Lane Road Segmentation Based on Improved UNet Architecture for Autonomous Driving

Hoang Tran Ngoc, Huynh Vu Nhu Nguyen, Khang Hoang Nguyen, Luyl-Da Quach
Software Engineering Department, FPT University, Cantho City, Vietnam

***Abstract*—This paper introduces a real-time workflow for implementing neural networks in the context of autonomous driving. The UNet architecture is specifically selected for road segmentation due to its strong performance and low complexity. To further improve the model's capabilities, Local Binary Convolution (LBC) is incorporated into the skip connections, enhancing feature extraction, and elevating the Intersection over Union (IoU) metric. The performance evaluation of the model focuses on road detection, utilizing the IOU metric. Two datasets are used for training and validation: the widely used KITTI dataset and a custom dataset collected within the ROS2 environment. Simulation validation is performed on both datasets to assess the performance of our model. The evaluation of our model on the KITTI dataset demonstrates an impressive IoU score of 97.90% for road segmentation. Moreover, when evaluated on our custom dataset, our model achieves an IoU score of 98.88%, which is comparable to the performance of conventional UNet models. Our proposed method to reconstruct the model structure and provide input feature extraction can effectively improve the performance of existing lane road segmentation methods.**

Keywords*—*Local binary patterns; feature extraction; UNet; semantic segmentation

I. Introduction

There has been a growing interest in autonomous driving research due to its significant impact on traffic management, the economy, and the development of self-driving cars.

The purpose of these vehicles is to imitate human driving actions through intelligent decision-making and executing various tasks such as switching lanes, preventing collisions, detecting objects, and issuing warnings for lane departure [1],[2],[3],[4],[5]. The design of autonomous driving cars involves three essential components: perception, path planning, and control [6],[7],[8],[9]. Recent advancements in sensor technology have greatly improved perception capabilities. While cameras are commonly used, the integration of additional sensors like GPS, radars, or LIDARs enhances the performance of self-driving systems [10]. The focus of autonomous navigation is on accurately detecting and identifying traffic participants, including cars, pedestrians, and surrounding objects/areas.

In particular, road detection and segmentation are crucial for autonomous driving and intelligent transportation systems as it ensures safe and efficient vehicle operation. Solutions in this area aim to reduce accidents, alleviate traffic congestion, and improve fuel efficiency.

Precise detection and recognition of roadways, encompassing boundaries and lanes, empower intelligent decision-making and enhance navigation efficiency. These advancements have the potential to greatly enhance overall transportation systems. Various datasets such as KITTI [11], Berkeley DeepDrive [12], A2D2, or those generated by the CARLA simulator [13] are utilized for a range of autonomous driving and lane segmentation tasks. Teichmann et al. Research was carried out to measure the computational time required for semantic segmentation tasks using the KITTI dataset [14]. Neven et al. focused on scene understanding using the Cityscapes dataset [15]. Similarly, real-time efforts utilizing the Cityscapes dataset involved the development of an ENet architecture [16]. Wang et al. utilized 3D LiDAR point clouds and the PointSeg architecture for real-time semantic segmentation [17]. Bai et al. explored time-critical task performance in road segmentation using the KITTI benchmark [18]. Additionally, Jang et al. aimed to explain and reduce the end-to-end delay for self-driving cars in their work [19].

In recent years, UNet is a fully convolutional network architecture that has gained popularity for lane segmentation in autonomous driving. It utilizes a U-shaped network design for accurate identification and delineation of road lanes. Studies have demonstrated its effectiveness, comparing it favorably to other methods in terms of accuracy and efficiency [20]-[21]. Giurgi et al. introduce a real-time implementation workflow for neural networks in autonomous driving, specifically focusing on road segmentation using the UNet structure with the KITTI dataset [22]. UNET's potential for improving autonomous driving systems may be seen in activities such as lane departure alerts and autonomous lane holding. However, when autonomous cars operate in tough traffic settings with high levels of noise and interference from elements such as dust, vibrations, rain, and wind, these algorithms become susceptible to disruptions, resulting in decreasing lane segmentation accuracy. The existence of these external elements severely impairs the effectiveness of lane segmentation algorithms, resulting in less than ideal results.

