Identify the Driving Space for Vehicle Movement by Lane Line and Road Object Detection Using Deep Learning Technique

Jayamani Siddaiyan[*], Kumar Poonusamy

Department of Electronics and Communication Engineering K S. Rangasamy College of Technology, Tiruchengode

Corresponding Author Email: drbright.sa@gmail.com

ABSTRACT

Lane and object detection is the major concern of an autonomous vehicle or driver assistance to mobilize continuously without making any traffic congestions and accidents. In a complex traffic scene countries like India facing many challenges to enabling the intelligent transport system in end-to-end customer connectivity. In this work the major district road (MDR) type is considered to identify the driveable space for the host vehicle. The proposed novel work is the combination of lane lines and object detection by LaneNet with sliding window and YOLOv5. Prior to the detection method, for computational complexity pre-processing methods, ROI and bird eye top down views are carried out. The object bottom corner coordinate points and lane boundary coordinate points on the reference line is considered to calculate the space on both sides of a front object of host vehicle parental lane.
Finally, we used the real-time data and the most available CULane, BDD100K and TuSimple public dataset to perform simulation of a proposed work. LaneNet with sliding window for lane detection and pertained YOLOv5 model for an object detection and localization with an accuracy of 97% and 98% respectively. The simulation's outcomes demonstrate that the precision of the driving space identification results, 80% to 92% on various datasets.

1. INTRODUCTION

In today's fast-paced society, everyone pursues their own lifestyle and sticks with a schedule to fulfill expectations and comfort. A minor accident or violation of the traffic laws on a highway or city street causes traffic and slows down the day's progress. 1.3 million Individuals pass away every year in a roadway accident report by the World Health Organization. The disciplined highways have made significant progress in the identification of lane and road borders for an autonomous vehicle, and driver support systems. The operation of an autonomous intelligent vehicle has not yet progressed to the full degree in the real-world scenario depicted in the figure 1.

Figure 1. Sample scene of major district road environment

As seen in the figure. 1 situation operating an autonomous, intelligent vehicle still presents difficulties, as a result of undesired roadside occupation, poor maintenance, and lane-avoiding behaviour. Connecting residential areas to roads with

an autonomous vehicle movement is challenging in this scenario. The Vision system is the one which helps to detect any type of an object and track them in front of the host vehicle. Computer vision is opted for such detection and tracking of an object on the road lane. The review of the literature provides information on how to design and develop an existing system such that it functions properly with appropriate road environment upkeep.

The overview of well-structured lanes recognition and results reported in [1, 2]. Such a system is not assured if there are less lane care and poor lane conduct. In such situation the features of an image boosted methods carried in [3] and [4]. On the other sides of the road, where there is less visibility of the border and light variance due to shadow cast by trees and buildings, difficult to distinguish [5]. The improvements obtained in various algorithms, based on feature, model, and machine learning to detect the lane markings [6-9]. Each of these have their own benefits and drawbacks depending on the situation. Therefore, the single algorithm cannot produce the optimal outcomes in every setting. This situation is addressed together for better results. Typically, an SVM-based technique is used in conjunction with supervised or semi-supervised online learning to identify roads [10]. For instance, [11] effective method of using self-supervised online learning with SVM for road detection. Both the correlation feature set and the raw feature set are evaluated and make use of the boosting, SVM, and random forest classifiers, recently neural networks were also used to identify roads [12].

Despite having great performance accuracy, it is dependent on intricate calculations and a substantial amount of training

data. The aforementioned examples concentrated on the organised and disciplined highways, which is not typical for all connected roads. On gravel roads, Lane detection is challenging because of the road's texture, lighting variations, and fuzzy borders. The overall performance of detection also degrades due to weather and ruts. The state of features of dirt roads as described in [13, 14] and correct, but unsubstantiated information obtained accurately in [15]. Identify the discovery of driveable space for a host vehicle's movement in another section of this work. In this situation, one must find the objects that suddenly emerged in front of the host vehicle. The method which located and detects the object is called object detection by making bounding boxes with labelling.

