Advanced Lane Detection with Ensemble Models: A Hybrid Approach Leveraging ResNet and VGG Variants for Feature Extraction

Sourav Mondal
Department of Computer Science and Engineering
Vignan's Foundation for Science, Technology and Research
Vadlamudi, India
svml_cse@gmail.com

M. Lavanya
Department of Computer Science and Engineering
Vignan's Foundation for Science, Technology and Research
Vadlamudi, India
221fa04230@gmail.com

Sunil Babu Melingi
Department of Computer Science and Engineering
Vignan's Foundation for Science, Technology and Research
Vadlamudi, India
sunil.babu.m@gmail.com

S. Naga Phanindra
Department of Computer Science and Engineering
Vignan's Foundation for Science, Technology and Research
Vadlamudi, India
221fa04577@gmail.com

Burla Sangeetha
Department of Computer Science and Engineering
Vignan's Foundation for Science, Technology and Research
Vadlamudi, India
221fa04634@gmail.com

Nikhil Tripathy
Department of Computer Science and Engineering
Vignan's Foundation for Science, Technology and Research
Vadlamudi, India
221fa04727@gmail.com

***Abstract*—Identifying lanes is an integral aspect of autonomous driving, which involves a road scene's effective feature extraction. Therefore, in this paper, a performance comparison of the different CNN models targeting lane detection tasks including, but not exclusive to, ResNet-101, ResNet-50, VGG-16, and VGG-19 is presented. The ResNet family of models performs well due to the residual connections, enabling better feature-extracting capability at greater depths. At the same time, spatial detail capture is the strong point of the VGG family of networks, but at a higher cost in processing power. In addition, the potential of XGBoost to boost lane detection performance while using features extracted from a CNN is investigated. This analysis shows the performance and efficiency trade-offs that exist and assists in determining the best models that can be integrated into autonomous driving systems.**

***Index Terms*—Lane detection, Autonomous driving, Convolutional neural networks (CNNs), ResNet-101, ResNet-50, VGG-16, VGG-19, Feature extraction, Residual connections, Deep learning, XGBoost, Spatial detail capture, Performance comparison, Efficiency trade-offs, Computer vision, Autonomous systems.**

I. Introduction

Over the last few years, the development of deep learning methods has allowed for lane detection systems to be very effective, which are extremely important for autonomous vehicles and advanced driver assistance systems (ADAS) as well. It is known that among the many neural network architectures convolutional neural networks such as ResNet and VGG have successfully extracted valuable information from complicated images of roads which has contributed to effective lane detection systems. Furthermore, this process is aided by boosting algorithms like XGBoost that improve the classification of features from the Regions of Interest CNNs making the detection system more efficient and precise. The other well-known computer vision model is the ResNet family of networks which utilizes residual connections to enable extremely deep networks to be trained. ResNet-101, and ResNet-50 are all popular models of different depths (101, and 50 layers respectively). Each of the models has its own advantages in terms of efficiency and computational complexity and provides different performance. These models employ skip layers, which allow the network to construct so-called residual mappings and thereby decrease the time required for training the network while allowing the training of deeper networks. The computation involved in training ResNet-101 is higher, although its coverage of intricate details is better than other less deep networks such as ResNet-50 which is a midpoint scenario between network depth and computation involved. Likewise, VGG (Visual Geometry Group) networks, mainly VGG-16 and VGG-19, are also widely used for image classification owing to their uncomplicated but highly efficient architecture. Marine Transportation Resource Catalog (MTRC) version VGG-16 has a total of sixteen layers with weights, while VGG-19 increases this number to nineteen layers. These architectures do not use large-sized convolution filters but instead rely on smaller structures (3x3), which enable deep features vertical and horizontal without compressing the image. In VGG models, which are less complex than Resnet, more parameters are needed thus creating higher costs in the operations, which in return makes them not suitable for use in real-time lane detection systems.