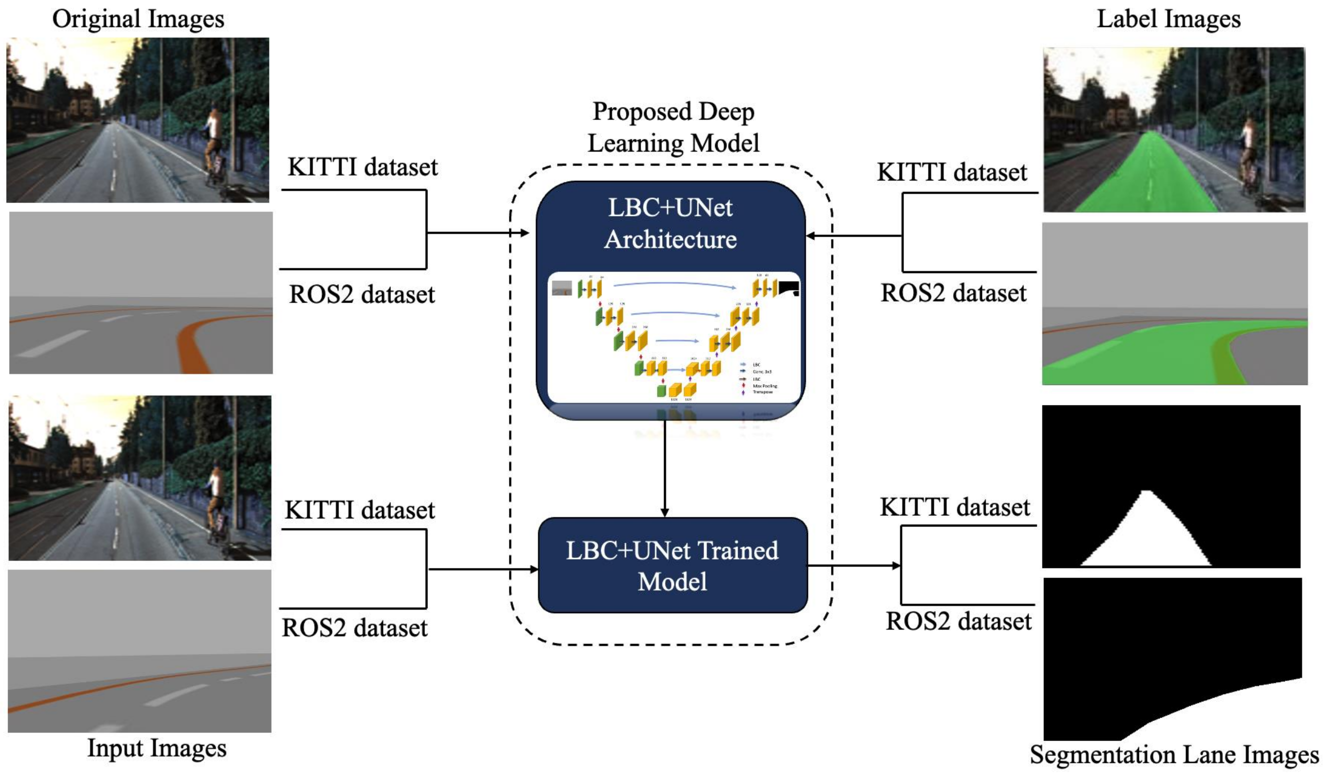

Fig. 1. Our proposed system.

To address these challenges, we propose the combination of LBC layers with UNet in this paper to improve lane segmentation performance in noisy traffic environments. The LBC layers integrate local binary patterns into the convolutional neural network (CNN) architecture [23], enhancing the ability to extract fine-grained structural information and model image representations. These layers have shown potential in applications that require robust feature extraction and learning, particularly in scenarios with limited training data or noisy environments. We will compare the performance of the proposed method to earlier approaches in order to quantify the improvement in accuracy on two datasets: KITTI and our own gathered dataset in a ROS2 robot simulation environment. Fig. 1 describes our proposed system.

The subsequent sections of the paper follow the following structure. Section II provides an introduction and summary of relevant research pertaining to lane segmentation. Section III describes the model's architecture in detail, outlining the integration of LBC and skip connections to enhance local feature extraction. Section IV focuses on the experimental implementation, dataset utilization, and a comparative analysis of various models. Finally, the paper concludes by summarizing key findings and proposing future avenues for advancement.

II. Related Work

In recent years, there have been several advancements in the field of road lane segmentation. One notable approach is the DeepLab method proposed by Chen et al. [24], which combines deep convolutional nets with fully connected conditional random fields for accurate semantic image segmentation. Another approach is the ENet architecture introduced by Paszke et al. [25], specifically designed for real-time semantic segmentation tasks. Additionally, Pan et al. [26] proposed LaneNet, a spatial CNN architecture for traffic scene understanding, focusing on lane segmentation. Fu et al. [27] presented SCNN, a parallel CNN model that explores the road scene in depth for precise road segmentation. In a recent study, Giurgi et al. developed a unique method employing the UNet architecture, which demonstrated appreciable increases in lane segmentation accuracy. These studies are a limited exploration of image segmentation in the complex context of autonomous vehicles. Lane recognition is a crucial task in autonomous driving, and existing approaches confront difficulties owing to the complexities of the input pictures. To address this, we propose the use of LBC layers to reduce complexity and increase processing speed. Building upon this, we present an enhanced UNet model incorporating skip connections and LBC layers for improved lane markings recognition while minimizing training time. Comparative analysis and evaluation metrics, such as IoU, Dice coefficient, and precision, are employed to assess the accuracy and efficiency of the proposed models. Our study focuses on developing efficient and accurate segmentation models for lane recognition in autonomous driving scenarios.

