Road Lane Detection and Classification in Urban and Suburban Areas based on CNNs

Nima Khairdoost, Steven S. Beauchemin and Michael A. Bauer
Department of Computer Science, The University of Western Ontario, London, ON, N6A-5B7, Canada

Keywords: Lane Detection, Lane Type Classification, CNN.

Abstract: Road lane detection systems play a crucial role in the context of Advanced Driver Assistance Systems (ADASs) and autonomous driving. Such systems can lessen road accidents and increase driving safety by alerting the driver in risky traffic situations. Additionally, the detection of ego lanes with their left and right boundaries along with the recognition of their types is of great importance as they provide contextual information. Lane detection is a challenging problem since road conditions and illumination vary while driving. In this contribution, we investigate the use of a CNN-based regression method for detecting ego lane boundaries. After the lane detection stage, following a projective transformation, the classification stage is performed with a RseNet101 network to verify the detected lanes or a possible road boundary. We applied our framework to real images collected during drives in an urban area with the RoadLAB instrumented vehicle. Our experimental results show that our approach achieved promising results in the detection stage with an accuracy of 94.52% in the lane classification stage.

1 INTRODUCTION

Nowadays, almost every new vehicle features some type of Advanced Driving Assistance System (ADAS), ranging from adaptive cruise control, blind-spot detection, collision avoidance, traffic sign detection, overtaking assistance, to parking assistance. ADASs generally increase safety and reduce driver workload. Lane detection constitutes one of the fundamental functions found in autonomous driving systems and ADASs. Lane boundaries provide the information required for estimating the lateral position of a vehicle on the road, enabling systems such as lane departure warning, overtaking assistance, intelligent cruise control, and trajectory planning.

Lane detection approaches are categorized into two groups: classical and deep learning methods. The traditional lane detection methods usually employ a number of computer vision and image processing techniques to extract specialized features and to identify the location of lane segments. Subsequently, post-processing techniques remove false detections and join sub-segments to obtain final road lane positions. In general, these traditional approaches suffer from performance issues when they encounter challenging illumination conditions and complex road scenes.

Recently, deep learning-based methods have been employed to provide reliable solutions to the lane detection problem. Methods based on CNNs fall into two categories, namely segmentation-based methods and Generative Adversarial Network based methods (GAN) (Yoo et al., 2020). (Chougule et al., 2018) proposed a regression-coordinate network based on CNN for lane detection in highway driving scenes in an end-to-end fashion. In this study, we followed their lane detection strategy in environments where there exists a greater variety of lane types as opposed to highways. We classify various types of lanes as they indicate traffic rules relevant for driving. Following the detection stage, we use a two-step algorithm to classify the lane boundaries into eight classes, considering road boundaries (no markings) as one particular type of lane.

The rest of this contribution is organized as follows: In Section 2, we review the related literature. Section 3 provides a summary of the datasets and the lane model. Results and evaluations are given in Section 4. Finally, we summarize our results in Section 5.

Khairdoost, N., Beauchemin, S. and Bauer, M.
Road Lane Detection and Classification in Urban and Suburban Areas based on CNNs.
DOI: 10.5220/0010241004500457
In *Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2021) - Volume 5: VISAPP*, pages 450-457
ISBN: 978-989-758-488-6

2 LITERATURE SURVEY

In this Section, we survey both traditional and Deep Learning methods for lane marking recognition and classification.

2.1 Traditional Approaches

Most traditional methods extract a combination of visual highly-specialized features using various elements such as color (Chiu and Lin, 2005), (Cheng et al., 2006), edges (Lee and Moon, 2018), ridge features (López et al., 2010), and template matching (Choi and Oh, 2010). These primitive features can also be combined by way of Hough transforms (Liu et al., 2010), Kalman filters (Mammeri et al., 2014), (Kim, 2008), and particle filters (Linarth and Angelopoulou, 2011). Most of these methods are sensitive to illumination changes and road conditions and thus prone to fail.