Various methods for the improvement of environmental perception are provided by recent deep learning algorithms[16] . Numerous sensors are crucial to an autonomous vehicle's ability to gather data about its surroundings. The most frequently used sensors, such as LiDAR, GNSS, GPS, and many others[17], are efficient and perform the required work. Computer vision and deep learning methods have become more popular in every area of Intelligent Transportation Systems (ITS) in various ways, including reserve vehicle finding, traffic control and sign recognitions and number plate recognition. By detecting computational behaviour and offering encouraging outcomes in numerous disciplines, such as picture arrangement [18] segmentation[19], and identification of moving objects and tracking [17], Totalling objects, detecting overtaking vehicles, classifying objects[20], and detecting lane changes [21].

In general, two types of deep learning algorithms such as Regression-based models and region-based models for object detection are both employed. The object had previously been located in two steps, which took more processing time. On the other hand, simple regression uses a single shot of the full image to provide bounding boxes and class probabilities [22]. In comparison to the double-stage object detector, this model performs faster. It has been suggested that an improved version of You Only Look Only (YOLO) be developed to detect electrical components[23], licence plates [24], and many other things. YOLO-UA for traffic flow monitoring[25] and YOLO-CA for accident detection are two examples. [26] To find the crop being harvested from the palm tree plantation, the YOLO-P model is employed. From the above One-shot trained model detect the various objects present in an image or video frame which helps identifying the object utilized area in an image. The overall existing methods as discussed in the earlier methods gives the clearest view of an objects and lane line detection disciplined highways road scenarios. As these detection methods not full fill the needs of driver assistance at traffic congested area as well as rural and urban road environment. By adding the essence of an object space occupied on the road in forward-facing of the host vehicle will help the driver avoid traffic congestion at crucial stages. The remaining part of the paper organized with support for the work to obtain the driveable space in the single road urban and city road environment.

2. SUPPORTING METHODS

Vision system is one of the major component in an autonomous vehicle/driver assistance system to gather data of in front of the host vehicle. For some reason, few photographs collected in a real-time scenario are distorted. The tangential distortion, which occurs when the images are not maintained parallel to the image plane with the camera lens, in other case the causes of warped edges belonging to the category of radial distortion. The actual driving environment reflects these distortions. The following are the mathematical model definition of radial and tangential distortion correction.

Distortion Parameters
$$D = (k_1, k_2, p_1, p_2, k_3) \quad (1)$$
Correction on radial manner
$$x_c = x_d(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \quad (2)$$
$$y_c = y_d(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \quad (3)$$
Correction on tangential manner
$$x_c = x_d + [(2p_1 x_d y_d + p_2(r^2 + 2x_d^2)] \quad (4)$$
$$y_c = y_d + [p_1(r^2 + 2x_d^2) + 2p_2 x_d y_d] \quad (5)$$
Where $r^2 = x^2 + y^2 (x_c, y_c)$ is the position's location in the image pixel coordinate at point d.

Suffice d and c – represents distorted and corrected coordinate respectively

r – represents the distance between the centre of the image to correct coordinate point.

k,p – represent radial and tangential distortion parameters respectively.

Before installing the camera, these types of known distortion are addressed. One of the greatest ways to examine camera distortion and calibration is with a checkerboard. The regimented regular forms made of a checkerboard pattern of black and white aid in adjusting the camera to detect scenes with distorted images. The photos below, as seen in figure 2,2.1 show how distorted original photographs are transformed into undistorted images.
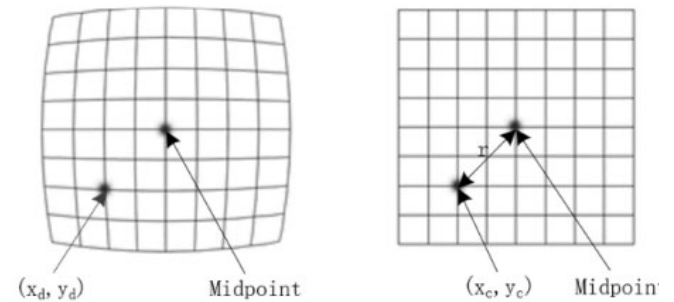

(x_d, y_d) Midpoint (x_c, y_c) Midpoin

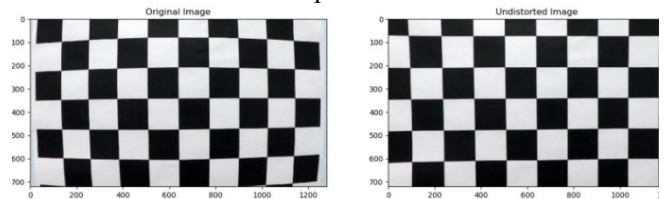**Figure 2**. Camera distortion correlation parameters.

