International Journal of Electrical and Computer Engineering (IJECE)
Vol. 12, No. 4, August 2022, pp. 4206~4216
ISSN: 2088-8708, DOI: 10.11591/ijece.v12i4.pp4206-4216

Efficient lane marking detection using deep learning technique with differential and cross-entropy loss

Abdullah Al Mamun[1], Poh Ping Em[1], Md. Jakir Hossen[1], Anik Tahabilder[2], Busrat Jahan[3]
[1]Faculty of Engineering and Technology, Multimedia University, Melaka, Malaysia
[2]Department of Computer Science, Wayne State University, Detroit, Michigan, USA
[3]Department of Computer Science and Engineering, Feni University, Feni, Bangladesh

Article Info

Article history:

Received Oct 18, 2020
Revised Mar 29, 2022
Accepted Apr 18, 2022

Keywords:

Deep neural network
Differential loss
Driver-assistance systems
E-Net
Lane marking detection

ABSTRACT

Nowadays, researchers are incorporating many modern and significant features on advanced driver assistance systems (ADAS). Lane marking detection is one of them, which allows the vehicle to maintain the perspective road lane. Conventionally, it is detected through handcrafted and very specialized features and goes through substantial post-processing, which leads to high computation, and less accuracy. Additionally, this conventional method is vulnerable to environmental conditions, making it an unreliable model. Consequently, this research work presents a deep learning-based model that is suitable for diverse environmental conditions, including multiple lanes, different daytime, different traffic conditions, good and medium weather conditions, and so forth. This approach has been derived from plain encode-decode E-Net architecture and has been trained by using the differential and cross-entropy losses for the backpropagation. The model has been trained and tested using 3,600 training and 2,700 testing images from TuSimple, a robust public dataset. Input images from very diverse environmental conditions have ensured better generalization of the model. This framework has reached a max accuracy of 96.61%, with an F1 score of 96.34%, a precision value of 98.91%, and a recall of 93.89%. Besides, this model has shown very small false positive and false negative values of 3.125% and 1.259%, which bits the performance of most of the existing state of art models.

Corresponding Author:

Abdullah Al Mamun
Faculty of Engineering and Technology, Multimedia University
Ayer Keroh, Melaka-7540, Malaysia
Email: mamun130203@gmail.com

1. INTRODUCTION

The collision of automobiles is considered to be one of the most fundamental causes of physical injury and sudden death on the road. According to the World Health Organization (WHO), about 1.2 M people have died, and a significant portion of the world's 100 M people have been severely injured due to a sudden road crash, resulting in substantial economic losses for the nation. In the world as a whole, the number of traffic crash deaths is rising day after day. A summary of accidental death in south Asian countries is depicted in Figure 1. According to the figures for 2013 and 1016, Thailand ranked first in traffic accidents rate among the Asian countries, and those caused about 7,152 innocent people to die [1].

Intelligence inspection and vision inspection for autonomous driving are one of the growing research domains in the field of artificial intelligence. The recent revolution in computer vision technology has strengthened and enriched the application of artificial intelligence (AI) in an intelligent investigation, resulting in a large revolution in self-driving vehicles. Thus, a thousand others are spared from death [2]. Autonomous driving research has been the center of attraction because it effectively reduces the risk of road accidents [3]. Especially advanced driver-assistance systems (ADAS) have been a very popular and growing research domain targeting apprehension, protection, and environmental awareness around vehicles [4].

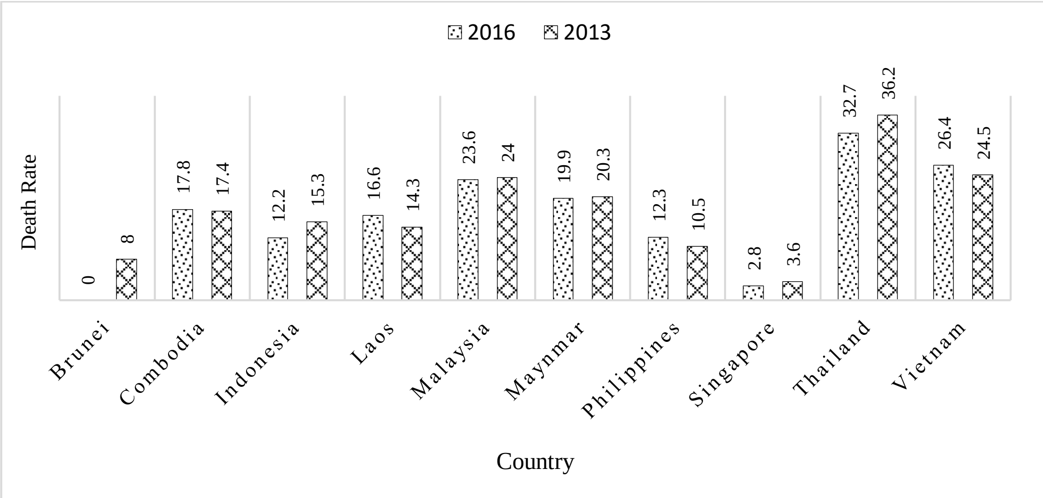

Figure 1. Death records in various Asian countries due to road accidents in 2013 and 2016

Lately, there has been a huge advancement in autonomous vehicles technology, including semi-autonomous driving technology, vehicle to vehicle communication, self-parking technology, lane markings detection, vehicle to infrastructure communication, lane-departure detection technology, and so on [5]. US National Highway traffic safety administration has declared lane markings detection as the primary requirement for any autonomous vehicles [6]. Besides, lane mark identification is considered to be the most significant innovation in road condition analysis by any autonomous vehicle [7]. In addition, knowing the lane position helps to avoid any accident that occurs during the lane changing [8]. Again, the application of the lane markings determination is not only for lane-keeping improvement but also for implementing traffic regulations based on the road's lane markings [9].