On the other hand, as opposed to these CNNs, XGBoost (Extreme Gradient Boosting) is a boosting algorithm that is very effective in terms of speed and performance when dealing with structured data. XGBoost was mostly used for tabular data but it can also be embedded in the lane detection systems as a classifier or a regressor on top of the deep bones of VGG and ResNet so as to assist in the interpretation of results even better. The purpose of this competitive analysis is to evaluate several lane detection models namely ResNet-101, ResNet-50, VGG 16, and VGG 19, based on their capacity to capture and efficiently process lane features. Also, the importance of XGBoost in this hybrid framework will be outlined showing its importance in improving the accuracy of lane detections by using the deep feature representations produced by these CNN models.

Fig. 1. Lane

II. Literature Review

Nowadays, lane tracking is increasingly useful, particularly because of the advent of autonomous driving technology. According to our study, we found that the LIDAR-RADAR fusion technology along with the CNN display, Artificial Intelligence, and training under environmental conditions have been the most successful methods that were put forward for lane detection improvement. Under the circumstances of the lane existing, some vehicles being on the other side of the road, missing lane markings, and trees' shadows reflecting SUVs as complicated with obstacles, these methods operate successfully, respectively.
For the CLLLD, on the one hand, the journal paper published in 2024 by way of a Contrastive Learning approach for the Lane Detection method, which is CLLLD, the technique has demonstrated positive results, achieving an accuracy of 96.81% on the TuSimple dataset. Of the analysis, when the CLL method was employed in the CuLane dataset, the accuracy was notably decreased (to 4%), marking the difficulty in the detection of hidden lanes, missing larger areas appearing in highly cluttered scenarios, and imitating the model which was not outperforming in terms of better results [1].

Automated deep neural network systems like PolyLaneNet and 3D ResNet50, when combined with temporal data, are also successful in lane detection even though the architecture does not provide real-time information. The percent of 91.34 was achieved, but these systems have shown to be chaotic, not finding their footing, and were very expensive in terms of computer resources, especially in real-time. However, the interpretation is that while the level of detail is high, the application of such devices in real time can still be problematic [2].

A highly efficient lane detection SGPLane which employs GANet and grid points for lane detection worked well with a 96.84% accuracy on TuSimple but the answer on the Shougang dataset was lower, about 76.85%. It has to fight with the problem of lanes adjacent to each other because the usual setting of the parameters will twist the lane boundaries until they are completely off sight. Clearly, to achieve precise lane detection along closely positioned lanes in real-world situations, additional attempts are to be made [3].

An Improved Lane Detection and Lane Departure Warning Framework for ADAS, Sliding Discrete Cosine Transform (SDCT), which is absorbed in mixed with the blob analysis, deals with the curve in the lane as well as the warning of the departure. These techniques appeared to be trustworthy and accurate and good picket wood for road safety. Systems reached good accuracy on datasets like Caltech and KIST, but the curve of road markings can be faint or sharp. Are models still in trouble with sustaining their performance? Our research also focused on hybrid models and thus not only Jetson but we also concentrated on several other methods [4].
. LATR: 3D Lane Detection from Monocular Images with Transformer, The conference paper of LATR, which has namely the year 2023 and is titled "3D Lane Detection from Monocular Images with Transformer", was written on the subject of the novel way of using transformers in 3D lane detection over monocular images. When tested on Apollo, OpenLane, and ONCE-3DLanes LATR worked and worked exceedingly well as an end-to-end detector through the use of 3D-aware front-view features without any view transformation. The incorporation of the cross-attention mechanism helped to get a higher F1 score of 11.4 points on the OpenLane dataset. On the other hand, the model enables us to deal with the depth ambiguity by analyzing the monocular images more profoundly; hence it helps to reduce misalignment between the surrogate feature map of the original image. [5].
The novel and emerging method, FLLENet, a real-time lane detection method for low-light conditions to improve advanced driver assistance systems, will be presented in 2024. A real-time lane detection method FLLENet is utilized under low-light conditions to achieve better advanced driver assistance

systems. FLLENet comprises a low-light enhancement module and an attention mechanism that are able to get through F1 scores of 76.90 in the CULane dataset and 78.91 in the NightLane dataset respectively. However, the tool has computational challenges of high costs and requires a large expenditure of data to be annotated. [6].