Figure 2.1 Chess board Image for camera calibration

2.1 Region of Interest (ROI)

A single image contains more information that is somehow unrelated to a specified work. Performance is delayed since the process takes into account all information about the photos. Finding the informative portion of an image is therefore crucial for accelerating the process. The part of the image consider for our work is devoted to important information and the remaining space considered as unrelated. The ROI (region of interest) aids in locating the image's such a manner in order to choose informative area, the figure.3 visualises the chosen zone from the original image.

ROI is the region enclosed by a circle, rectangle, or other shape surrounding an interesting area. In this study, ROI is calculated using a rectangular pair and a trapezoid mask, which may be stated as follows.

C1. Using a trapezoid mask, create a rectangle couple

C2. Merging an image as well as a mask

C3. Choosing only the region that interacts with the mask and the picture

The inside the green box is the drivers most attend able area, which are considered as important information as related to road scene.

Figure 3. Visualization of the original scene with ROI

2.2 Inverse Perspective Mapping (IPM)

In real-world scenario, the detected object appears in dynamically varying, while it is larger when it near to the camera and smaller if it keep the distance from the camera. The disappearing point is harmful to active lane borders where two parallel lines meet at the far range of vision. IPM (Inverse perspective mapping) restores the link between the parallel line and the world coordinate seen from above [27, 28]. IPM is then used to transform the forward camera's vision into a bird's eye perspective. The road apparent point $(X_w\ Y_w\ Z_w)$ that projects to construct the homograph relations of the four corners of the picture plane (u, v) is necessary to produce a top-down view. Figure 4 pictures the projection principle and the equation as shown in (6).
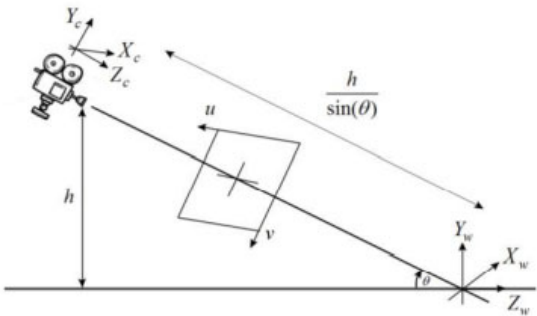

Figure 4. Invers Perspective Projection View

$$(u, v, 1)^T = K\ T\ R\ (X_W, Y_W, Z_W, 1)^T \qquad (6)$$

R – Rotation matrix

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad (7)$$

T – Translation Matrix

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \frac{-h}{\sin\theta} \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad (8)$$

Considering the above equation 7, 8 with K (camera matrix)

and road plane $(Y_w = 0)$ the equation becomes

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} p_{11} & p_{13} & p_{14} \\ p_{21} & p_{23} & p_{24} \\ p_{31} & p_{33} & p_{34} \end{pmatrix} \begin{pmatrix} X_w \\ Z_w \\ 1 \end{pmatrix} \qquad (9)$$

Using this equation, it is possible to map each pixel on the image plane to a top down view as shown the result in figure 5, which helps to detect the lane line.

Figure 5. Visualization of IPM

3. PROPOSED WORK

3.1 Lane Detection

The driverless vehicle fared best on highways, where lane markings are brighter and more pronounced. However, in practise, the connectivity of the roads between urban and rural areas is not regularly maintained. In this case, the autonomous vehicle operation is not successfully detecting the lane line. Finding the lane line with the use of a lane line predictor and lane localization is one of the simplest ways to locate the road area. In order to successfully detect lanes, it is typically necessary to first employ pertinent methods to extract the pixel features of the lane line. Next, the proper pixel fitting algorithm is then required to complete the process.
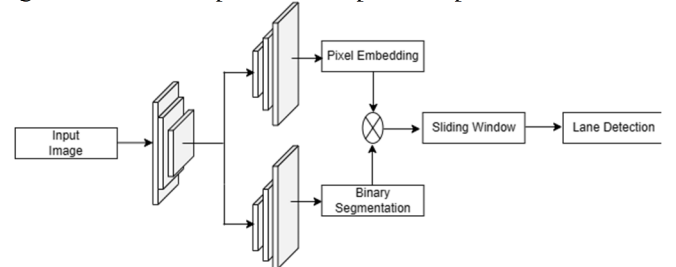