Even though a lot of research has been done on landmark detection, there are still many unsolved challenges, especially for autonomous driving in very diverse environmental conditions [10]. Factors like occlusion, rain, shadow, fog, and sunlight create a lot of illusions and result in a challenging and utterly unknown environment taught to be handled by autonomous vehicles [11]. Diverse computer vision technologies like [12] and [13] are being implemented, targeting various problems. Existing methods mostly use the handcrafted and highly specialized feature that suffers from computational complexity and struggle to truckle diverse environmental conditions [11].

2. RELATED WORK

Though lane marker detection is a key study issue for self-driving automobiles, it may be difficult and time-consuming under a variety of situations and impacts [14]. As a result, a lot of research is being conducted to develop more precise, accurate, and reliable lane mark detection technology [15]. Human error causes thousands of innocent deaths, including pedestrians and other drivers during driving and especially during lane changing. To solve this problem, computer vision techniques like hough transformation [16], edge detection [17], template matching [18] are being implemented based on low-level features including color, texture, and so on. But none of these techniques is perfect due to the constraints of light, shades, clouds, weather, and environmental changes [19]. In many cases, texture feature is used in landmark detection, and computer vision technologies like [20] and [21] are frequently being used with machine learning classifiers, including AdaBoost [22], support vector machine (SVM) [23], and so on, which sometimes results in the inefficient output. On the other hand, some viable methods, like [24], are slow to apply in real-life applications, and others [25] suffer from a lack of accuracy. So, to overcome both of the abovementioned limitations, researchers are using deep learning models to develop a fast and accurate solution. Levi *et al.* [25] used convolution neural network (CNN) to solve the detection problem by splitting the images into a distinct column called SitxelNet and improving by setting sleekness to constrain between the neighbors. However, it was not suitable for all the cases. The specific affinity between the segmentation and recognized boundary is insignificant and has fewer numerical results than other methods [25]. In their research, Mamun *et al.* demonstrated a Seg-Net-based model for lane marker detection, though it has an overfitting problem and only focuses on lane space [26]. He *et al.* had developed a unique CNN-based solution to alleviate the issues under various illumination effects. The divergence of the road markings influenced the result of lane markings detection [27]. Huang combined the spatial and temporal data in the CNN framework to detect the lane markings by selecting the lane boundaries [28].

Nevertheless, this makes up for low illumination conditions, such as night and rainy times. Li aims to develop a more sophisticated method for detecting the low-level lanes in the traffic scene by applying a noble framework of the hierarchical neural network, the fusion of CNN and recurrent neural networks (RNN). However, it vastly depends on the target, and it will collapse if it misses the target detection [29]. Also, the whole process is time-consuming as it did not use any preprocessing for the algorithm [30]. Nguyen *et al.* have developed a fully connected network (FCN) based model for detecting multi-lane boundary features and hence determining multiple lanes in the road. But, this model still has a bit higher false-positive rate, and it fails to detect the lane if an identical obstacle is positioned nearby the lane [10]. Aiming at real-time performance with efficiency, Paszke *et al.* proposed an agnostic lane detection process based on an E-net [31] and Neven *et al.* layout [32]. Though it improves the application in real-time, it is still unable to provide higher performance in the distinct traffic scene condition [11]. The embedding loss-driven generative adversarial network (GAN) model has been introduced to avoid the computational cost and complexity in the pixel-wise model. The successful detection rate varies significantly because this model can only work in fixed scenarios. The presented model is based on a basic encode-decode E-Net architecture that was trained to employ differential and cross-entropy loss. The model has been trained and tested with a robust publicly available dataset entitled TuSimple and contains more than 3,600 training images and 2,700 testing images belonging to very diverse weather and environmental conditions, including diversity in daytime weather light, shades, rain, and so on.

3. RESEARCH METHODOLOGY

The proposed deep learning approach provides a novel technique for landmark determination for automated vehicles. It is a simple deep learning model that has been derived from the basic encoder-decoder-based model inspired by E-Net [31]. A customized ENet architecture has been trained by the processed TuSimple data with a combination of discriminative and cross-entropy losses. The complete layout of the proposed model is shown in Figure 2.