A 2022 paper embedded within the Jetson Xavier NX by NVIDIA "Deep embedded hybrid CNN–LSTM network for lane detection" deals with a deep hybrid technology that uses a combination of the CNN and the LSTM for lane detection. It is a system that is being built for a Lane Departure Warning System (LDWS) using the NVIDIA Jetson Xavier NX. This hierarchical CNN–LSTM model, in which the learned features with dropout layers are used to enhance the system's performance, achieves an accuracy of 96.36%, a recall of 97.54%, and an f1 score of 97.42 respectively. [7].

Fast Multi-Lane Detection Based on CNN Differentiation for ADAS/AD, 2023 journal paper presents a fast multi-lane detection method utilizing CNN differentiation for advanced driver assistance systems (ADAS) and automated driving (AD). In this case, the primary focus is on the semantic lane line rather than the legal lane marking of the road and the method has been evaluated on TuSimple and CuLane datasets with an accuracy of 96.01% on the TU simple dataset. Nevertheless, the validity of the approach may decrease in highly difficult environments, or if the lane markings are too faded or blurry [8].

A Hybrid Lane Detection Model for Wild Road Conditions 2023 journal paper presents a hybrid lane detection model specifically designed for wild road conditions using TuSimple and Indian Lane Dataset (ILD). This approach integrates a CNN model along with an encoder-decoder network in which lane marking detection is resolved via binary segmentation. It reports a 95.19% accuracy score on the TuSimple dataset which deteriorates distinctly under low light accuracy dropping by 1.73% from best-existing accuracies. [9].

A two-stage noise features filtering and clustering approach for lane detection systems was used together with deep learning systems in the journal paper published in 2022. Classic systems rely on the manipulation of edges or colors by applying various thresholds while deep learning systems consider lane markings either as objects or as something that can be segmented. It is the two stages of filtering that this paper presents which reduce the level of noise in the features thus increasing the level of precision. The model was tested on standard datasets such as TuSimple and CuLane, where it attained 94.4% and 89.3% accuracy, respectively. While the model is quite accurate, it still has some issues with real-time capabilities that must be worked on for it to be useful in practical situations. [10].

Citadis review paper explains studies considering lane detection algorithms with a cutting-edge review of tests' performances of common lane detection algorithms, 2024. Studies show that CNN plays its role by recognizing what lane markings would look like in the images and classifying together similar patterns, Murray et al 1995. While this method was successful on the TuSimple and CuLane datasets with the rates of 93.1% and 91.25% respectively, it worked poorly under adverse conditions particularly due to heavy rains and fog, hence some revisions should be done in those aspects. [11].

This study aims to develop a lightweight lane detection method built upon image processing and deep learning, recorded in the journal titled MLP Finally Efficient Convolutional Neural Network for Lane Detection, 2023. A hybrid MLP block is defined to be used with a convolutional neural network to perform lane detection based on segmentation, classification, and performance parameter approaches. Though specific performance measures are not included in the paper, the authors claim there is a slight drop in the accuracy performance because of the smaller number of parameters and hence lower processing time which is geared towards improving the efficiency of lane detection systems. [12].