Figure 6. LaneNet Architecture

The network, termed LaneNet and depicted [29] in figure 6, combines the benefits of binary lane segmentation and identifies the pixels that are a part of lanes and those that are not. In order to capture the coordinates of pixels, the sliding window is scanned from the bottom up. The centre of the next sliding window is determined by the mean value of the pixels when the number of pixels exceeds a predetermined threshold. After threshold filtering, the hybrid incompatible operator and graph, along with the sliding window fit the lane. The pixels from the preceding frames are used to fit the curve, while the pixels from the recorded position information are used in a quadratic polynomial[30] . LaneNet model is used in this case, segmentation to detect the lane in any situation, such as when it is blocked by objects or when there are no visible lane line marks. The LaneNet outcome is shown in the figure. 7

Figure 7. Lane detection – Proposed LaneNet Model

3.2 Object Detection

The very first and most essential function of the visual system is to locate and categorise an object within an entire image. The visualisation of this detection uses a bounding box with a label and probability score. Tracking an item in vision applications is another difficulty. The most common traditional approach incorporates modules for extracting features, region selection, and classification. The basic feature extraction models for road object recognition are Haar [31] and HOG [32] discussed. The characteristics are manually retrieved using this traditional method, which takes more time. Recent improvements in vision sensors have increased data collection over a time as well as the rapid growth of hardware, particularly parallel computing, which helps deep learning outperform older algorithms. The object detection algorithm used in the deep learning approach contains two stages in addition to a single stage [33] to ensure that real-time performance requirements for timeliness and accuracy are met. Models with two stages of detection move more slowly than those with one stage in the prior manner. Double stage object detection models with high complexity and calculation time included in RFCN [34] and Mask R-CNN [35]. When compared to two-stage detection models, one stage networks like YOLO and SSD offer faster calculation times. Redmon et al. presented the YOLO algorithm [36] is significantly quicer than existing approaches, then YOLOv3 in 2018, which drastically improves finding speed and precision. The YOLO sequence of algorithms will have evolved to YOLOv5 by 2020. Additionally, the sequences of models can be identified based on the network's breadth and feature map's size. More specifically, all previous four models share the identical input, backbone, neck, and prediction components of the network.

The one stage You only Look once architecture has a convolutional [37] layer and an input picture are the two main components of the regression-based whole YOLO structure. In a single step, the convolution layer processed a picture and returned the spot and grouping the detected objects. The input size for three-layer YOLO architecture is 640x640x3. The Conv2D + Batch Normal + Leaky subcomponents make up the Basic Convolution Module (CBL). 32 convolution kernels are used in RELU to slice the input images, which produces a final product with the dimensions 320x320x32. The feature extraction, gradient information elimination, and optimization processes are handled by the CSP bottleneck module. Different feature scales are acquired via the Spatial Pyramid Pooling (SPP) Figure 8 shows the precise YOLOv5 architecture.
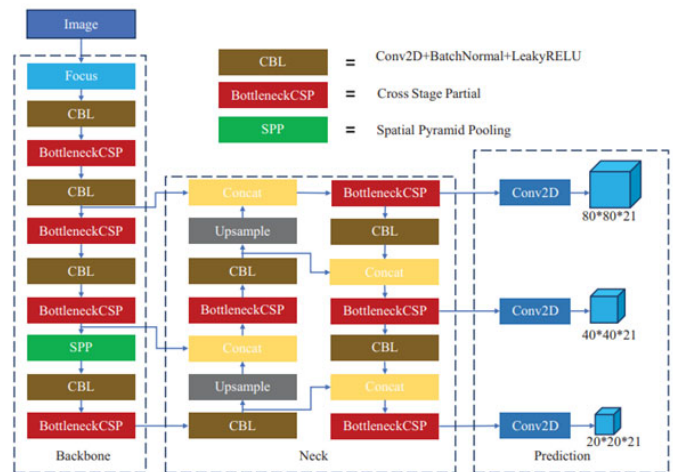