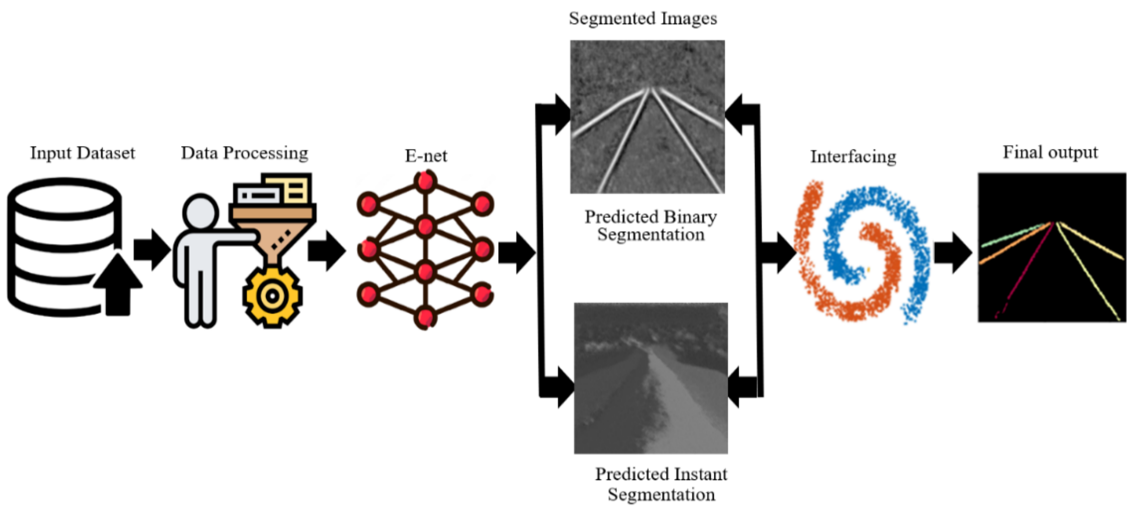

Figure 2. Complete workflow diagram of the proposed approach

3.1. Input dataset and preprocessing

TuSimple dataset has been used to develop the model because this dataset has adequate images with accurate annotation. The training images and the testing images are of good quality, and it ensures the standards of the model. Images of this dataset are annotated with full lane boundary instead of just annotating the lanemark. Each image of this dataset is of good quality in 720×1280 px dimension. The dataset comes with 3 JSON files that indicate the clips' path having 3,626 image frames, lane position, and the lane's height as a list. After the lane features get extracted, a hyper line is drawn to fit every lane's data points. Here, the corresponding lane pixel has been converted into binary value for the pixels that do not belong to the lanes to create the binary and instant label images. Eventually, the image frames were resized to 224×224 px dimension to maintain a constant aspect ratio and to reduce the computational cost. After completing all the above data preprocessing steps, the output data is the original image with an instant label and binary label.

3.2. E-Net architecture

The original ENet architecture is an encode-decode network in which there are three stages for the encoding section and two stages for the decode section. However, the decode section only upsamples the output information obtained from encoding stages. Sharing all the data from the encode stages towards the decode stages will lead to a lower result considering irrelevant information of inputs data like excluding the lane information. Therefore, the original ENet architecture has been customized by dividing the encode section into binary and instant segments. Consequently, each unit can carry one particular information regarding lane and perform the individual task. Binary segmentation provides information about the pixels inherent to the lanes, whereas instant segmentation ensures the proper pixel position of the lane on the images. The layout of the customized E-Net architecture is shown in Figure 3. Several small Bottlenecks have been formed before going to the main encode-decode sections, for example, normal bottleneck (NB), encode bottleneck (EB), and decode bottleneck (DB). Hence, it will reduce the number of features in every layer to reduce computational complexity, learn the relevant features more deeply and find the best possible training loss.

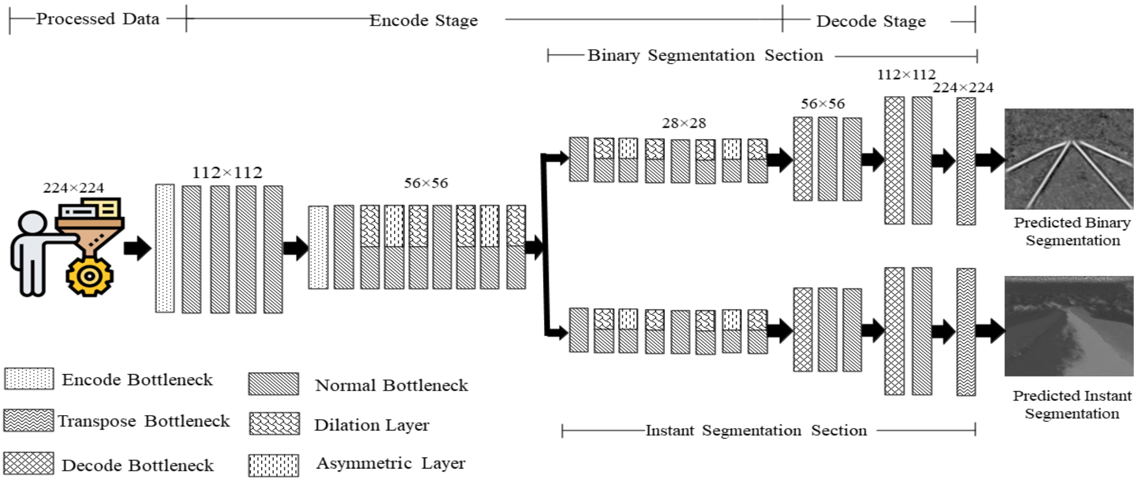

Figure 3. The architecture of the customized E-Net model

3.2.1. Normal bottleneck (NB)

In NB, 1×1 convolution has been applied to decrease the channel numbers. 1×1 convolutional layer reduces the computational complexity and has the efficiency of the embedding process in the pooling section [33]. Dilation convolution with kernel 3×3 has been used to maintain the input data's constant dimension and keep the resolution high. Asymmetric convolution has been fixed with kernel 5×1 to reduce the probability of overfitting and computational complexity as it convolutes the separate input channel by different filter channels. A regular convolution with kernel 3×3 has been utilized for continuing the convolution with the same dimension. Besides, a 1×1 dimensional convolution has been implemented to maintain the channel number back to the initial numbers of channels. Furthermore, the dropout layer also has been utilized, considering regularization to reduce the overfitting.

3.2.2. Encode bottleneck (EB)

Max pooling has been executed with the kernel of 2×2 and stride 2×2 in the EB. Besides, 2D convolution with the same kernel and stride has been applied to reduce the channel numbers. Also, 1×1 convolution and dropout layer have been utilized as like as NB.

3.2.3. Decode bottleneck (DB)

Max unpooling has been executed with the same resolution as EB. Transposed convolutional layer with kernel 3×3 and stride 2×2 has been used to uplift the encoding features. Also, 1×1 convolution and dropout layer have been utilized as like as NB.

3.2.4. Encoder stage

The encoder initially received the original, binary label, and instant label images from the data processing section. It separated the image information into two segments, binary and instant segmentation, and extracted the lane features and instances individually. Two bottlenecks have been used in the encoder stage to split the input dataset regarding binary segmentation and instant segmentation. There are one EB and four NB in the first bottleneck. In constrain, one EB and seven NB with three dilation and two asymmetric convolution layers have been applied to the second bottleneck. Besides, eight NB has four dilations, and two asymmetric convolutions have been examined in the bottleneck for binary segmentation and instant segmentation.