2.2 Deep Learning-based Approaches

There are mainly two groups of segmentation methods for lane marker detection: 1) Semantic Segmentation and 2) Instance Segmentation. In the first group, each pixel is classified by a binary label indicating whether it belongs to a lane or not. For instance, in (He et al., 2016), the authors presented a CNN-based framework that utilizes front-view and top-view image regions to detect lanes. Following this, they used a global optimization step to reach a combination of accurate lane lines. (Lee et al., 2017) proposed a Vanishing Point Guided Net (VPGNet) model that simultaneously performs lane detection and road marking recognition under different weather conditions. Their data was captured in a downtown area of Seoul, South Korea.

Conversely, Instance Segmentation approaches differentiate individual instances of each class in an image and identify separate parts of a line as one unit. (Pan et al., 2018) proposed the Spatial CNN (SCNN) to achieve effective information propagation in the spatial domain. This CNN-analogous scheme effectively retains the continuity of long and thin shapes such as road lanes, while its diffusion effects enable it to segment large objects. LaneNet (Neven et al., 2018) is a branched, instance segmentation architecture that produces a binary lane segmentation mask and pixel embeddings. These are used to cluster lane points. Subsequently, another neural network called H-net with a custom loss function is employed to parameterize lane instances before the lane fitting.

GANs have been used for lane detection. (Liu et al., 2020) presented a style-transfer-based data enhancement approach, which used GANs (Goodfellow et al., 2014) to create images in low-light conditions that raise the environmental adaptability of the model. Their method does not require additional annotation nor extraneous inference overhead. (Ghafoorian et al., 2018) proposed an Embedding Loss GAN (EL-GAN) framework for lane boundary segmentation. The discriminator receives the source data, a prediction map, and a ground truth label as inputs and is trained to minimize the difference between the training labels and embeddings of the predictions. In (Kim et al., 2020), a data augmentation method with GAN was proposed for oversampling minority anomalies in lane detection. The GAN network is employed to address the imbalance problem by synthesizing the anomalous data. It learns the distribution of the falsely detected lane by itself, without domain knowledge.

2.3 Approaches for Lane Type Classification

Different types of lane markings exist. Generally, a lane marking is categorized by its color, with dashed or solid, and single or double segments. In (Hoang et al., 2016), a method is presented for road lane detection that discriminates dashed and solid lane markings. Their method outperformed conventional lane detection methods. Several other approaches such as (Sani et al., 2018), (de Paula and Jung, 2013), and (Ali and Hussein, 2019), recognize five lane marking types including Dashed, Dashed-Solid, Double Solid, Solid-Dashed, and Single Solid. In (Sani et al., 2018), a method that utilizes a two-layer classifier was proposed to classify these lane markings using a customized Region of Interest (ROI) and two derived features, namely; the contour number, and the contour angle. In (de Paula and Jung, 2013), the authors presented a method to detect lane markers based on a linear parabolic model and geometric constraints. To classify lane markers into the aforementioned five classes, a three-level cascaded classifier consisting of four binary classifiers was developed. In (Ali and Hussein, 2019), the ROI is divided into two subregions. To identify the lane types, a method based on the Seed Fill algorithm is applied to the location of the lanes. (Lo et al., 2019b) proposed two techniques, Feature Size Selection and Degressive Dilation Block to extend an existing semantic segmentation network called EDANet (Lo et al., 2019a) to discriminate the road from four types of lanes, including double solid yellow, single dashed yellow, single solid red, and single solid white.

3 PROPOSED METHOD

In this Section, we present our approaches to the problem of lane marking recognition and classification, with their respective datasets extracted from the RoadLAB experiments.

3.1 Lane Detection Stage

3.1.1 Regression-based Lane Detection Model

To identify the ego lane boundaries in the road image, a regression-based network is utilized that outputs two vectors representing the coordinate points of the left and right boundaries from the ego lane. Each coordinate vector consists of 14 coordinates (x, y) on the image plane indicating sampled positions for the ego lane boundary. To construct this model, a pre-trained AlexNet architecture is utilized. First, the last two fully connected layers are removed from the network and then four-level cascaded layers are added to the first six layers of AlexNet to complete the lane detection model. These four-level cascaded layers contain two branches of two back-to-back fully connected layers, a concatenation layer and a regression layer, as shown in Figure 1. This branched architecture minimizes misclassifications of the detected lane points (Chougule et al., 2018). Moreover, this architecture is capable of detecting the road boundary as an assumptive ego lane left/right boundary when there is no actual lane marking.