Figure 8. YOLOv5 Architecture

Algorithmic tasks include things like classifying and identifying one or more objects in a picture. Before creating our own, we used a YOLOv5 model that had already been trained using 80 classes from the COCO dataset. Vehicles, traffic lights, pedestrians, motorbikes, bicycles, and even animals can act as roadblocks. The following are the process's specifics and the result specifies in figure 9:

Step 1. The trained YOLOv5 model has been used. The input image to the neural network needs to have a particular shape, like a blob. Using the blob and the image function, the frame is translated into a blob for a neural network in order to be understood from a video sequence. The process also changes the pixel values between 0 and 1 and resizes an image to the required size (640, 640). The YOLO network is then used for blob forwarding. The YOLOv5 method generates bounding boxes to aid in detection prediction.

Step 2. A vector is used to represent the class, five network components (the centre x, centre y, breadth, height, and sureness box), and each output bounding box. The output of the YOLOv5 algorithm consists of many boxes. Though, the majority of the boxes are superfluous. They must be screened and removed. In the first step, a box will be removed if there is a low likelihood that an object will be detected. Only those boxes are maintained that have probabilities greater than the specified threshold. The remaining boxes will do non max suppression. Fewer boxes will overlap as a result.

Step 3. Combining lane detection techniques. Hazy lanes, losing track, and interference from immovable objects on the road are the main causes of inaccurate lane identification. Therefore, using the obstacle detection result in the lane detecting approach would be advantageous.
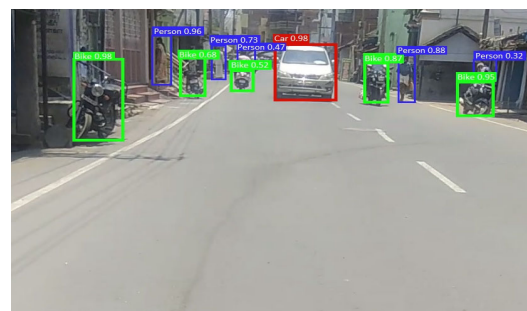

Figure 9. Object detection – YOLOv5

3.3 Driving Space localization

The road image is a complex scene with a variety of objects in and around the host vehicle. It is quite difficult for a novice driver of a host vehicle to continue driving without colliding when they assume driveable space. Connecting roads between urban and rural areas are bidirectional, and some sections of the road are filled and crowded. Such an environment is focused on the effort in order to provide the driveable space for vehicles continuous movement. For obtaining the non-occupancy space for driving, a reference line is fixed and the coordinate points of a lane border and front object is considered. The figure's 10 represents graphically to estimate the coordinate of an object.
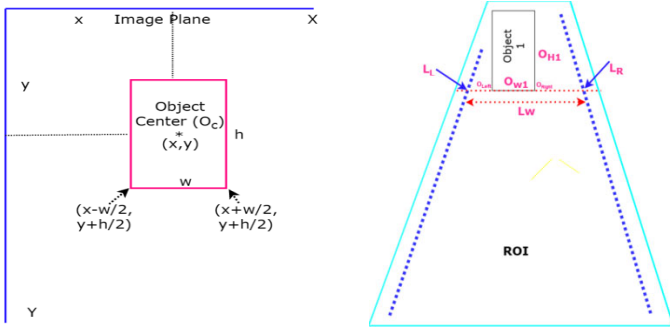

Figure 10. Driveable space detection and coordinate Findings

The estimation of a single object coordinate point within the ROI region is shown in figure 10. Each object's centre point (x, y), object width (w), and object height (h) is provided by the YOLOv5. The schematic in figure.11 provides a perspective on how to estimate the driveable area for an eco-vehicle. The endpoints of the boundary lane are selected on this reference line, noted as lane left edge (L_L) and lane right edge (L_R), and the width of the lane is determined (Lw).

$$D_{path} = L_w - O_w \qquad (10)$$
$$D_{Rpath} = L_R - O_{right} \qquad (11)$$
$$D_{Lpath} = O_{left} - L_L \qquad (12)$$