3.2.5. Decode stage

The information that has been obtained from the encode stages needs to decode to have the final result from the network. One up-sampling and two NB have been applied in the bottleneck for the binary segmentation part and for the instant segmentation part for uplifting the lane information respectfully. Again, one up-sampling and NB have been used for the same purpose of boosting the lane information. Eventually, the network has provided the predicted output as binary segmentation and instant segmentation images after implementing a transposed convolutional layer in the last bottleneck.

3.2.6. Loss measurement

Like other machine learning models, the loss has been computed using backpropagation, and the model has been updated accordingly to optimize the model. There are two types of casualties that have been executed for the two segmented images such as cross-entropy and discriminative loss. The binary segmented images preserve the data as 0 and 1. Here the computation of cross-entropy loss has been performed using (1).

$$-\left(y(\log(p) + (1-y)\log(1-p))\right) \tag{1}$$

Since instant-based segmentation finds the exact location of the lane, the discriminative loss has been computed in this stage. The model has been designed so that the same label pixel takes a nearby position, and different label pixel maintains distances. So, it puts pixels of the same lane in the same cluster, and pixels from different lanes go onto other indifferent perspective lanes. The whole process has been modeled in three different stages: separation, neighborhood, and regulation. The separation process extends the distance between two lene clusters based on a threshold value. On the other hand, the neighborhood section reduces the distance of pixels in a particular lane cluster based on a threshold value δ_{neighb}. Additionally, the regularization section maintains the origin of any cluster. Decisively, the discriminative loss functions are computed by using the following formula.

$$Disc_{loss} = Loss_{sep} + Loss_{neighb} + Loss_{Regu}$$

$$= \frac{1}{N_c}\sum_{N=1}^{N}\frac{1}{N_e}\sum_{j=1}^{N_e}\left[\| M - x_i\| - \delta_{neighb}\right]_+^2 + \frac{1}{N_c(N_c-1)}\sum_{N_{ca}=1}^{N_c}\sum_{N_{ca}=1}\left[\delta_{sep} - \|M_{ca} - M_{cb}\|\right]_+^2 + \frac{1}{N_c}\sum_{N_c=1}^{N_c}\|M\|$$

Here, N_c represents the number of lane clusters, N_e represents the number of elements in the lane cluster, M represents the mean value of the instance in the cluster, and x_i represents instances. The total loss of the network has been computed by accumulating the cross-entropy loss and the discriminative loss. Backpropagation has updated the weight of this neural network model by operating on this total loss of the network.

3.2.7. Interfacing

The output of our proposed model is the immediate segmentation of pixels in each lane on the anticipated pictures. The final goal is to superimpose the lane pixels on top of the original input images. Hence, the densely based spatial clustering of application with noise (DBSCAN), has been employed to interface the predicted lane image with the original input image. DBSCAN performs better and more efficiently than most common clustering techniques like K-means and so on, especially for noisy or arbitrary clusters [34]. If the lanes are positioned close and random in nature, DBSCAN performs better interface the lane pixels. In this DBSCAN model, the nearest distance has been considered is 0.05 for the same lane pixels. A point is considered to be in the same lane if it is close to the mentioned threshold distance. Otherwise, the point is considered to belong to another cluster. This whole process iterates until all the points on the lanes are converged.

4. RESULTS AND DISCUSSION

The preprocessed data has been fed into the model to train, and then the model performance has been tested in terms of accuracy and other relevant metrics. Accuracy itself can't guarantee the reliability of this model, so relevant parameters like positive rate, false-negative rate, F1 score are also considered in this type of classification task [35]–[38]. The equations regarding performance parameters have been mentioned in (2) to (5).

$$\text{Accuracy value} = \frac{\text{Quantity of actual prediction}}{\text{Total sample data}} \tag{2}$$

$$\text{Precision value} = \frac{\text{actual positive prediction}}{\text{true positive+false positive}} \tag{3}$$

$$\text{F1 score value} = 2 \times \frac{\text{Precision value} \times \text{Recall value}}{\text{Precision value+Recall value}} \tag{4}$$

$$\text{Recall value} = \frac{\text{actual number of positive prediction}}{\text{Number of true positive+number of false}-\text{negative}} \tag{5}$$

Our model has been trained for a total of hundred epochs with images of dimension 224×224 px having a batch size of 16. In this work, PReLu activation has been used to develop this model. Adam optimizer has been used with a learning rate of 0.0001 and strides, valid padding setting to optimize the proposed model. Eventually, the whole model was developed on an Ubuntu-based Linux environment with GTX 1080 Ti GPU unit. The key parameters of the performance of the proposed model have been listed below in Figure 4. This model has recorded the highest accuracy value of 96.61 with an F1 score of 96.34, a precision of 98.91, and a recall value of 93.89 percent. In addition, the model has shown minimal false-positive and false-negative rates, 3.125% and 1.259%, respectively.