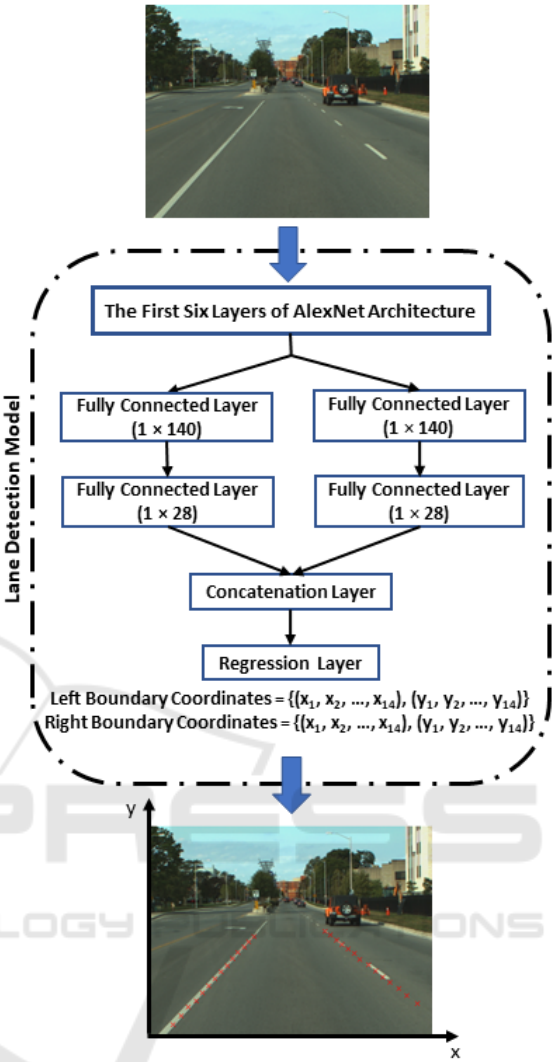

Figure 1: The lane detection model provides two lane vectors, each consisting of 14 coordinates in the image plane that represent the predicted left and right boundaries of the ego lane.

3.1.2 Our Dataset for Lane Detection

In this Section, we introduce our lane detection dataset extracted from the driving sequences, captured with the RoadLAB instrumented vehicle (Beauchemin et al., 2012), (see Figure 2). Our experimental vehicle was used to collect driving sequences from 16 drivers on a pre-determined 28.5km route within the city of London Ontario, Canada. (see Figure 3). Data frames were collected at a rate of 30Hz with a resolution of 320×240. We used 12 driving sequences, as described in Table 1, to derive our dataset containing 5782 images along with their corresponding lane annotations. Figure 4 illustrates examples from our derived dataset.

Figure 2: Forward stereoscopic vision system mounted on rooftop of the RoadLAB experimental vehicle.

An essential element of any deep learning-based system is the availability of large numbers of sample images. Data augmentation is a commonly used strategy to significantly expand an existing dataset by generating unique samples through transformations of images in the dataset. The exploitation of data augmentation strategy reduces overfitting from the network. We employed data augmentation techniques to enrich the dataset, resulting in an improved performance at the lane detection stage.

Table 1: Summary of driving conditions of our data (Each row belongs to one driver).

Seq. #	Capture Date	Time	Temperature	Weather
2	2012-08-24	15:30	31 °C	Sunny
4	2012-08-31	11:00	24 °C	Sunny
5	2012-09-05	12:05	27 °C	Partially Cloudy
8	2012-09-12	14:45	27 °C	Sunny
9	2012-09-17	13:00	24 °C	Partially Cloudy
10	2012-09-19	09:30	8 °C	Sunny
11	2012-09-19	14:45	12 °C	Sunny
12	2012-09-21	11:45	18 °C	Partially Sunny
13	2012-09-21	14:45	19 °C	Partially Sunny
14	2012-09-24	11:00	7 °C	Sunny
15	2012-09-24	14:00	13 °C	Partially Sunny
16	2012-09-28	10:00	14 °C	Partially Sunny