$L_w - Lane\ Width$
$O_w - Object\ Width$
L_L
$- Left\ Lane\ Boundary\ and\ Reference\ line\ meeting\ point$
L_R
$- Right\ Lane\ Boundary\ and\ Reference\ line\ meeting\ point$

According to the equation (11), (12), where the difference is greater gap as chosen as a positive driveable area for the host vehicle to approach in advance. In case the front object close to the host vehicle or crossed the reference line the 'Y' coordinate point of an object has to be consider to decide the driving space by drawing a new reference line to collect the corresponding L_L L_R points

4. EXPERIMENTAL RESULTS AND ALALYSIS

4.1 Experimental Setting

a. Data set
BDD100K: The 100K video frames in the BDD100K dataset are divided into 70000 for training, 10000 for validation, and 20000 for testing. We employ the original validation set for testing but leave the training set unmodified because the testing partition's ground-truth labels are not accessible to the general public..

TuSimple: The TuSimple dataset, which is comprised of 6K road photos taken from US highways, is now the most often used for lane recognition. The image has a resolution of 1280 x 720. 3K records for training, 358 records for justification, and 2K records for testing make up the dataset. The photographs in the dataset were captured on the highway in primarily favourable weather and lighting conditions.

CULane: The CULane dataset 20 times more than the size of the TuSimple dataset contained a total of 133,235 frames. The training set is 88K the validation set is 9K and the test set is 34K. There are nine distinct scenarios included in it, including regular, throng, curve, dazzle night, night, no line, and arrow in the city.

TRLane: The Mixture of rural road in china with TuSimple dataset known as TRLane[38]. Only between lanes and backdrop are distinguished in its annotation; instances of lanes are not. This dataset includes 1.1K rural road image arrangements, each of which consists of 20 successive frames, in addition to a portion of the TuSimple dataset.

b. Evaluation Metrics
The final segmentation result for the lane task requires an estimation index to estimate the effectiveness of the test. We utilise the F-Measure assessment index to assess the model's capacity to forecast the lane line since it can take into account both the precision and recall measurements from the experiment. Positive samples, identified by the letters P and background, are assumed to exist in each pixel of the lane line. Using the letter N, a pixel is a undesirable sample. If the expected picture output and individually pixel of the actual label are equal, the pixel prediction is said to be accurate. Using the F-Measure index, the formulas for evaluating picture prediction outcomes are shown in (13) and (14)

$$Precision = \frac{True\ Positive}{True\ Positive + Faulse\ Positive} \qquad (13)$$

$$Recall = \frac{True\ Positive}{True\ Positive + Faulse\ Negative} \qquad (14)$$

The proportion of the number of accurate forecasts to all predictions serves as a measure of precision, which represents the model's accuracy. Recall, a measure of coverage that compares the number of goals found in the data to the whole amount of goals that really exists, represents the model's recall rate. TP stands for the proportion of correctly recognised positive instances, FP for instances that were detected as positive cases, and FN for instances that were detected as negative cases.

F1 Scores: F scores are a statistician's tool for evaluating a binary model's precision. Both recall and precision are considered simultaneously. The F1 score, which ranges from 0 to 1, can be understood of as a periodic mean of precision and recall. It is defined as

$$F_1 = 2 \times \frac{precision \times recall}{precision + recall} \qquad (15)$$

The average number of valid points per image serves as the basis for calculating accuracy:

$$Accuracy = \sum_{im} \frac{C_{im}}{S_{im}} \qquad (16)$$

C_{im} the number of true points and S_{im} the number of ground-truth points. When there is less than a predetermined amount of difference among the projected point and the ground truth, the point is considered right.

5. RESULTS AND DISCUSSION

The recommended method is applied on the Google cloud stage with an OpenCV library. Real-time data is collected from a rural and urban MDR environment in and around the Salem region, Tamilnadu India with varying traffic conditions and congested roads. The combined proposed method of driving space identification on real time data results as shown in figure 12. Table 1 and 2 gives the comparative results with the planned method of lane line and object finding separately.

The model is validated through the various datasets and their results are represented in figure 13. From the test result the driveable space is identified based on the reference line and bounding box. In various scenarios, the space identification as exposed in figure 13.b and c the model identifying the maximum gap on any of the sides of the front object detected. In figure 13 a model did not concluded the maximum space on either side of the front object so it does not identify the driving space.

