classify the types of markers due to the different types are available in the different countries. Thus the 3 different types of markers in the urban are marks for the classification process.

2 Methodology

2.1 Hardware Preparation and Setup

Camera positioning is crucial as incorrect position of camera mounting will result in recording failure. The camera is mounted same angle as the x-axis, there may be some illumination issue on the video frame when strong sunlight is projected into the view of camera. The best angle is to mount the camera slightly lower about 5 to 10 degree with respect to the x-axis as shown as in Figure 2. Image acquisition tool in MATLAB is used with Logitech C310. The acquisition parameters such as RGB format and resolution at 1021 x 720 with 30 frame per second.

Fig. 2 Camera Position at The Car

2.2 Pre-Processing Module

There are 5 stages of processing on the sample frames which are grayscale conversion, image sharpening, noise removal, cropping and resizing as shown in Figure 3. This preprocessing module is created to make sure the process of feature extraction smooth and easier.

Fig. 3 Stages of the Pre-processing Methods

The first step to process the frame is to convert the frame colour from RGB format to grayscale format. The step is known as grayscale conversion and often widely used in image processing. Next, all the image samples are sharpened and becoming clear compared to before. The principle of it is to extract the high frequency components by subtracting the blurred version from the original. Noise removing will be applied to the images to remove any unwanted noise from the image and once all the process had been completed, it will crop to the region of interest for the feature extraction. The results for the pre-processing as shown as in Figure 4.

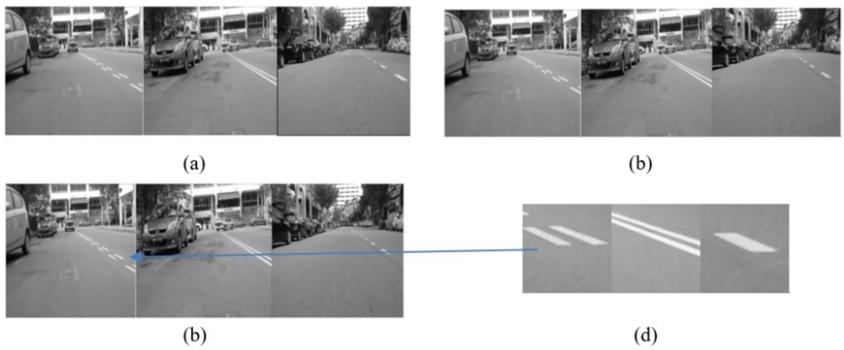


Fig. 4 Results from the Pre-Processing Module

2.3 Feature Extraction

After the sample images going through the preprocessing module, the output image samples are ready for the feature extraction. In this research, two types of feature extraction have been used, Histogram of Oriented Gradient (HOG) and Local Binary Pattern (LBP). These two types of feature extraction are known as common method to extract the feature from the image and the feature data are in the form of feature vector. The only difference between these two methods is the value in the feature vector. For HOG, the value in the feature vector is float number whereas the value in the feature vector for LBP is scalar. Example of the sample frames on HOG visualisation are shown in Figure 5.

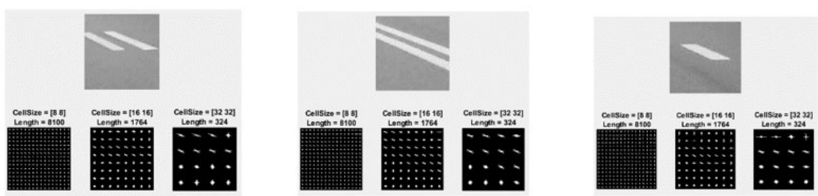


Fig. 5 HOG Visualization for Different Markers

2.4 Training and Classification Module

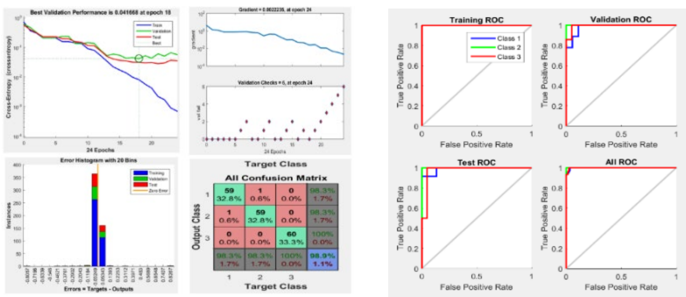
Features extracted are used to build models for accurate classification. Artificial neural-network (ANN) method was used to classify the type of the road lane markers. In this case, supervised learning was used for training purpose. This research aimed to classify 3 types of road marker which meant there were 3 types of target data.