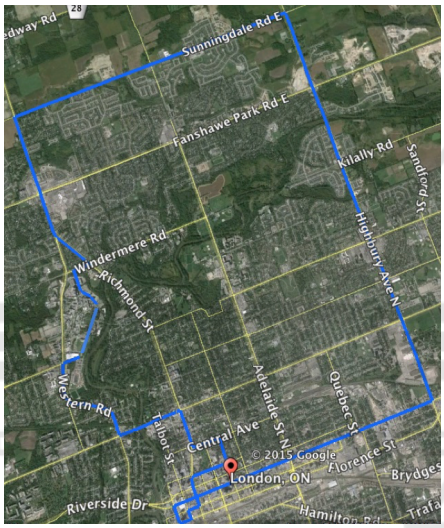

Figure 3: Map of the predetermined course for drivers, located in London, Ontario, Canada. The path includes urban and suburban driving areas and is approximately 28.5 kilometers long.

3.2 Lane Type Classification Stage

Lane type information is of great importance in guiding drivers to safely decide either to keep course in the ego lane, to change lane, to overtake, or to turn around. Our goal is to classify the detected ego lane boundaries into eight classes including dashed white, dashed yellow, solid white, solid yellow, double solid yellow, dashed-solid yellow, solid-dashed yellow, and road boundary. The road boundary type specifies the edge of the road when an actual lane marking does not exist.

Figure 4: Examples of annotated samples of our lane detection dataset.

3.2.1 ResNet101-based Lane Type Classification Model

The lane type classification stage receives the output of lane detection (14 coordinates in the image plane for each predicted ego lane boundary) as input. We first identify the ROI for each lane boundary separately. Each ROI fits the detected ego lane boundary as per its corresponding predicted coordinates. Next, we apply a projective transformation to each ROI to obtain an image where the lane marking align in the center of the resulting image. Afterwards, we crop the middle rectangular part of the transformed image that contains the lane type information. Finally, we apply our trained ResNet101 network to classify the resulting images obtained for each lane boundary into the aforementioned eight classes. Figure 5 illustrates how the lane type classification stage performs the above steps on a sample road image.

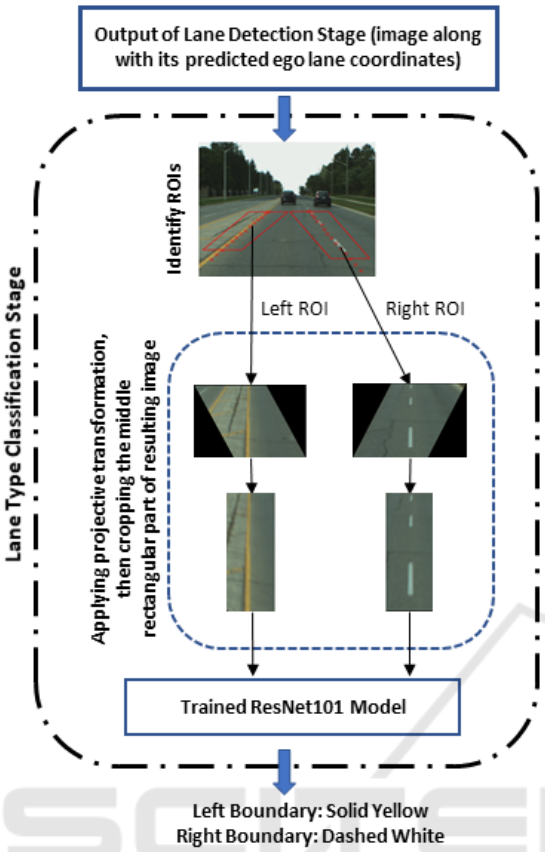

Figure 5: Visualization of the lane type classification stages, from road images to ego lane boundaries.

3.2.2 Our Dataset for Lane Boundary Types

In order to train and test our lane type classification model, we collected 10571 sample lane boundary images from the outputs of the lane detection model. These samples are inputs to our ResNet101 model, as they contain the lane type information. Figure 6 shows samples of our dataset for the eight lane boundary types.