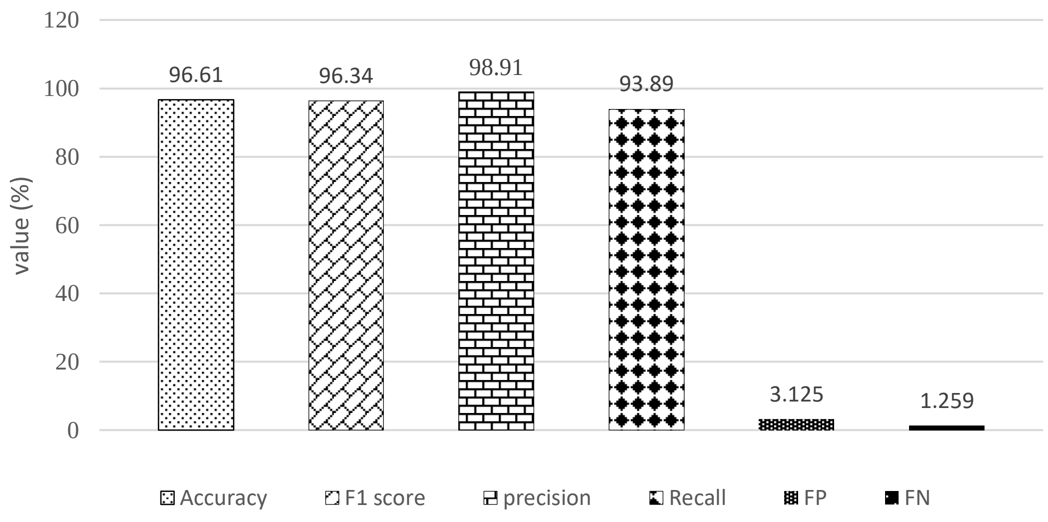

Figure 4. Performance metrics visualization of the proposed model

The proposed approach has also been evaluated in distinct epochs like 20, 40, 60, 80, and 100. The performance has been listed in Table 1. The performance of this model has increased gradually with respect to the epochs. Eventually, this model has gained maximum accuracy at around the 100th epoch, as depicted in Table 1.

Table 1. Performance result comparison with different epochs

Epoch	Accuracy	F1 score	Precision	Recall
20	90.1	93.4	94.85	91.2
40	92.32	94.21	95.01	91.47
60	94.35	95.66	95.5	92.39
80	95.45	95.79	97.36	93.2
100	**96.61**	**96.34**	**98.91**	**93.45**

The loss measurement and optimization are very central steps for developing a functional deep learning-based model as the lowest possible loss ensures the best-optimized architecture. The overall loss of the model has been visualized in Figure 5, representing the gradual decrease of the loss during the training process. The accumulated minimal loss of the model has been noted as 5.39%, which indicates the efficiency of the proposed model having a minimal loss. Again, as the loss gets reduced in every epoch, the basic features of lane marking get extracted from the input images. This process also ensures the minimization of the false-positive rate.

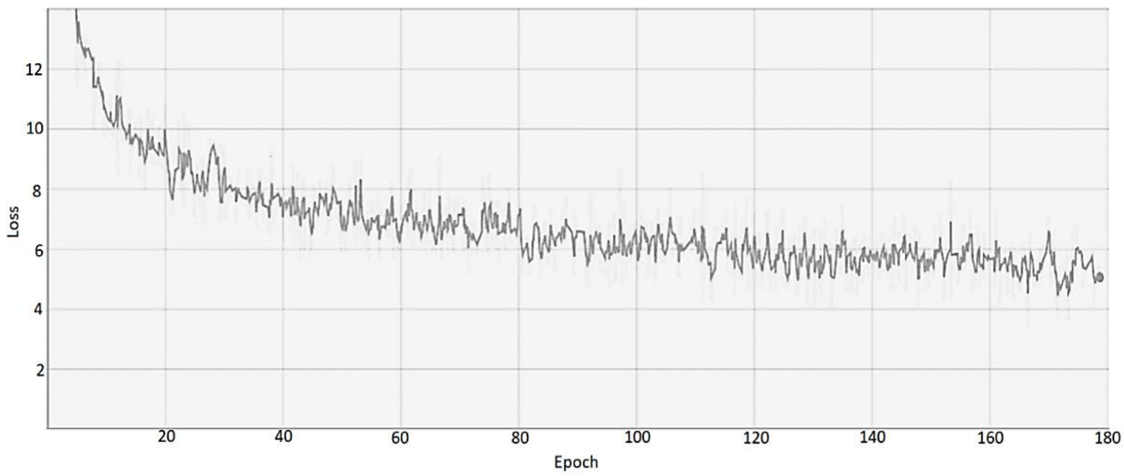

Figure 5. Loss visualization with respect to the number of epochs

The overall performance of the proposed model has been compared with the recent state-of-the-art model shown in Table 2. The table shows that our approach is superior to other recent deep learning-based lane mark recognition models. The aforementioned model has shown the best performance in terms of accuracy, F1 score, precision, and recall, compared to most of the existing relevant research articles. Besides, this work has shown the lowest false positive and false negative values compared to the other existing work in this domain. As the proposed method's evolutionary result is superior to the current deep learning techniques, this method is more efficient for detecting lane marking than others.

Table 2. Performance comparison of the proposed method and some existing methods

Methods	Accuracy (%)	Recall	Precision	False Positive	False Negative	F1 score
Pizzati *et al.* [39]	95.24	-	-	9.42	0.033	-
Hoe *et al.* [11]	96.29	-	-	3.21	4.28	-
Mamidala *et al.* [40]	96.10	-	-	-	-	94.45
Yoo *et al.* [41]	96.02	-	-	7.22	2.18	-
Tabelini *et al.* [42]	93.36	-	-	6.17	-	-
He *et al.* [27]	-	93.80	95.49	-	-	-
Zhe *et al.* [43]	-	-	94.94	2.79	4.99	-
Tian *et al.* [19]	-	66.4	83.5	-	-	-
Proposed Method	**96.61**	**93.89**	**98.91**	**3.125**	**1.259**	**96.34**

As the whole architecture of the model has been trained through a small bottleneck of convolutional layers with fewer parameters, it has reduced the complexity to compute the training process. Besides, the proposed approach has utilized the asymmetric convolutional layers, which reduced the probabilities of overfitting and computational cost. Besides, DBSCAN has been executed to interface the predicted segmented images instead of using the different convolution-based networks to reduce computational cost as well as to provide compatibility with both straight and curved lines. This model was trained and tested with the TuSimple dataset, which contains roads frames in very different and complex environmental conditions. A few of the input and output images from the lane mark have been shown in Figure 6 to visualize the proposed method's outcome. Screen copies of the original image predicted lanemark, corresponding color image, and projected lanemark have been shown in the figure. This figure ensures that our proposed model can determine lane marking more accurately and precisely than other existing models. We strongly believe that this research will significantly contribute to lane mark detection research.