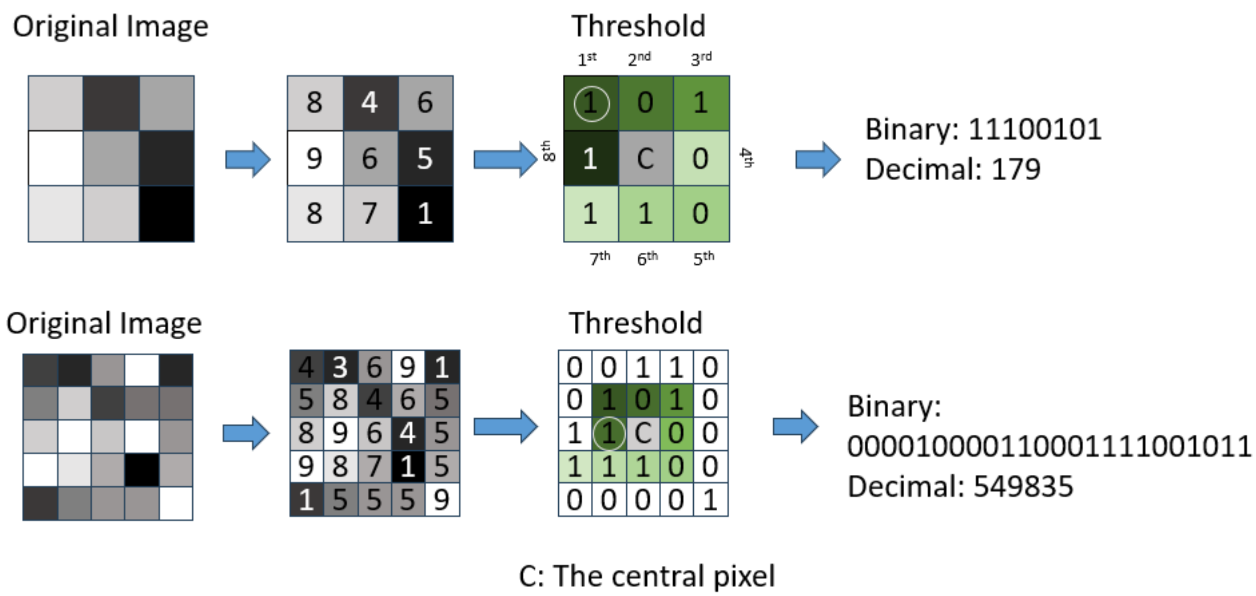

Fig. 2. Visualizing local binary patterns (LBP) operation: exploring 3x3 and 5x5 local dimensions.

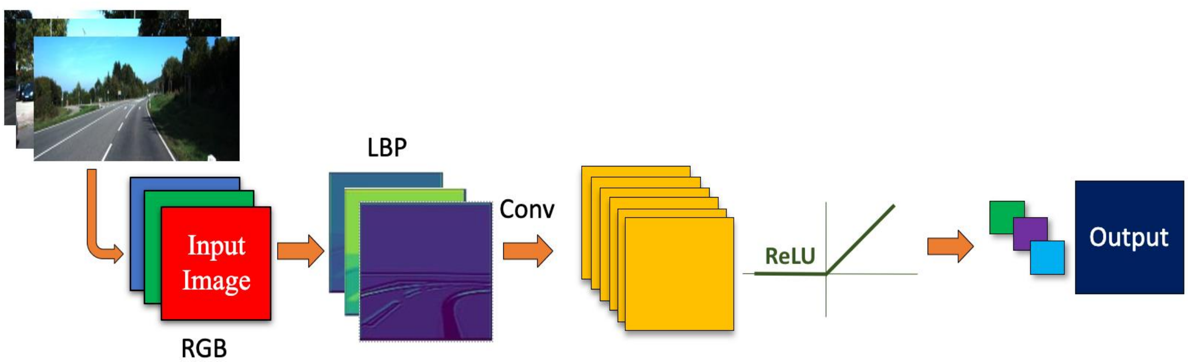

Fig. 3. Basic module local binary convolution.

III. Proposed Method

A. Local Binary Pattern and its Convolution Variants

1) Local binary patterns: Local Binary Patterns (LBPs) is an image processing technique used for capturing local patterns by comparing pixel values in small neighborhoods [28]. It is commonly employed in face recognition and object detection. LBPs operate by selecting a neighborhood around each pixel and converting the pixel values into a binary string. By comparing the values of surrounding pixels with the central pixel, the binary string is constructed. The equation for calculating the brightness intensity of LBPs can be described as follows:

$$LBP_P = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \quad (1)$$

$$s(z) = \begin{cases} 1, & if\ z \geq 0 \\ 0, & if\ z < 0 \end{cases}$$

where g_p is the neighbor pixel intensity value; g_c is the center pixel intensity value. P is the number of neighbor pixels. Z is the result of g_p minus g_c.

This string represents the local spatial patterns within the neighborhood. Analyzing the distribution of these binary patterns provides valuable insights into local variations in brightness, which can be utilized for tasks like contrast enhancement and object recognition. LBPs are a compact representation of local patterns and find wide usage in computer vision applications. Fig. 2 illustrates the basic operation of LBPs, demonstrating their functionality for each pixel in an image with local dimensions of 3x3 and 5x5.

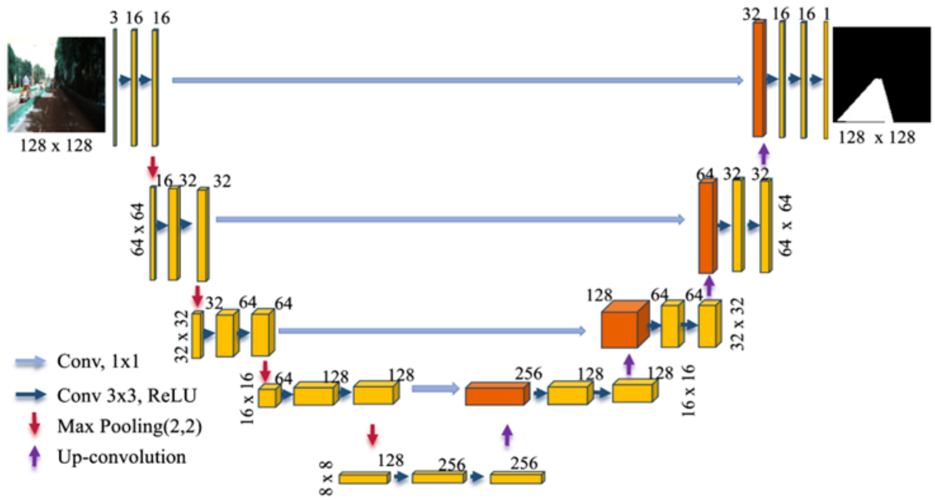

Fig. 4. Conventional UNet architecture.