3 Results and Discussion

3.1 Training and Testing Set Images

In this research, a total number of 180 image frames are selected. 60 image frames for each type of road markers are selected. These 180 selected image frames are stored in database to be used in training and classification process. For using HOG as feature, a set of 60 samples are used for each type of road marker and 180 samples for 3 classes of road marker. The extracted 180-by-1764 features vector is used as an ANN input data set for training, testing, and validating the network.

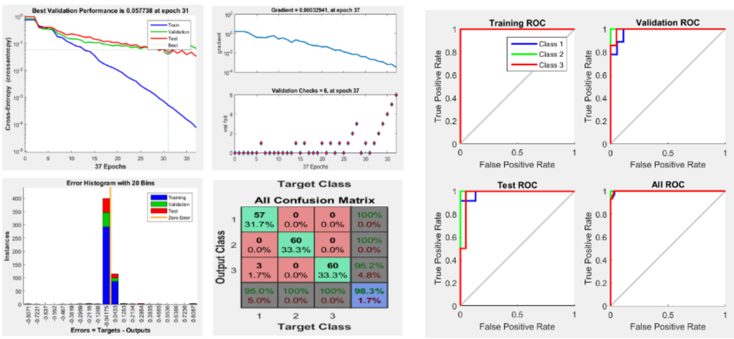
Figure 6(a) shows the HOG as feature for neural network's performance, training state, error histogram, and overall confusion matrix. From the performance plot, the cross-entropy error is maximum at the beginning of training. For this proposed system, the best validation performance is at epoch 18, and at this point the cross-entropy error is very close to zero. On the training state plot, the maximum validation checks 6 at epoch 24 and at this point, the neural network halts the training process to give best performance. The error histogram plot represents that the error of this system is very close to zero. An overall confusion matrix is three sets of combined confusion matrices, which are the training confusion matrix, validation confusion matrix, and testing confusion matrix. This overall confusion matrix plot shows 98.9% correct classification for this system. Receiver Operating Characteristic (ROC) curve of the network which illustrates true positive rate verses false positive rate at various threshold settings of the network, is shown in Figure 6(b) Area under the curve (AUC) shows a slightly result for this proposed system. At the neural network train, test and validation conclusion, this network performs 98.9% correct classification of 3 classes of road marker.



(a) (b)

Fig. 6 (a) Neural Network (HOG): Performance, Training State, Error Histogram, and Overall Confusion Matrix (b) HOG Network Receiver Operating Characteristic (ROC)

For using LBP as feature, the extracted 180-by-256 features vectors are used as an ANN input data set for training, testing, and validating the network. Figure 7(a) shows the LBP as feature for neural network's performance, training state, error histogram, and overall confusion matrix. From the performance plot, the cross-entropy error is higher at the beginning of training compared to HOG. For this system, the best validation performance is at epoch 31, and at this point the cross-entropy error is closer and more stable to zero compared to HOG. On the training state plot, the maximum validation checks 6 at epoch 37 and at this point, the neural network halts the training process to give best performance. The error histogram plot represents that the error of this system is very close to zero. The overall confusion matrix plot shows 98.3% correct classification for this system. Receiver Operating Characteristic (ROC) curve of the network which illustrates true positive rate verses false positive rate at various threshold settings of the network, is shown in Figure 7(b). Area under the curve (AUC) shows a slightly perfect result for this proposed system. At the neural network train, test and validation conclusion, this network performs 98.3% correct classification of 3 classes of road marker.