To further enrich our lane type dataset for training, we employed two different techniques including data augmentation and a boosting method. By means of data augmentation, we expanded our dataset by creating the translated, rotated, sheared, and scaled versions of our original samples. Table 2 represents the techniques we have used to augment our data with their descriptions and ranges. To boost the performance of our trained model, we used an advanced learning method called Hard Examples Mining (HEM). HEM refers to the examples that have been misclassified by the current trained version of the model. We trained the ResNet101 model in an iterated procedure, and at each iteration, the model was applied to a number of new samples from the training data. We then added the corrections of misclassified outputs to the training set for the next iteration. Finally, the model is provided with more key samples to increase its robustness.

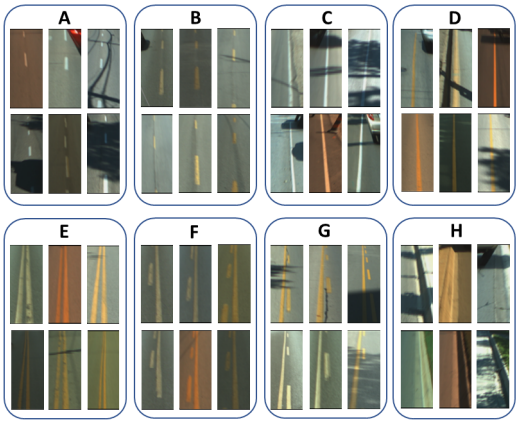

Figure 6: Lane boundary samples of our train-and-test data a) Dashed White, b) Dashed Yellow, c) Solid White, d) Solid Yellow, e) Double Solid Yellow f) Dashed-Solid Yellow, g) Solid-Dashed Yellow, h) Road Boundary.

4 EXPERIMENTAL RESULTS

To perform the experiments, we applied the model to the unseen test data extracted from our driving sequences (Beauchemin et al., 2012). To evaluate the performance of the lane detection stage, we used a metric suggested by (Chougule et al., 2018): we compute the mean error between the predicted lane coordinates generated by the lane coordinate model with the corresponding ground truth values as a Euclidean distance (in terms of pixels), for each lane boundary. For each single lane boundary, the Mean Prediction Error (MPE) is computed as follows (see Figure 7):

$$\text{MPE} = \frac{1}{14}\sum_{i=1}^{14}\sqrt{(xp_i - xg_i)^2 + (yp_i - yg_i)^2} \quad (1)$$

where (xp_i, yp_i) and (xg_i, yg_i) indicate the predicted lane coordinates and the corresponding ground truth coordinates respectively. Additionally, we investigated the performance of the following two L_1 and L_2 loss functions at the lane detection stage:

$$L_1 = \sum_{i=1}^{14}|xp_i - xg_i| + \sum_{i=1}^{14}|yp_i - yg_i| \quad (2)$$

$$L_2 = \sum_{i=1}^{14}(xp_i - xg_i)^2 + \sum_{i=1}^{14}(yp_i - yg_i)^2 \quad (3)$$

Table 2: Description of data augmentation.

Augmentation Method	Description	Range
Translate	Each image is translated in the h/v direction by a distance, in pixels	[-20, 20]
Rotate	Each image is rotated by an angle, in degrees	[-25, 25]
Shear	Each image is sheared along the h/v axis by an angle, in degrees	[-25, 25]
Scale	Each image is zoomed in/out in the h/v direction by a factor	[0.5, 1.5]

where (xp_i, yp_i) and (xg_i, yg_i) indicate the predicted lane coordinates and the corresponding ground truth coordinates respectively.

In Table 3, we report the performance of the lane detection stage described in Section 3.1 for the ego lane left/right boundaries using the aforementioned loss functions. As observed from Table 3, the L_1 loss function is superior to L_2.

Table 3: Description of our lane detection results based on the prediction error.

Loss Function	Ago Lane Boundary	MPE	Standard Deviation
L_1	Left	5.96	4.70
	Right	5.79	4.85
L_2	Left	7.39	5.55
	Right	7.16	5.42

Figure 7: Visualization of the Euclidean error between the predicted lane coordinates and the corresponding ground truth coordinates.