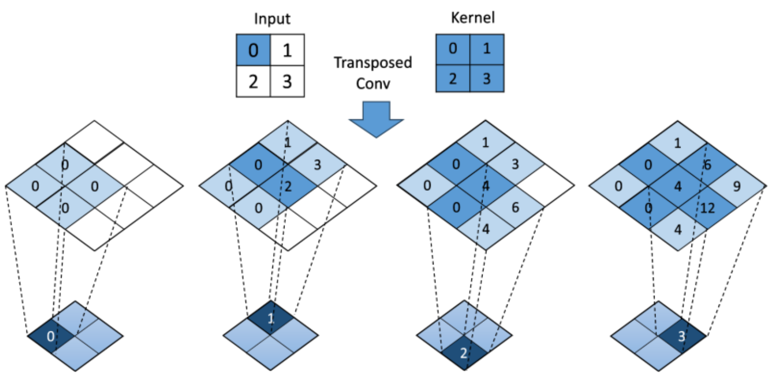

Fig. 5. Transposed convolution with a 2x2 kernel.

2) Local binary convolution: The Local Binary Convolution (LBC) layer convolves a filter over an input image, converting pixel values into binary patterns called Local Binary Patterns (LBPs). These patterns record structural, morphological, and textural information. During training, gradients can be backpropagated through the layer's anchor weights, while the learnable 1x1 filters are updated. The anchor weights can be generated deterministically or stochastically, allowing for diversified filters and fine-grained control overweight sparsity. The LBC layer efficiently extracts meaningful features for tasks like object detection and recognition in computer vision. In Fig. 3, we present the basic model of LBC that we use in this paper. The input image is first separated into three RGB channels through the LBP 3x3 local dimensions. After that, a convolutional layer is used for additional processing, followed by the ReLU activation function to extract key traits of road segments.

B. Improved Lane Road Segmentation

1) Conventional UNet model

a) *UNet architecture:* UNet is an architecture for semantic segmentation introduced by Olaf Ronneberger et al. [23]. This is a widely used architecture for road segmentation that combines encoding and decoding paths. It utilizes max pooling for down-sampling and transposed convolution for up-sampling. Skip connections play a crucial role in preserving information between the encoding and decoding stages.

Fig. 4 illustrates the structure of basic UNet architecture. It is made up of a left side encoding path and a right-side decoding path. Max pooling techniques are used in the encoding process to gradually lower the spatial resolution while raising the number of feature channels. This helps extract abstract features related to road structures. The decoding path employs transposed convolutions to up-sample the feature maps and increase the spatial resolution. This results in a dense output map representing the road segmentation mask. Skip connections establish direct connections between corresponding encoding and decoding layers, allowing detailed information to flow between them. This facilitates the reconstruction of accurate road segmentations.

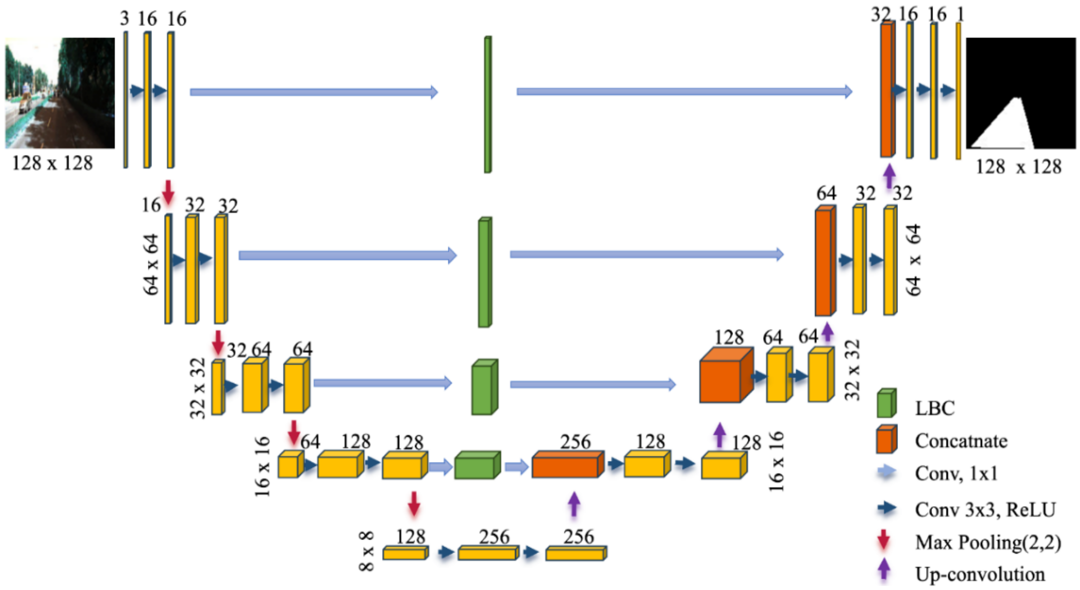

Fig. 6. Proposed UNet architecture.

UNet's combination of encoding and decoding paths with skip connections enables exceptional performance in road segmentation. It effectively captures both local and global context information, enabling precise delineation of road regions in images.

b) Up-convolution with transposed convolution approach: Transposed convolution, also known as deconvolution or fractionally stridden convolution, is a technique used to up-sample feature maps in convolutional neural networks. It is the reverse operation of the standard convolution operation and is commonly used in the decoding path of architectures like UNet. Fig. 5 shows the operation of Transposed Convolution with a 2x2 kernel. The parameters required to design a Transposed Convolution to achieve the desired output size can be described using (2):

$$O_{\text{size}} = (T_{\text{size}} - 1) \cdot s + H_{\text{size}} - 2 \cdot p \quad (2)$$

where O_{size} is the desired size of the output feature map; T_{size} refers to the size of the input feature map; s is the stride value used in the Transposed Convolution operation; H_{size} represents the size of the kernel used in the operation; and p refers to the padding applied to the input feature map. However, in noisy and challenging environments such as transportation, a lane segmentation system with a robust feature extractor needs to be investigated. Therefore, we have developed an algorithm that we propose in the next section.