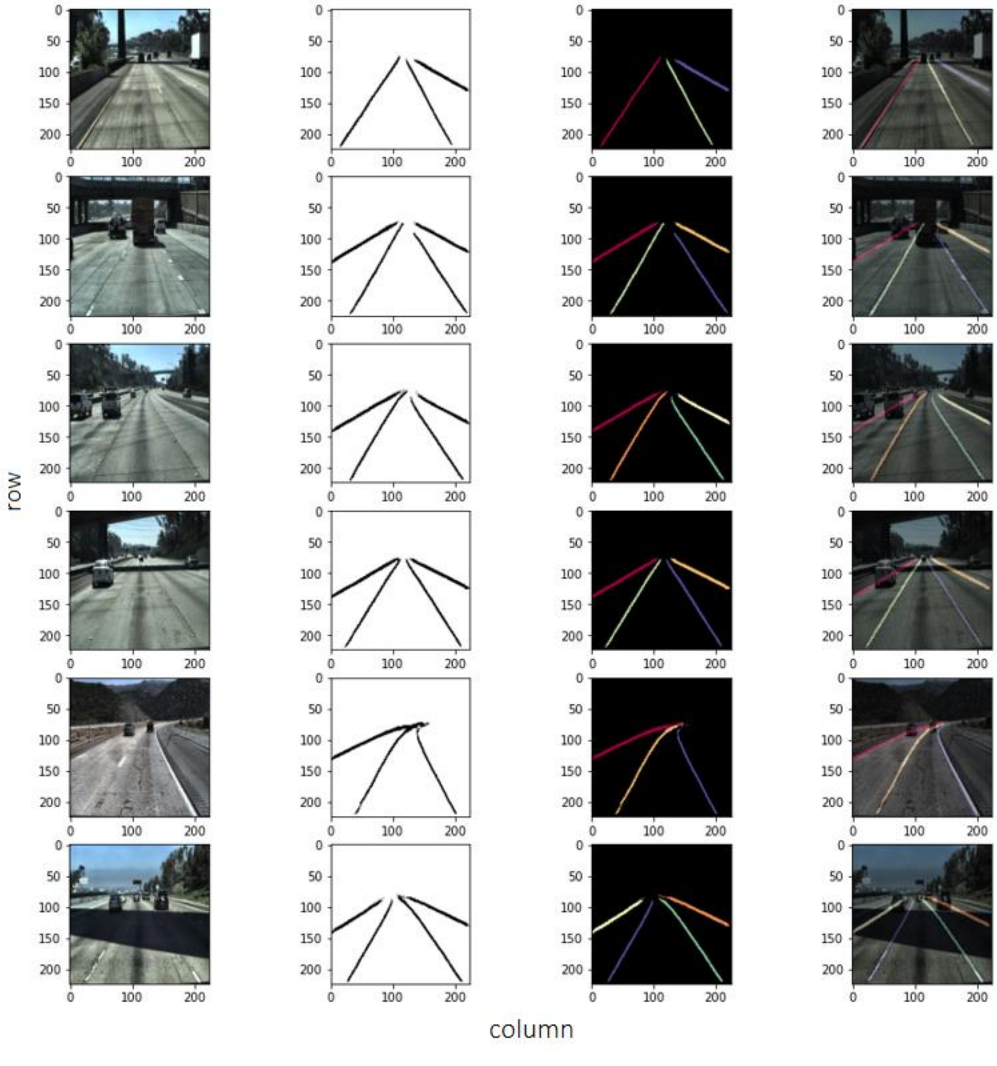

Figure 6. The output of the proposed model: original image (1st column), lane prediction in black (2nd column), lane prediction in color (3rd column), lane projection in the original image (4th column)

5. CONCLUSION

ADAS are mainly supporting autonomous vehicles technology, and lane mark detection is one of the core components of ADAS. In this research work, a simple encode-decode basis customized E-Net architecture has been used to find lane marks in very diverse and practical environmental conditions. TuSimple, a robust dataset that includes frames of very diverse environmental conditions, including straight lane, low light, shadow, curve lane, and so on, has been used to develop and justify the performance of the model. The proposed model has shown better accuracy, F1 score, precision, and recall than the other existing model. In addition, the proposed model has less computation complexity compared to the existing models.

Besides, our model showed a minimal loss during the training process. Hence, the proposed architecture is a better model for lane mark detection, outperforming the state-of-the-art technologies in every performance parameter. So, this model will create a significant positive impact in the AI-based ADAS research area. However, this result may be improved by training the model with a more robust dataset that contains frames of more diverse environmental conditions.

ACKNOWLEDGEMENTS

The research work has been supported by the Malaysian Ministry of Higher Education (MOHE) for Fundamental Research.