(a) (b)

Fig. 7(a) Neural Network (HOG): Performance, Training State, Error Histogram, and Overall Confusion Matrix (b) HOG Network Receiver Operating Characteristic (ROC)

For using HOG and LBP as feature, the extracted 180-by-2020 features vectors are used as an ANN input data set for training, testing, and validating the network. Figure 8(a) shows the HOG and LBP as feature for neural network's performance, training state, error histogram, and overall confusion matrix. From the performance plot, the cross-entropy error is higher at the beginning of training compared to HOG. For this system, the best validation performance is at epoch 23, and at this point the cross-entropy error is closer and more stable to zero compared to HOG. On the training state plot, the maximum validation checks 6 at epoch 29 and at this point, the neural network halts the training process to give best performance. The error histogram plot represents that the error of this system is very close to zero. The overall confusion matrix plot shows 99.4% correct classification for this system. Receiver Operating Characteristic (ROC) curve of the network which illustrates true positive rate verses false positive rate at various threshold settings of the network, is shown in Figure 8(b) Area under the curve (AUC) shows a slightly perfect result for this proposed system. At the neural network train, test and validation conclusion, this network performs 99.4% correct classification of 3 classes of road marker.

(a) (b)

Fig. 8 (a) Neural Network (HOG): Performance, Training State, Error Histogram, and Overall Confusion Matrix (b) HOG Network Receiver Operating Characteristic (ROC)

Based on the Table 1, it is clearly that the combination of HOG and LBP feature extraction method will result in higher accuracy in the classification process.

Table 1 Accuracy of the classification determined by feature extraction method

Feature Extracted	Accuracy (%)
HOG	98.9
LBP	98.3
HOG and LBP	99.4

3.2 Video Testing

After getting the trained network, the system was tested with new video input on the real time situation. Different types of road markers situation were tested with the trained network. The accuracy and result were shown in Table 2 and Table 3.

Table 2 Accuracy of the classification on real time video

Experiment	Type	Frame	Accuracy
1	Double Dashed	91	41.76%
2	Double Solid	89	100.00%
3	Single Dashed	180	68.33%

Table 3 Confusion matrix of the overall accuracy on real time video

	1	2	3	
1	38 (10.56%)	35 (9.72%)	18 (5%)	41.76%
2	0	89 (24.72%)	0	100.00%
3	48 (13.33%)	9 (2.5%)	123 (34.17%)	68.33%
	44.19%	66.92%	87.23%	69.44%

Based on Table 2, three experiments were carried out on a recorded video. Each experiment was tested with the trained network on each type of the road marker. In experiment 1, the type of the road marker was double dashed and 91 samples of frames were tested and classified. The accuracy of the classification on real time for experiment 1 was 41.76%. In experiment 2, the type of the road marker was double solid and 89 samples of frames were tested and classified. The accuracy of the classification on real time for experiment 2 was 100.00%. In experiment 3, the type of the road marker was single dashed and 180 samples of frames were tested and classified. The accuracy of the classification on real time for experiment 3 was 68.33%.

In Table 3, the first three diagonal cells showed the number and percentage of correct classifications by the trained network. For example, 38 samples were correctly classified as Class 1 which was double dashed road marker. This corresponded to 10.56% of all 360 samples. Secondly, 89 samples were correctly classified as double solid road marker. This leaded to 24.72% of all 360 samples. Similarly, 123 samples were correctly classified as single dashed road marker. This corresponded to 34.17% of all 360 samples.

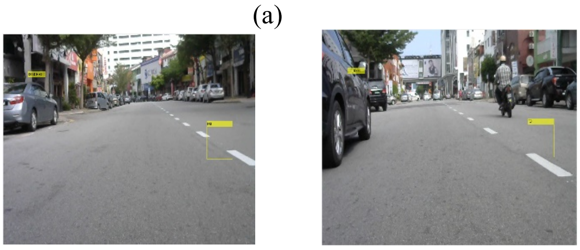
35 samples of the double dashed markers were incorrectly classified as double solid marker. This corresponded to 9.72% of all 360 samples in the data. 18 samples of the double dashed markers were wrongly classified as single dashed marker. This corresponded to 5% of all data. At the same time, 48 samples of single dashed markers were wrongly classified as double dashed marker and this corresponded to 13.33% of all data. Other 9 samples of single dashed marker were incorrectly classified as double solid marker which corresponded to 2.5% of all the samples.

Out of 91 double dashed marker predictions, 41.76% were correct and 58.24% were wrong. Out of 89 double solid marker predictions, all 89 samples were predicted correctly which indicated 100% correct. Out of 180 single dashed marker predictions, 68.33% were correct and 31.67% were wrong.

On the other hand, out of 86 double dashed marker cases, 44.19% were correctly predicted as double dashed marker and 55.84% were wrongly predicted. Out of 133 double solid marker cases, 66.92% were correctly classified as double solid marker and 33.08% were classified as other type of road marker. Out of 141 single dashed marker cases, 87.23% were correctly classified as single dashed marker and 12.77% were classified as other type of road marker. Overall, 69.44% of the predictions were correct and 30.56% were wrong.