2) Improved design of the UNet model: The improved design of the UNet model aims to enhance the segmentation accuracy compared to the classic UNet architecture. To achieve this, we have introduced the Local Binary Convolution (LBC) layer into the skip connections of the UNet model. The design structure is illustrated in Fig. 6.

The encoding structure (left side) consists of four blocks. Each block includes a series of convolutional layers with ReLU activation, followed by max-pooling operationsThe output of each block is created by gradually applying the pooling and convolutional layer operations. The LBC layers are integrated into the skip connections of the UNet model to capture local binary patterns and improve the segmentation performance. These skip connections establish direct connections between the corresponding encoder and decoder layers. The four blocks handle the up-sampling and concatenation operations necessary for the skip connections. They take the inputs from the corresponding pooling layers and transpose convolutional layers to up-sample the feature maps. Finally, the model is compiled with the Adam optimizer and binary cross-entropy loss. The metrics used for evaluation include the Intersection over Union (IoU). This improved UNet model with LBC layers in the skip connections offers enhanced capabilities for accurately segmenting road images. In the upcoming section, we will evaluate the results and accuracy of this model using two datasets. The purpose is to demonstrate the superiority of the proposed method in comparison to existing approaches.

IV. Experimental Results

A. Experimental Setting

The experiments were conducted to train and evaluate multiple models using two distinct datasets. Each model underwent 100 epochs of training, and performance was assessed based on metrics such as IOU, Validation IOU, Loss, and Validation Loss. The training process utilized a computer with Ubuntu 20.04, an Intel i7 3.4 GHz CPU, an Nvidia GTX 3060 Laptop, and 32 GB RAM. The implementation was carried out in Python 3.9.13, employing the Conda 22.9.0, CUDA 11.7, Tensorflow 2.10.0, and Keras 2.10.0 libraries.

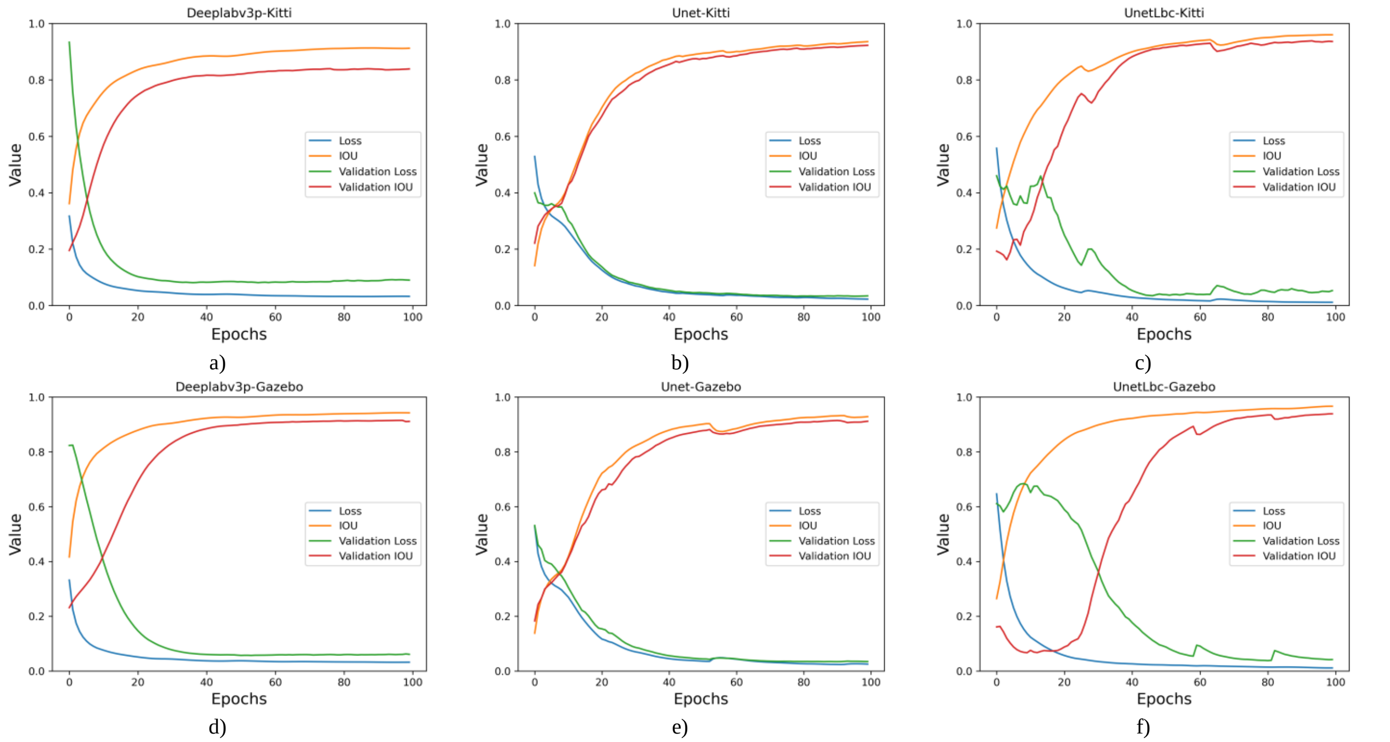

Fig. 7. Comparison of segmentation performance among DeepLabv3, UNet, and proposed model based on IOU, Validation IOU, Loss, and Validation Loss. a), b), c) using KITTI dataset, and d), e), f) using Gazebo/ROS2 dataset.