REFERENCES

[1] WHO, "WHO global status report on road safety 2013," World Health Organization, 2014.
[2] W. Q. Yan, *Introduction to intelligent surveillance: Surveillance data capture, transmission, and analytics: Second edition*. 2017.
[3] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An open urban driving simulator," *arXiv:1711.03938*, Nov. 2017.
[4] L. A. Tran and M. H. Le, "Robust u-net-based road lane markings detection for autonomous driving," in *Proceedings of 2019 International Conference on System Science and Engineering, ICSSE 2019*, 2019, pp. 62–66, doi: 10.1109/ICSSE.2019.8823532.
[5] R. Van der Heijden and K. Van Wees, "Introducing advanced driver assistance systems: Some legal issues," *European Journal of Transport and Infrastructure Research*, vol. 1, no. 3, pp. 309–326, 2001, doi: 10.18757/ejtir.2001.1.3.3486.
[6] P. Szikora and N. Madarasz, "Self-driving cars - the human side," in *2017 IEEE 14th International Scientific Conference on Informatics, INFORMATICS 2017 - Proceedings*, Mar. 2018, vol. 2018, pp. 383–387, doi: 10.1109/INFORMATICS.2017.8327279.
[7] S.-Y. Lo, H.-M. Hang, S.-W. Chan, and J.-J. Lin, "Multi-class lane semantic segmentation using efficient convolutional networks," in *2019 IEEE 21st International Workshop on Multimedia Signal Processing (MMSP)*, Sep. 2019, pp. 1–6, doi: 10.1109/MMSP.2019.8901686.
[8] W. Wang, D. Zhao, W. Han, and J. Xi, "A Learning-based approach for lane departure warning systems with a personalized driver model," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 10, pp. 9145–9157, Oct. 2018, doi: 10.1109/TVT.2018.2854406.
[9] Y. Sun, J. Li, and Z. P. Sun, "Multi-lane detection using CNNs and a novel region-grow algorithm," *Journal of Physics: Conference Series*, vol. 1187, no. 3, Apr. 2019, doi: 10.1088/1742-6596/1187/3/032018.
[10] T. P. Nguyen, V. H. Tran, and C. C. Huang, "Lane detection and tracking based on fully convolutional networks and probabilistic graphical models," in *Proceedings - 2018 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2018*, 2019, pp. 1282–1287, doi: 10.1109/SMC.2018.00224.
[11] Y. Hou, "Agnostic lane detection," *arXiv1905.03704v1*, pp. 1–6, May 2019.
[12] A. Borkar, M. Hayes, and M. T. Smith, "A novel lane detection system with efficient ground truth generation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 1, pp. 365–374, Mar. 2012, doi: 10.1109/TITS.2011.2173196.
[13] H. Deusch, J. Wiest, S. Reuter, M. Szczot, M. Konrad, and K. Dietmayer, "A random finite set approach to multiple lane detection," in *2012 15th International IEEE Conference on Intelligent Transportation Systems*, Sep. 2012, pp. 270–275, doi: 10.1109/ITSC.2012.6338772.
[14] Q. Zou, H. Jiang, Q. Dai, Y. Yue, L. Chen, and Q. Wang, "Robust lane detection from continuous driving scenes using deep neural networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 1, pp. 41–54, Jan. 2020, doi: 10.1109/TVT.2019.2949603.
[15] W. Zhang, H. Liu, X. Wu, L. Xiao, Y. Qian, and Z. Fang, "Lane marking detection and classification with combined deep neural network for driver assistance," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 233, no. 5, pp. 1259–1268, 2019, doi: 10.1177/0954407018768659.
[16] X. Yu and Y. Sun, "Research on parking detecting analysis based on projection transformation and Hough transform," *Journal of Physics: Conference Series*, vol. 1187, no. 4, Apr. 2019, doi: 10.1088/1742-6596/1187/4/042068.
[17] M. Juneja and P. S. Sandhu, "Performance evaluation of edge detection techniques for images in spatial domain," *International Journal of Computer Theory and Engineering*, pp. 614–621, 2009, doi: 10.7763/IJCTE.2009.V1.100.
[18] D. A. Zuehlke, T. A. Henderson, and S. A. H. McMullen, "Machine learning using template matching applied to object tracking in video data," in *Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications*, May 2019, doi: 10.1117/12.2518982.
[19] Y. Tian *et al.*, "Lane marking detection via deep convolutional neural network," *Neurocomputing*, vol. 280, pp. 46–55, Mar. 2018, doi: 10.1016/j.neucom.2017.09.098.
[20] X. Wen, L. Shao, W. Fang, and Y. Xue, "Efficient feature selection and classification for vehicle detection," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 3, pp. 508–517, 2015, doi: 10.1109/TCSVT.2014.2358031.
[21] R. Gopalan, T. Hong, M. Shneier, and R. Chellappa, "A learning approach towards detection and tracking of lane markings," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 3, pp. 1088–1098, Sep. 2012, doi: 10.1109/TITS.2012.2184756.
[22] M. Everingham, S. M. A. Eslami, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes challenge: a retrospective," *International Journal of Computer Vision*, vol. 111, no. 1, pp. 98–136, 2015, doi: 10.1007/s11263-014-0733-5.
[23] S. Kawaguchi, Eric, I. Shen, Jenny, G. Li, and A. Suchard, Marc, "A fast and scalable implementation method for competing risks data with the R package fastcmprsk," *The R Journal*, vol. 12, no. 2, pp. 163–171, 2020, doi: 10.32614/RJ-2021-010.
[24] B. Huval *et al.*, "An empirical evaluation of deep learning on highway driving," *arXiv:1504.01716*, vol. abs/ 1504, Apr. 2015.
[25] D. Levi, N. Garnett, and E. Fetaya, "StixelNet: a deep convolutional network for obstacle detection and road segmentation," 2015, doi: 10.5244/C.29.109.
[26] A. Al Mamun, P. P. Em, and J. Hossen, "Lane marking detection using simple encode decode deep learning technique: SegNet,"