III. Methodology

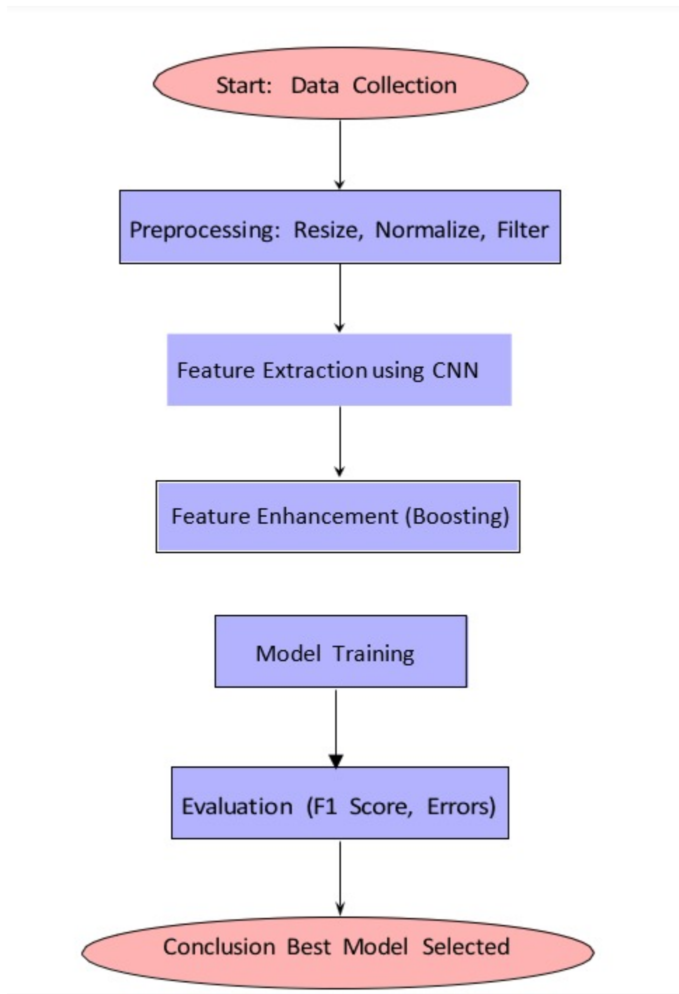

Fig. 2. Proposed work flow for lane detection

A. Data Collection

The dataset used in this study on lane detection is from open source benchmarks such as TuSimple and CULane. Such datasets are used largely in research on lane detection as they contain marked lane images in various environments (day, night, raining, and other difficult situations including occlusions). The input data includes video streams and/or pictures taken by on-board cameras. These images are used for lane processing based on CNN architectures.

B. Preprocessing

1) Image Resizing: The input images from TuSimple and CULane datasets are not uniform in size. To avoid such variances across the datasets and to save on computational costs, it is a common practice to resize all images to a standard image size (for instance, 224x224 for VGG and ResNet models). This resizing step helps in making all input image of the same size, which is a necessity for the operations involving convolution neural networks.

$$\text{I'} = \text{Resize(I, (h,w))}$$

2) Normalization: The purpose of normalization is to adjust the position of the image pixels in an appropriate range that neural networks can work with. In most cases the pixel values are normalized between the values of [0, 1], with the other option being that these values are standardized with a mean of 0 and standard deviation of 1. This avoids the situation where very large input values take control of the training and also assists in speeding up the convergence of the network.

$$I' = \frac{I - \mu}{\sigma}$$

3) Color Space Conversion: For the purpose of enhancing the prominence of lane demarcations, certain input images are modified from RGB (Red, Green, Blue) to other color spaces such as HSV (Hue, Saturation, Value) or even greyscale. This alteration restricts the distracting effects of colors irrelevant to the lane lines, which is most useful in images captured under different lighting states.

$$\text{Igray} = 0.299\text{R} + 0.587\text{G} + 0.114\text{B}$$

4) Gaussian Blur: Gaussian blur is used to lessen image noise and detail, which aids in the identification of continuous lane markings by reducing the minor variations in pixel intensity. This stage is particularly effective in eliminating noise of a high frequency that might otherwise disturb the model while detecting lanes.

$$I' = I * G(x, y, \sigma)$$

5) Canny Edge Detection: Canny edge detection is considered effective in marking the boundaries of lane markings. It functions through the detection of abrupt pixel intensity changes that indicate the edges of lanes. This approach is useful in minimizing unrelated details and enhances the lane edges thus helping the CNN models to concentrate on lane looking patterns.

$$E = \text{Canny}(I', T_{\text{low}}, T_{\text{high}})$$

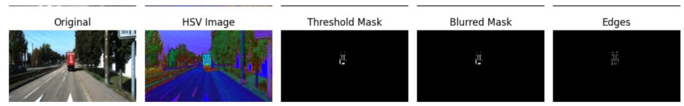

Fig. 3. Feature map

6) Thresholding: An operation is carried out to create a binary mask of the white lanes in the image by applying the function inRange of the OpenCV library, which masks the HSV image to obtain a picture of the white pixels.