B. Datasets and Evaluation Metrics:

1) Datasets: The research paper utilizes two datasets, namely KITTI [29] and a dataset created from simulated lanes in the Gazebo/ROS2 environment of our laboratory. The KITTI dataset is employed primarily for unmarked lane segmentation in urban areas, comprising 800 training images and 200 test images. On the other hand, the second dataset involves the Donkey self-driving car, which operates in the Gazebo/ROS2 3D simulation environment. The car is controlled using a driving wheel joystick to maintain lane position. The car is equipped with a front-facing camera that captures images, and the ROS2 controller records these images at a rate of 5 frames per second for training data. A total of 1000 images were collected for this dataset. The data split ratio is 80% for training data and 20% for validation data.

2) Evaluation metrics: In image segmentation, Intersection over Union (IoU) [25] is a primary metric used to evaluate the accuracy of models. Unlike in object detection, where IoU serves as a supplementary metric, it plays a crucial role in the pixel-level analysis of segmentation masks. The definitions of true positive (TP), false positive (FP), and false negative (FN) differ slightly in image segmentation, considering the pixel-wise intersection and logical operations between the ground truth and segmentation masks. IoU is determined in image segmentation by dividing the intersected area by the sum of the ground truth and prediction areas using the TP, FP, and FN areas, or pixel counts.

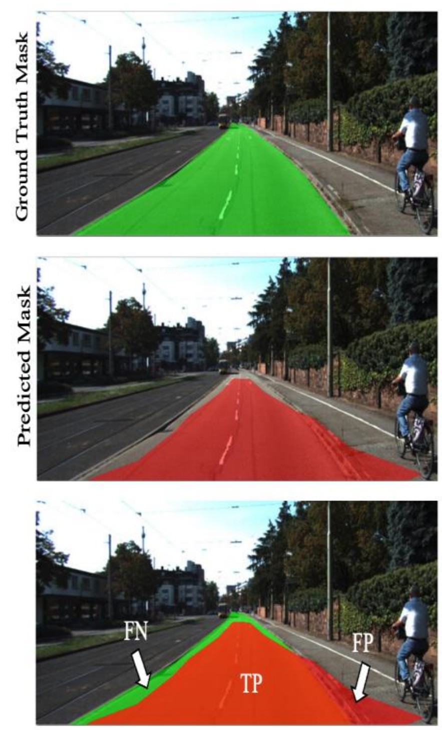

Fig. 8. Example of the IOU equation.

This metric helps assess the effectiveness of models in accurately segmenting objects and regions of interest in images. The equation is shown below:

$$IoU = \frac{TP}{FP + TP + FN} \qquad (3)$$

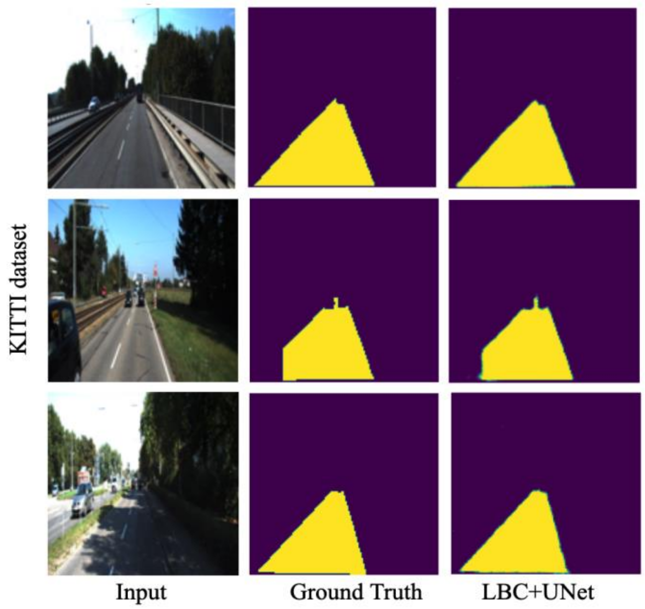

Fig. 9. Lane road segmentation results of the proposed model using the KITTI dataset.

TABLE I. SEGMENTATION RESULTS ON KITTI DATASET

Approach	IOU	Validation IoU	Loss	Validation Loss
DeeplabV3+	0.93	0.9292	0.014	0.0670
UNET	0.95	0.9211	0.021	0.0742
LBC+UNET	0.97	0.9442	0.012	0.0437

with:

$$TP = GT.X$$

$$FP = (GT + X) - GT$$

$$FN = (GT + X) - X$$

where TP, FP, and FN indicate the True Positive, False Positive, and False Negative numbers, respectively; GT is the region's Ground Truth; X is segmentation mask overlap. Fig. 8 shows an example of IoU on the actual input image.

During training, the Loss function is used to measure the difference between the model's predicted output and the actual output value. The goal is to find a way to minimize the loss function to make a more accurate prediction. The equation is shown below (4):

$$Log\ L = \frac{1}{E}\sum_{i=1}^{E} -[y_i log(v_i) + (1 - y_i) log(1 - v_i)] \qquad (4)$$

where L represents the Binary Cross Entropy Loss, E is the number of samples in the dataset, y_i represents the true label (ground truth) for the i^{th} sample (0 or 1), v_i represents the projected probability for the i^{th} sample, and log represents the natural algorithm.