Figure 13.d one of the vehicles crossed the reference line, such condition the model compares the L_L L_R coordinate points on reference line with the closest object coordinate point and reassign the driving space.

Table 1. Comparison of proposed model with other Lane detection Methods

Method	Algorithm	Average Accuracy (%)	Recall	Precision	F1
Traditional Method [39]	Hough Transform	95.70	0.9194	0.9409	0.9302
Traditional Method [40]	RANSAC + HSV	86.21	0.9028	0.9213	0.9021
Deep Learning [41]	FastDrae ResNet	95.00	0.9525	0.9622	0.9537
Deep Learning (Our Proposal)	LaneNet + Sliding Window	97.03	0.9679	0.9767	0.9723

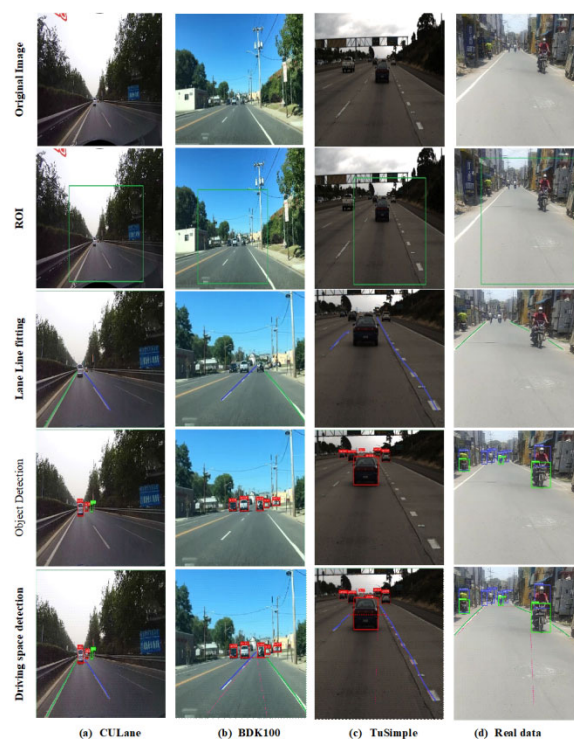**Figure 11.** Result of Driveable space identification in MDR type environment

Figure 12. Model Performance of various data sets with real time data

Table 2. Comparison of different Object detection methods

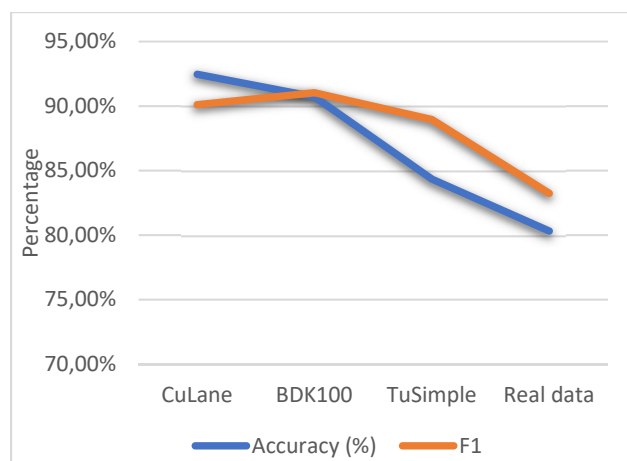| Model	Average Accuracy (%)	Recall
HOG + SVM [42]	57.0	0.4892
Horizontal filter + Otus [39]	75.00	0.6833
Faster R-CNN [43]	81.60	0.8138
YOLOv5x [44]	98.57	0.9812

Figure 13. Proposed Model Accuracy and F1 score for various Data sets

6. CONCLUSION

This paper designed for identifying the driveable space for the host vehicles, continuous movement and collision avoidance. The novel method utilizes the LaneNet with sliding window for lane detection and YOLOv5 pertained model for

object detection and localization with an accuracy of 97% and 98% respectively. The combined proposed method examined on TuSimple, CUlane, BDD100K and real-time data, the model produced the promising output with an accuracy of 82 % to 92% on various dataset scenes and these variation depends on the data quality. In future work the host vehicle size will be considered to identify the suitable driveable space on either side of the front vehicle which assist the driver to take the decision on right path to move forward.