As described in Section 3.2, the lane type classification stage is applied to the output of the lane detection stage to recognize the detected lane boundaries and to provide a classification result. We trained a ResNet101 CNN using our dataset to verify and categorize the localized lane boundaries into eight classes of lane types. To verify the accuracy of the lane type classification stage, we computed the confusion matrix from the ResNet101 model on the test data (See Figure 8). The results show that the model reaches 94.52% of overall correct classification. This model is able to discriminate the eight lane types with less than 4.2% of mislabeling error. The lowest degree of correctly categorized classes belongs to class dashed-solid yellow, while class double solid yellow obtained 97.7%. Figure 9 displays small portions of the visual outputs from our system for the eight classes of lane boundary types.

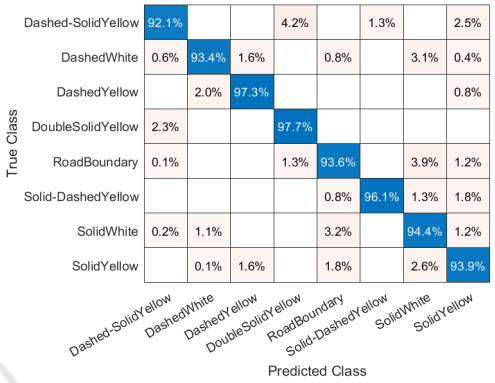

Figure 8: Confusion matrix from ResNet101 for lane type classification.

5 CONCLUSIONS

In this study, we presented a CNN-based framework to detect and classify lane types in urban and suburban driving environments. To perform lane detection and classification stages, we created an image dataset for each from sequences captured in different illumination conditions created by the RoadLAB initiative (Beauchemin et al., 2012). We also enriched our training data using data augmentation and a hard example mining strategy. To detect lanes, we used a network which generates lane information in terms of image coordinates in an end-to-end way. In the lane type classification stage, we utilized our trained ResNet101 network to categorize the detected lane boundaries into eight classes including dashed white, dashed yellow, solid white, solid yellow, double solid yellow, dashed-solid yellow, solid-dashed yellow, and road boundary. Finally, our results showed that the ResNet101 model achieved over 94% of correct lane type classifications.

Figure 9: Output samples of our experiments on the RoadLAB dataset.

REFERENCES

Ali, A. A. and Hussein, H. A. (2019). Real-time lane markings recognition based on seed-fill algorithm. In *Proceedings of the International Conference on Information and Communication Technology*, pages 190–195.

Beauchemin, S., Bauer, M., Kowsari, T., and Cho, J. (2012). Portable and scalable vision-based vehicular instrumentation for the analysis of driver intentionality. *IEEE Transactions on Instrumentation and Measurement*, 61(2):391–401.

Cheng, H.-Y., Jeng, B.-S., Tseng, P.-T., and Fan, K.-C. (2006). Lane detection with moving vehicles in the traffic scenes. *IEEE Transactions on intelligent transportation systems*, 7(4):571–582.

Chiu, K.-Y. and Lin, S.-F. (2005). Lane detection using color-based segmentation. In *IEEE Proceedings. Intelligent Vehicles Symposium, 2005.*, pages 706–711. IEEE.

Choi, H.-C. and Oh, S.-Y. (2010). Illumination invariant lane color recognition by using road color reference & neural networks. In *The 2010 International Joint Conference on Neural Networks (IJCNN)*, pages 1–5. IEEE.

Chougule, S., Koznek, N., Ismail, A., Adam, G., Narayan, V., and Schulze, M. (2018). Reliable multilane detection and classification by utilizing cnn as a regression network. In *Proceedings of the European Conference on Computer Vision (ECCV)*.

de Paula, M. B. and Jung, C. R. (2013). Real-time detection and classification of road lane markings. In *2013 XXVI Conference on Graphics, Patterns and Images*, pages 83–90. IEEE.