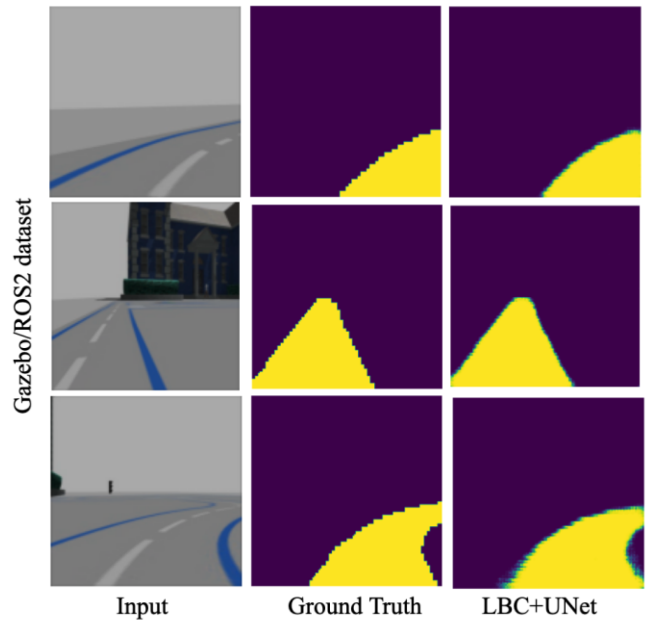

Fig. 10. Lane road segmentation results of the proposed model using the Gazebo/ROS2 dataset.

TABLE II. SEGMENTATION RESULTS ON GAZEBO/ROS2 DATASET

Approach	IOU	Validation IOU	Loss	Validation Loss
DeeplabV3+	0.94	0.933	0.0223	0.0319
UNET	0.94	0.948	0.0238	0.0388
LBC+UNET	0.988	0.96	0.0122	0.0144

3) Results and discussion: DeepLabV3+ [30], UNet, and LBC+UNet segmentation outcomes were compared using Intersection over Union (IoU) and loss values on two datasets, the KITTI dataset and the ROS2 dataset. Afterwards 100 training epochs, the three techniques' highest IoU values were as follows: On the KITTI dataset, DeepLabV3+, UNet, and LBC+UNet each had an IoU of 0.93, 0.95, and 0.97, respectively. In addition, using the KITTI dataset, the three techniques produced the following loss values: LBC+UNet had a loss of 0.012, DeepLabv3+ had a loss of 0.014, and UNet had a loss of 0.021.

The models' performance was further examined using the Gazebo/ROS2 dataset. The IoU values attained by the three methods were as follows after 100 training epochs: DeepLabV3+ had an IoU of 0.94, UNet had an IoU of 0.94, and LBC+UNet had an IoU of 0.988. DeepLabV3+ had a loss of 0.0223, UNet had a loss of 0.0238, and LBC+UNet had a loss of 0.0122. These were the loss numbers produced by the three methodologies.

The proposed method, LBC+UNet, achieved the highest segmentation results in terms of IoU for both the KITTI dataset and the Gazebo/ROS2 dataset. This can be observed from the results presented in Fig. 7, where the performance of each epoch is displayed. The IoU Validation and Loss Validation metrics are also included in Tables I and II, respectively, to facilitate a clearer comparison. The proposed method demonstrated the highest accuracy in terms of IoU and the lowest error in terms of the loss function.

The experimental results of road lane segmentation using the proposed method are illustrated in Fig. 9 and Fig. 10. We can observe a high level of accuracy, exceeding 95%, which can be attributed to the utilization of the feature extraction capabilities of LBC combined with the UNet architecture. The performance of the proposed method is consistently strong on both the KITTI and Gazebo/ROS2 datasets, demonstrating good segmentation results and high accuracy. Moreover, the proposed method was tested on both simulated and real-world datasets, confirming its effectiveness in road lane segmentation. This advancement supports autonomous driving systems and contributes to reducing accidents by providing higher accuracy.

V. Conclusion

We have successfully used a combination of the UNET architecture and the LBC feature extractor in this article to improve the accuracy of road lane segmentation for autonomous driving support. The proposed method has demonstrated superior accuracy compared to conventional approaches such as DeepLabV3+ and the classical UNET method. Through comprehensive evaluations using well-known datasets, including KITTI and our custom-built dataset based on the ROS2 robot simulation model, the proposed model has proven its effectiveness in both simulated and real-world scenarios. The application of our proposed model holds great potential for various domains, including simulation and practical implementations. Looking ahead, further advancements in road lane segmentation will focus on fulfilling the demand for more refined lane segmentation, particularly the differentiation between drivable and non-drivable areas. This ongoing development will significantly contribute to the improvement of self-driving systems by providing precise lane segmentation for enhanced decision-making and safer navigation.

References

[1] Yaqoob, L. U. Khan, S. M. A. Kazmi, M. Imran, N. Guizani, and S. C. Hong, ‘‘Autonomous driving cars in smart cities: Recent advances, requirements, and challenges,’’ IEEE Netw., vol. 34, no. 1, pp. 174–181, Jan./Feb. 2020.

[2] S. P. Narote, P. N. Bhujbal, A. S. Narote, and D. M. Dhane, ‘‘A review of recent advances in lane detection and departure warning system,’’ Pattern Recognit., vol. 73, pp. 216–234, Jan. 2018.