7) Region of Interest (ROI): That is an optional step in which a polygonal mask is drawn to restrict the region of concern in the picture. This ensures that the detection is aimed at only the pertinent portions of the image.

$$\text{ROI}(x, y) = I[x_{\text{min}} : x_{\text{max}}, y_{\text{min}} : y_{\text{max}}]$$

C. ResNet Architectures

1) ResNet-101: ResNet-101 has a total of 101 layers which helps in better feature extraction due to the more significant depth. The model architecture employs residual learning to overcome the challenges associated with the vanishing gradients problem. The residual mapping can be written as follows:

$$y = F(x, \{W_i\}) + x$$

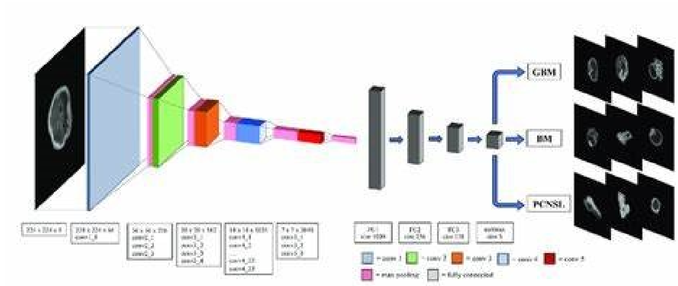

Fig. 4. Resnet-101 Architecture diagram

2) ResNet-50 : The sake of realization by using a deep neural net using 50-layer ResNet is working better for road detection. ResNet-50, the challenge of vanishing gradients is solved simply by introducing the idea of residual learning, which assists in the successful training and the extraction of rich and robust features of road lanes. Therefore, those who wish for a well-calibrated recognition model with satisfactory use of resources will get to benefit from the very well-balanced nature of ResNet-50 in terms of calculation time, accuracy level and channels used. It should be seen as level-headed in terms of efficiency but balanced in performance, best suited for real-time lane detection systems where speed is a necessity, as ResNet-50 gives the best trade-offs between computational power and accuracy. This ability helps ResNet-50 in lowering operational costs without additional performance loss, making it a very practical and competitive option among modern autonomous driving applications.

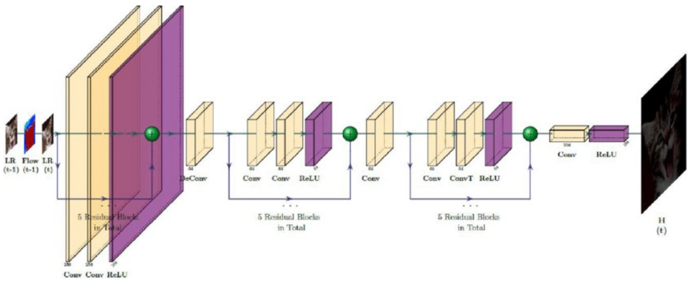

Fig. 5. Resnet-50 Architecture diagram

D. VGG Architectures

1) VGG-16: The VGG-16 comprises a total of 16 layers, each of which is a convolutional layer, followed by ReLU activations and pooling operations. This kind of structure of the network has a greater emphasis on the spatial features extraction rather than reducing the dimensionality of an image. Convolution operation's main equation is provided below:

$$y = f(x) = W * x + b$$

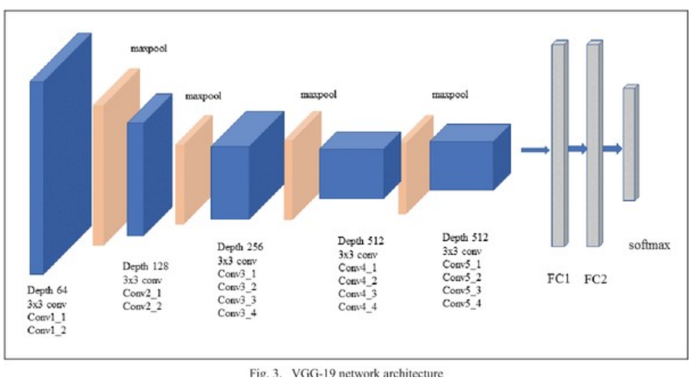

Fig. 6. VGG 16 Architecture

2) VGG-19: VGG-19 can be considered as an advanced variant of the base VGG-16 architecture as it comprises of 19 layers. It operates on the same principles but has additional layers for enhanced feature extraction. VGG-19 provides richer and more detailed feature maps, but this comes at an economically expensive cost similar to ResNet-101.