International Journal of Electrical and Computer Engineering (IJECE), vol. 11, no. 4, pp. 3032–3039, 2021, doi: 10.11591/ijece.v11i4.pp3032-3039.
[27] B. He, R. Ai, Y. Yan, and X. Lang, "Lane marking detection based on convolution neural network from point clouds," in *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 2016, pp. 2475–2480, doi: 10.1109/ITSC.2016.7795954.
[28] Y. Huang, S. Chen, Y. Chen, Z. Jian, and N. Zheng, "Spatial-temproal based lane detection using deep learning," in *IFIP Advances in Information and Communication Technology*, 2018, vol. 519, pp. 143–154, doi: 10.1007/978-3-319-92007-8_13.
[29] J. Li, X. Mei, D. Prokhorov, and D. Tao, "Deep neural network for structural prediction and lane detection in traffic scene," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 3, pp. 690–703, 2017, doi: 10.1109/TNNLS.2016.2522428.
[30] Y. Y. Ye, X. L. Hao, and H. J. Chen, "Lane detection method based on lane structural analysis and CNNs," in *IET Intelligent Transport Systems*, 2018, vol. 12, no. 6, pp. 513–520, doi: 10.1049/iet-its.2017.0143.
[31] A. Paszke, A. Chaurasia, S. Kim, and E. Culurciello, "ENet: a deep neural network architecture for real-time semantic segmentation," *arXiv:1606.02147*, pp. 1–10, Jun. 2016.
[32] D. Neven, B. De Brabandere, S. Georgoulis, M. Proesmans, and L. Van Gool, "Towards end-to-end lane detection: an instance segmentation approach," in *IEEE Intelligent Vehicles Symposium, Proceedings*, 2018, vol. 2018, pp. 286–291, doi: 10.1109/IVS.2018.8500547.
[33] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2015, pp. 1–9, doi: 10.1109/CVPR.2015.7298594.
[34] P. H. Ahmad and S. Dang, "Performance evaluation of clustering algorithm using different datasets," *International Journal of Advance Research in Computer Science and Management Studies*, vol. 3, no. 1, pp. 167–173, 2015.
[35] M. Rashid *et al.*, "The classification of motor imagery response: an accuracy enhancement through the ensemble of random subspace k-NN," *PeerJ Computer Science*, vol. 7, Mar. 2021, doi: 10.7717/peerj-cs.374.
[36] A. Al Mamun, M. S. Hossain, P. P. Em, A. Tahabilder, R. Sultana, and M. A. Islam, "Small intestine bleeding detection using color threshold and morphological operation in WCE images," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 4, pp. 3040–3048, Aug. 2021, doi: 10.11591/ijece.v11i4.pp3040-3048.
[37] A. Al Mamun, P. P. Em, T. Ghosh, M. M. Hossain, M. G. Hasan, and M. G. Sadeque, "Bleeding recognition technique in wireless capsule endoscopy images using fuzzy logic and principal component analysis," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 3, pp. 2688–2695, Jun. 2021, doi: 10.11591/ijece.v11i3.pp2688-2695.
[38] A. Al Mamun, M. S. Hossain, M. M. Hossain, and M. G. Hasan, "Discretion way for bleeding detection in wireless capsule endoscopy images," in *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*, May 2019, pp. 1–6, doi: 10.1109/ICASERT.2019.8934589.
[39] F. Pizzati, M. Allodi, A. Barrera, and F. García, "Lane detection and classification using cascaded CNNs," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2020, vol. 12014, pp. 95–103, doi: 10.1007/978-3-030-45096-0_12.
[40] R. S. Mamidala, U. Uthkota, M. B. Shankar, A. J. Antony, and A. V Narasimhadhan, "Dynamic approach for lane detection using google street view and CNN," in *TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON)*, Oct. 2019, vol. 2019, pp. 2454–2459, doi: 10.1109/TENCON.2019.8929655.
[41] S. Yoo *et al.*, "End-to-end lane marker detection via row-wise classification," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2020, vol. 2020, pp. 4335–4343, doi: 10.1109/CVPRW50498.2020.00511.
[42] L. Tabelini, R. Berriel, T. M. Paixao, C. Badue, A. F. De Souza, and T. Oliveira-Santos, "PolyLaneNet: lane estimation via deep polynomial regression," in *2020 25th International Conference on Pattern Recognition (ICPR)*, Jan. 2021, pp. 6150–6156, doi: 10.1109/ICPR48806.2021.9412265.
[43] Z. Chen and Z. Chen, "RBNet: A deep neural network for unified road and road boundary detection," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2017, vol. 10634, pp. 677–687, doi: 10.1007/978-3-319-70087-8_70.

BIOGRAPHIES OF AUTHORS

Abdullah Al Mamun has received a B.Sc. degree in Electrical and Electronic Engineering from Pabna University of Science and Technology in 2018. Now he is pursuing M.Eng.Sc. at Multimedia University (MMU) in the Faculty of Engineering and Technology since 2019. He is also working as a lecturer (on study leave) since November 2019 in EEE at Feni University, Feni, Bangladesh. His research interest includes computer vision, image processing, signal processing, deep learning, and machine learning. He can be reached at the email address mamun130203@gmail.com.

Poh Ping Em currently is a senior lecturer at the Faculty of Engineering and Technology, Multimedia University. Poh Ping does research in vehicle safety, data fusion, fuzzy control, and image processing. His research mainly focuses on developing advanced driver assistance systems. He can be contacted at email: ppem@mmu.edu.my.

Md. Jakir Hossen is graduated in Mechanical Engineering from the Dhaka University of Engineering and Technology (Former Name is BIT Dhaka,1997), Higher Diploma in Computer Science and Engineering (1999), master's in communication and Network Engineering from Universiti Putra Malaysia (2003), and Ph.D. in Smart Technology and Robotic Engineering at Universiti Putra Malaysia (2012). He is currently a Senior Lecturer at the Faculty of Engineering and Technology, Multimedia University, Melaka, Malaysia. His research interests are in the area of artificial intelligence, data analytics, robotics and control. He can be contacted at email: jakir.hossen@mmu.edu.my.

Anik Tahabilder is currently pursuing his Ph.D. in Computer Science at Wayne State University, Detroit, Michigan, USA. He had his master's degree in Engineering Technology from Western Carolina University, Cullowhee, NC-28723, USA. He also got a bachelor's degree in Electrical and Electronic Engineering from Pabna University of Science and Technology, Pabna-6600, Bangladesh. His research interest includes deep learning, big data workflow, and blockchain. He can be contacted at email: tahabilderanik@gmail.com.

Bushrat Jahan is working as a Senior Lecturer and Department head of CSE since 2019 at Feni University, Feni, Bangladesh. Currently, she is working on Data Science, Information Security, Natural Language Processing, and Machine Learning Approaches. She can be contacted at email: hossenbipasa980@gmail.com.