[3] Hoang Tran Ngoc and Luyl-Da Quach, “Adaptive Lane Keeping Assist for an Autonomous Vehicle based on Steering Fuzzy-PID Control in ROS” International Journal of Advanced Computer Science and Applications(IJACSA), 13(10), 2022

[4] V. D. Nguyen, T. D. Trinh and H. N. Tran, "A Robust Triangular Sigmoid Pattern-Based Obstacle Detection Algorithm in Resource-Limited Devices," in IEEE Transactions on Intelligent Transportation Systems, vol. 24, no. 6, pp. 5936-5945, June 2023

[5] H. K. Hua, K. H. N., L. Quach, and H. N. Tran. 2023. “Traffic Lights Detection and Recognition Method using Deep Learning with Improved YOLOv5 for Autonomous Vehicle in ROS2”. In Proceedings of the 2023 8th International Conference on Intelligent Information Technology (ICIIT '23). Association for Computing Machinery, New York, NY, USA, 117–122..

[6] J. Vargas, S. Alsweiss, O. Toker, R. Razdan, and J. Santos, “An overview of autonomous vehicles sensors and their vulnerability to weather conditions,” Sensors (Basel, Switzerland), vol. 21, pp. 1–22, August, 2021.

[7] M. Buehler, K. Iagnemma, and S. Singh, “The darpa urban challenge: Autonomous vehicles in city traffic, george air force base, victorville, california, usa,” in The DARPA Urban Challenge.

[8] H. T. Vo, H. N. Tran, and L. Quach, “An Approach to Hyperparameter Tuning in Transfer Learning for Driver Drowsiness Detection Based on Bayesian Optimization and Random Search” International Journal of Advanced Computer Science and Applications(IJACSA), 14(4), 2023.

[9] P. H. Phan, A. Q. Nguyen, L. Quach, and H. N. Tran. 2023. “Robust Autonomous Driving Control using Auto-Encoder and End-to-End Deep Learning under Rainy Conditions”. In Proceedings of the 2023 8th International Conference on Intelligent Information Technology (ICIIT '23). Association for Computing Machinery, New York, NY, USA, 271–278.

[10] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, “Carla: An open urban driving simulator,” in Proceedings of the 1st Annual Conference on Robot Learning, pp. 1–16, 2017.

[11] J. Fritsch, T. Kuhnl, and A. Geiger, “A new performance measure and evaluation benchmark for road detection algorithms,” 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013), pp. 1693–1700, 2013.

[12] F. Yu, H. Chen, X. Wang, W. Xian, Y. Chen, F. Liu, V. Madhavan, and T. Darrell, “Bdd100k: A diverse driving dataset for heterogeneous multitask learning,” 2018.

[13] M. Teichmann, M. Weber, J. Zollner, R. Cipolla, and R. Urtasun, ¨ “Multinet: Real-time joint semantic reasoning for autonomous driving,” pp. 1–10, 12 2016.

[14] D. Neven, B. Brabandere, S. Georgoulis, M. Proesmans, and L. Van Gool, “Fast scene understanding for autonomous driving,” pp. 1–5, 08 2017.

[15] A. Paszke, A. Chaurasia, S. Kim, and E. Culurciello, “Enet: A deep neural network architecture for real-time semantic segmentation,” pp. 1–10, 06 2016.

[16] Y. Wang, T. Shi, P. Yun, L. Tai, and M. Liu, “Pointseg: Real-time semantic segmentation based on 3d lidar point cloud,” pp. 1–10, 07 2018.

[17] L. Bai, Y. Lyu, and X. Huang, “Roadnet-rt: High throughput cnn architecture and soc design for real-time road segmentation,” IEEE Transactions on Circuits and Systems I: Regular Papers, pp. 1–11, 11 2020.

[18] W. Jang, H. Jeong, K. Kang, N. Dutt, and J.-C. Kim, “R-tod: Realtime object detector with minimized end-to-end delay for autonomous driving,” pp. 1–14, 10 2020.

[19] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. Cham, Switzerland: Springer, pp. 234–241, 015.

[20] L. -A. Tran and M. -H. Le, "Robust U-Net-based Road Lane Markings Detection for Autonomous Driving," 2019 International Conference on System Science and Engineering (ICSSE), Dong Hoi, Vietnam, pp. 62-66, 2019,

[21] D. -V. Giurgi, T. Josso-Laurain, M. Devanne and J. -P. Lauffenburger, "Real-time road detection implementation of UNet architecture for autonomous driving," 2022 IEEE 14th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP), Nafplio, Greece, pp. 1-5,2022.

[22] F. J. Xu, V. N. Boddeti, and M. Savvides, "Local Binary Convolutional Neural Networks," Machine Learning, Jul. 2017.

[23] L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Semantic Image Segmentation with Deep Convolutional Nets and Fully

Connected CRFs," in International Conference on Learning Representations (ICLR), 2015.

[24] A. Paszke, A. Chaurasia, S. Kim, and E. Culurciello, "ENet: A Deep Neural Network Architecture for Real-Time Semantic Segmentation," in Conference on Neural Information Processing Systems (NIPS), 2016.

[25] X. Pan, J. Shi, P. Luo, X. Wang, and X. Tang, "Spatial As Deep: Spatial CNN for Traffic Scene Understanding," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.

[26] X. Fu, J. Cao, and Z. Li, "Look Deeper into the Road: A Parallel CNN for Road Scene Segmentation," in IEEE International Conference on Computer Vision (ICCV), 2017.

[27] T. H. Rassem and B. E. Khoo, “Completed local ternary pattern for rotation invariant texture classification,” Sci. World J., vol. 2014, pp. 1–10, Jan. 2014.

[28] Ronneberger O., Fischer P., Brox T. “U-net: Convolutional networks for biomedical image segmentation,” International Conference on Medical image computing and computer-assisted intervention, Springer, Cham, pp. 234–241, 2015.

[29] The KITTI Vision Benchmark Suite (cvlibs.net).

[30] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. 2018. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE transactions on pattern analysis and machine intelligence Vol. 40, pp. 834-848, April (2018).