E. Boosting Techniques

Boosting is a method of ensemble learning that aims to improve the predictive performance of a model by combining a set of weak predictors. For instance, in this research that deals with lane detection, all boosting algorithms such as Gradient Boosting, XGBoost, CatBoost, and AdaBoost can be used along with the Deep Learning Models like ResNet and VGG for better classification and detection of lane markings. In this segment however, each of the above boosting strategies is reviewed in relations to their performance in lane detection systems.

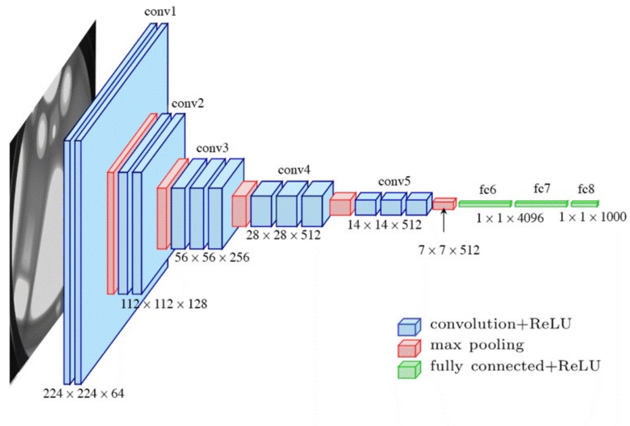

Fig. 7. VGG 19 Architecture

1) XGBoost: XGBoost (Extreme Gradient Boosting) is an ensemble algorithm based on decision trees which is effective for structed data processing. In the hybrid framework, XGBoost is employed over deep CNN feature representations (i.e.: from ResNet or VGG) as an additional step. Classification accuracy is reinforced by feeding features resulting from CNN layers into XGBoost.

$$\text{Obj} = \sum_{i=1}^{n} \left[l(y_i, \hat{y}i) + \sum k = 1^K \Omega(f_k) \right]$$

2) AdaBoost (Adaptive Boosting): AdaBoost is a straightforward, foundational-aspect boosting algorithm that modifies the weights of samples that have been misclassified, thereby compelling the model to pay closer attention to difficult Examples. With respect to lane detection, classification predictions can be enhanced by use of AdaBoost in difficult cases such as latitudinal obscurations of lanes, low light conditions, and old lane markings.

$$F(x) = \sum_{m=1}^{M} \alpha_m h_m(x)$$

3) Gradient Boosting: Gradient Boosting is a method that builds an ensemble of decision trees over several iterations. Here, each new tree fits within the excess error made by the previously trained trees in the ensemble. This is done using the approach of optimizing the objective function by employing the gradient descent technique. In gradient boosting, lane detection is performed to improve the classification of lane markings by dealing with difficult cases such as occluded or faded lanes.

$$F_m(x) = F_{m-1}(x) + \eta h_m(x)$$

4) CatBoost (Categorical Boosting): CatBoost is a boosting algorithm designed especially to include categorical features without having to do lots of preprocessing, for example, no need for one hot encoding. This is very relevant in lane detection which involves such features that are often categorical and discrete in values and types, for instance, in road conditions, weather types or types of lane markings. Therefore, by managing categorical data proficiently, CatBoost improves the feature extraction process in lane detection.

$$\text{Obj} = \sum_{i=1}^{n} \left[l(y_i, \hat{y}i) + \sum k = 1^K \Omega(f_k) \right]$$

IV. Result

In this lane detection project, we have thoroughly assessed the performance of a number of existing state-of-the-art models using a purpose-built lane detection dataset. This dataset poses various challenges such as straight lanes, bending lanes, crossways, and outdoor conditions. The models assessed include ResNet-50, ResNet-101, VGG16 and VGG19 all of which are incorporated with different boosting techniques XGBoost, AdaBoost, Gradient Boosting and CatBoost. From the results it was observed that the F1 score for ResNet-50 was the highest at 0.64, which proves its efficiency in lane marking detection in a wide range of conditions. This can be explained by the underlying deep architecture that the model uses, making it easy to extract profound features that deal with complex patterns of road lanes.