In addition, there are also some errors. Example of these errors are shown in Figure 9(b) showed that the classification on the road without the marker is classified as double solid marker. This is due to the trained network had only 3 classes of target markers which are double dashed marker, double solid marker and single dashed marker as shown in example of Figure 9(a). The road without marker is not included in the classification process. Due to this reason, the feature extracted from the image in Figure 9(b) may have almost the same value as in feature vector in double solid class and so it is classified in the double solid class even though it is not, which is not accurate.

(a)



(b)

Fig. 9 Testing on Video and Error on Recorded video (a) With Correct Results (b)With Wrong Results

4 Conclusion

Among all the feature extraction methods, the result of the combination of HOG and LBP as the feature extraction method showed the highest accuracy which was 99.4%. The testing was also carried out using a trained network against the real time situation. The result of classification accuracy on the real time road marker of double dashed, double solid and single dashed markers are 41.76%, 100% and 68.33% respectively. The overall accuracy of the classification on road marker is 69.44%. However, there are some errors in the real time situation which is due to the region selection. The errors can be solved by implementing detection system on the road marker in the future.

References

1. Department of Statistics Malaysia Official Portal “Social Statistic Bulletin, 2019” 26 November 2019. https://www.dosm.gov.my/
2. ASEAN Road Safety Centre, “General Road Accident Data in Malaysia (1997 – 2016)”, Malaysian Institute of Road Safety Research (MIROS) official website, 2017.
3. Subramaniam, S. K., Ganapathy, V. R., Subramonian, S., & Hamidon, A. H. (2010). Automated traffic light system for road user’s safety in two lane road construction sites. WSEAS Transactions on Circuits and Systems, 2(9), 71-80.
4. Gannapathy, V. R., Subramaniam, S. K., Diah, A. M., Suaidi, M. K., & Hamidon, A. H. (2008). Risk factors in a road construction site. Proceedings of the World Academy of Science, Engineering and Technology, 46, 640-643.
5. de Paula, M. B., & Jung, C. R. (2015). Automatic Detection and Classification of Road Lane Markings Using Onboard Vehicular Cameras. IEEE Transactions on Intelligent Transportation Systems, 16(6), 3160–3169. http://doi.org/10.1109/TITS.2015.2438714.
6. Hoang, Toan Minh, Hyung Gil Hong, Husan Vokhidov, and Kang Ryoung Park, ‘Road Lane Detection by Discriminating Dashed and Solid Road Lanes Using a Visible Light Camera Sensor’, Sensors (Switzerland), 16.8 (2016) <https://doi.org/10.3390/s16081313>
7. Zhang, Xiang, Wei Yang, Xiaolin Tang, and Jie Liu, ‘A Fast-Learning Method for Accurate and Robust Lane Detection Using Two-Stage Feature Extraction with YOLO V3’, Sensors (Switzerland), 18.12 (2018) <https://doi.org/10.3390/s18124308>
8. Tang, Luliang, Tuo Deng, and Chang Ren, ‘Semantic Information Extraction of Lanes Based on Onboard Camera Videos’, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 42.3 (2018), 1643–49 <https://doi.org/10.5194/isprs-archives-XLII-3-1643-2018>
9. Ali, Abduladhem Abdulkareem, and Hussein Alaa Hussein, ‘Real-Time Lane Markings Recognition Based on Seed-Fill Algorithm’, ACM International Conference Proceeding Series, 2019, 190–95 <https://doi.org/10.1145/3321289.3321306>
10. Z. M. Sani, H. A. Ghani, R. Besar and L. W. Sen, (2015), "Real-time daytime road marker recognition using features vectors and neural network," IEEE Conference on Sustainable Utilization and Development In Engineering and Technology (CSUDET),1-6. doi: 10.1109/CSUDET.2015.744622
11. Z. M. Sani, H. A. Ghani, R. Besar, A. Azizan., H. Abas, (2018) Real-Time Video Processing using Contour Numbers and Angles for Non-urban Road Marker Classification, International Journal of Electrical and Computer Engineering (IJECE) 4(8), 2540-2548. DOI: 10.11591/ijece.v8i4