Ghafoorian, M., Nugteren, C., Baka, N., Booij, O., and Hofmann, M. (2018). El-gan: Embedding loss driven generative adversarial networks for lane detection. In *Proceedings of the European Conference on Computer Vision (ECCV)*.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680.

He, B., Ai, R., Yan, Y., and Lang, X. (2016). Accurate and robust lane detection based on dual-view convolutional neutral network. In *2016 IEEE Intelligent Vehicles Symposium (IV)*, pages 1041–1046. IEEE.

Hoang, T. M., Hong, H. G., Vokhidov, H., and Park, K. R. (2016). Road lane detection by discriminating dashed and solid road lanes using a visible light camera sensor. *Sensors*, 16(8):1313.

Kim, H., Park, J., Min, K., and Huh, K. (2020). Anomaly monitoring framework in lane detection with a generative adversarial network. *IEEE Transactions on Intelligent Transportation Systems*.

Kim, Z. (2008). Robust lane detection and tracking in challenging scenarios. *IEEE Transactions on Intelligent Transportation Systems*, 9(1):16–26.

Lee, C. and Moon, J.-H. (2018). Robust lane detection and tracking for real-time applications. *IEEE Transactions on Intelligent Transportation Systems*, 19(12):4043–4048.

Lee, S., Kim, J., Shin Yoon, J., Shin, S., Bailo, O., Kim, N., Lee, T.-H., Seok Hong, H., Han, S.-H., and So Kweon, I. (2017). Vpgnet: Vanishing point guided network for lane and road marking detection and recognition. In *Proceedings of the IEEE international conference on computer vision*, pages 1947–1955.

Linarth, A. and Angelopoulou, E. (2011). On feature templates for particle filter based lane detection. In *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 1721–1726. IEEE.

Liu, G., Wörgötter, F., and Markelić, I. (2010). Combining statistical hough transform and particle filter for robust lane detection and tracking. In *2010 IEEE Intelligent Vehicles Symposium*, pages 993–997. IEEE.

Liu, T., Chen, Z., Yang, Y., Wu, Z., and Li, H. (2020). Lane detection in low-light conditions using an efficient data enhancement: Light conditions style transfer. *arXiv preprint arXiv:2002.01177*.

Lo, S.-Y., Hang, H.-M., Chan, S.-W., and Lin, J.-J. (2019a). Efficient dense modules of asymmetric convolution for real-time semantic segmentation. In *Proceedings of the ACM Multimedia Asia*, pages 1–6.

Lo, S.-Y., Hang, H.-M., Chan, S.-W., and Lin, J.-J. (2019b). Multi-class lane semantic segmentation using efficient convolutional networks. In *2019 IEEE 21st International Workshop on Multimedia Signal Processing (MMSP)*, pages 1–6. IEEE.

López, A., Serrat, J., Canero, C., Lumbreras, F., and Graf, T. (2010). Robust lane markings detection and road geometry computation. *International Journal of Automotive Technology*, 11(3):395–407.

Mammeri, A., Boukerche, A., and Lu, G. (2014). Lane detection and tracking system based on the mser algorithm, hough transform and kalman filter. In *Proceedings of the 17th ACM international conference on Modeling, analysis and simulation of wireless and mobile systems*, pages 259–266.

Neven, D., De Brabandere, B., Georgoulis, S., Proesmans, M., and Van Gool, L. (2018). Towards end-to-end lane detection: an instance segmentation approach. In *2018 IEEE intelligent vehicles symposium (IV)*, pages 286–291. IEEE.

Pan, X., Shi, J., Luo, P., Wang, X., and Tang, X. (2018). Spatial as deep: Spatial cnn for traffic scene understanding. In *Thirty-Second AAAI Conference on Artificial Intelligence*.

Sani, Z. M., Ghani, H. A., Besar, R., Azizan, A., and Abas, H. (2018). Real-time video processing using contour numbers and angles for non-urban road marker classification. *International Journal of Electrical & Computer Engineering (2088-8708)*, 8(4).

Yoo, S., Seok Lee, H., Myeong, H., Yun, S., Park, H., Cho, J., and Hoon Kim, D. (2020). End-to-end lane marker detection via row-wise classification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 1006–1007.