TABLE I
Performance Comparison of Model Configurations

Model Configuration	F1 Score	X Error	Z Error
ResNet-50	0.64	0.5131	0.5131
ResNet-101	0.3758	0.4562	0.3941
VGG16 + XGBoost	0.48	43	41
VGG16 + AdaBoost	0.4939	44	40
VGG16 + Gradient Boosting	0.4596	43	44
VGG16 + CatBoost	0.49	40	41
VGG19 + XGBoost	0.5082	0.4828	0.4915
VGG19 + AdaBoost	0.5126	0.4828	0.4874
VGG19 + Gradient Boosting	0.4969	0.50	0.50
VGG19 + CatBoost	0.4383	0.5343	0.5617

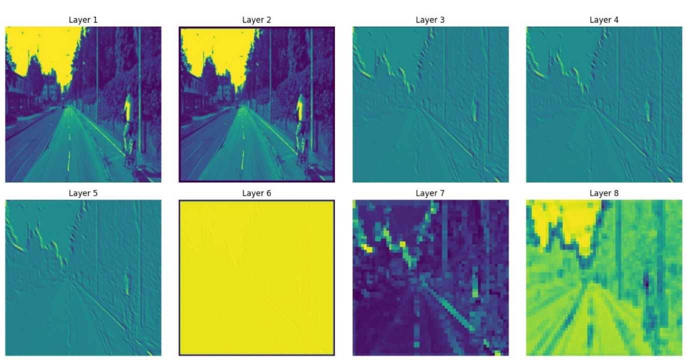

Fig. 8. Feature map

V. Evaluation of the Model

The last stage of the evaluation of our models designed for lane detection was their comprehensive evaluation by means of multiple metrics to comprehend their efficiency in various conditions. In this situation, the main attention was drawn to the F1 score which is calculated to measure the balance between precision and recall, whereas X and Z error rates were used to determine the accuracy and robustness of our models respectively.

As evident with the results obtained, the ResNet-50 network also earned a satisfactory F1 score of 0.64, in the clear presence of lane marking, suggesting an efficient performance in understanding lane marks even in the best case scenarios. This high score indicates that the ResNet-50 network is able to control false positive and negative rates very well and therefore reliable lane detection is made possible in simple cases.

On the other hand, the models had differences in accuracy when the lane markings were either partially or completely hidden. In these harsh conditions however ResNet-50 was still the best among all competitors but the F1 score dropped significantly. This finding points to the strength of the model, as it is able to learn and perform well even in cases that are different, for instance with different light, weather and visibility.We also decentralized further our analysis on the identified models to specific scenarios such as curved lanes, extreme weather, night time and interchanges, merges and splits. In scenarios of curved lanes, we observed that ResNet-50 performance did not decrease significantly indicating healthy adaptation to the changing lane geometry. In extreme conditions that hinder the clear view of the lanes such as heavy rain or fog, the model still performed very well showing its competence in adverse conditions. Also, in night time conditions that impair vision, the ResNet-50 had better detection performance which makes it an ideal candidate for practical systems that require operation at driving is common. Deep architecture and feature extraction are some of the main advantages that ResNet-50 encompass. However, there are also several drawbacks that must be considered. The prediction of the model on Z error, which accounts for the vertical shift, was lower than expected when relating to comparable models. Currently, the height loss is computed using the average height for the respective grid, leading to this shortcoming. This will be taken into consideration in the course of our activities in order to improve the model's performance on three-dimensional lane detection.

When evaluating the performance of our models, in addition to the parameter accuracy, we also searched for the parameter concerning the inference speed as we appreciated its significance in real-time applications. ResNet-50 kept relatively high accuracy whilst managing noticeably rapid inference speeds. EfficientNet family on the other hand, suited EfficientNet-b2 has shown beyond reasonable doubt even more speed without much compromise on the accuracy aspect. This allows EfficientNet-b2 to be used for applications such as real time lane detection because it is able to handle data at high speed without much loss of accuracy.

These evaluation results indicate that ResNet-50 is a good candidate solution for inclusion in lane detection systems of autonomous cars. Its impressive performance under diverse conditions means that it can be used in navigation systems with the confidence that it will improve the safety and efficiency of such systems in vehicles. Looking ahead towards the future, improvements such as attention mechanisms and speed of inference optimization will make ResNet-50 even more relevant in the context of moving cars.

VI. Conclusion

The author of this work has exhibited the effectiveness of using deep CNN architectures (ResNet and VGG) and coupling it with another approach which is a boosting algorithm (XGBoost) for lane detection. ResNet-50 delivers the highest precision, however ResNet-101 give a humble efficiency and f1 score. VGG models are also effective but take much time and hence cannot be used for real time applications. The robustness of the model, particularly for difficult lane detection situations, is also boosted by the inclusion of XGBoost.

References

[1] X. Pan, J. Shi, P. Luo, X. Wang, X. Tang, Spatial as deep: Spatial CNN for traffic scene understanding, Proc. AAAI Conf. Artif. Intell. 32 (2018).

[2] R. Liu, Z. Yuan, T. Liu, Z. Xiong, End-to-end lane shape prediction with transformers, in: 2021 IEEE Winter Conference on Applications of Computer Vision, WACV, 2021, pp. 3693–3701.

[3] Y.B. Özçelik, A. Altan, Overcoming nonlinear dynamics in diabetic retinopathy classification: A robust AI-based model with chaotic swarm intelligence optimization and recurrent long short-term memory, Fract. Fract. 7 (8) (2023).

[4] İ. Yağ, A. Altan, Artificial intelligence-based robust hybrid algorithm design and implementation for real-time detection of plant diseases in agricultural environments, Biology 11 (12) (2022).

[5] T. Zheng, H. Fang, Y. Zhang, W. Tang, Z. Yang, H. Liu, D. Cai, RESA: Recurrent feature-shift aggregator for lane detection, Proc. AAAI Conf. Artif. Intell. 35 (2021) 3547–3554.

[6] Z. Xie, Y. Lin, Z. Zhang, Y. Cao, S. Lin, H. Hu, Propagate yourself: Exploring pixel-level consistency for unsupervised visual representation learning, in: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 2021,pp. 16679–16688 .

[7] T. Chen, S. Kornblith, M. Norouzi, G. Hinton, A simple framework for contrastive learning of visual representations, in: H.D. III, A. Singh (Eds.), Proceedings of the 37th International Conference on Machine Learning, in: Proceedingsmof Machine Learning Research, vol. 119, PMLR, 2020, pp. 1597–1607,

[8] A. Bardes, J. Ponce, Y. LeCun, VICRegL: Self-supervised learning of local visual features, in: S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, A. Oh (Eds.), Advances in Neural Information Processing Systems, Vol. 35, Curran Associates, Inc., 2022, pp. 8799–8810.

[9] Y. Shi, N. Siddharth, P. Torr, A.R. Kosiorek, Adversarial masking for self supervised learning, in: K. Chaudhuri, S. Jegelka, L. Song, C. Szepesvari, G. Niu, S. Sabato (Eds.), Proceedings of the 39th International Conference on Machine Learning, in: Proceedings of Machine Learning Research, vol. 162, PMLR, 2022, pp. 20026–20040.

[10] L. Shen, H. Tao, Y. Ni, Y. Wang, V. Stojanovic, Improved YOLOv3 model with feature map cropping for multi-scale road object detection, Meas. Sci. Technol.34 (4) (2023) 045406,

[11] X. Song, Z. Peng, S. Song, V. Stojanovic, Anti-disturbance state estimation for PDT-switched RDNNs utilizing time-sampling and space-splitting measurements, Commun. Nonlinear Sci. Numer. Simul. 132 (2024) 107945, http:// dx.doi.org/10.1016/j.cnsns.2024.107945 Regression;Decision Tree;Random Forest;Gradient Boosting

[12] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional networks for biomedical image segmentation, in: N. Navab, J. Hornegger, W.M. Wells, A.F. Frangi(Eds.), Medical Image Computing and Computer-Assisted Intervention– MICCAI 2015, Springer International Publishing, Cham, 2015, pp. 234–241.