REFERENCES

[1] McCall, J.C. and M.M. Trivedi, Video-based lane estimation and tracking for driver assistance: survey, system, and evaluation. IEEE transactions on intelligent transportation systems, 2006. 7(1): p. 20-37.

[2] Zheng, T., et al. CLRNet: Cross Layer Refinement Network for Lane Detection. in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022.

[3] Narayan, A., et al., A dynamic colour perception system for autonomous robot navigation on unmarked roads. Neurocomputing, 2018. 275: p. 2251-2263.

[4] Yang, F., et al., Feature pyramid and hierarchical boosting network for pavement crack detection. IEEE Transactions on Intelligent Transportation Systems, 2019. 21(4): p. 1525-1535.

[5] Zürn, J., W. Burgard, and A. Valada, Self-supervised visual terrain classification from unsupervised acoustic feature learning. IEEE Transactions on Robotics, 2020. 37(2): p. 466-481.

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

[7] Tian, Y., et al., Lane marking detection via deep convolutional neural network. Neurocomputing, 2018. 280: p. 46-55.

[8] Yu, Y., M. Xu, and J. Gu, Vision-based traffic accident detection using sparse spatio-temporal features and weighted extreme learning machine. IET Intelligent Transport Systems, 2019. 13(9): p. 1417-1428.

[9] Qiu, Z., J. Zhao, and S. Sun, MFIALane: Multiscale Feature Information Aggregator Network for Lane Detection. IEEE Transactions on Intelligent Transportation Systems, 2022. 23(12): p. 24263-24275.

[10] Huang, D.-Y., et al., Vehicle detection and inter-vehicle distance estimation using single-lens video camera on urban/suburb roads. Journal of Visual Communication and Image Representation, 2017. 46: p. 250-259.

[11] Yuan, Y., Z. Jiang, and Q. Wang, Video-based road detection via online structural learning. Neurocomputing, 2015. 168: p. 336-347.

[12] Conrad, P. and M. Foedisch. Performance evaluation of color based road detection using neural nets and support vector machines. in 32nd Applied Imagery Pattern Recognition Workshop, 2003. Proceedings. 2003. IEEE.

[13] Bertozzi, M., et al., Artificial vision in road vehicles. Proceedings of the IEEE, 2002. 90(7): p. 1258-1271.

[14] Shang, E., et al., Robust unstructured road detection: the importance of contextual information. International Journal of Advanced Robotic Systems, 2013. 10(3): p. 179.

[15] Yan, C., et al., A highly parallel framework for HEVC coding unit partitioning tree decision on many-core processors. IEEE Signal Processing Letters, 2014. 21(5): p. 573-576.

[16] Ahangar, M.N., et al., A survey of autonomous vehicles: Enabling communication technologies and challenges. Sensors, 2021. 21(3): p. 706.

[17] Huang, Y. and Y. Chen. Survey of state-of-art autonomous driving technologies with deep learning. in 2020 IEEE 20th International Conference on Software Quality, Reliability and Security Companion (QRS-C). 2020. IEEE.

[18] Garcia-Garcia, A., et al., A survey on deep learning techniques for image and video semantic segmentation. Applied Soft Computing, 2018. 70: p. 41-65.

[19] Minaee, S., et al., Image segmentation using deep learning: A survey. IEEE transactions on pattern analysis and machine intelligence, 2021.

[20] Haris, M. and J. Hou, Obstacle detection and safely navigate the autonomous vehicle from unexpected obstacles on the driving lane. Sensors, 2020. 20(17): p. 4719.

[21] Haris, M. and A. Glowacz, Lane line detection based on object feature distillation. Electronics, 2021. 10(9): p. 1102.

[22] Le, T.-T. and C.-Y. Lin, Deep learning for noninvasive classification of clustered horticultural crops–A case for banana fruit tiers. Postharvest Biology and Technology, 2019. 156: p. 110922.

[23] Huang, R., et al., A rapid recognition method for electronic components based on the improved YOLO-V3 network. Electronics, 2019. 8(8): p. 825.

[24] Min, W., et al., New approach to vehicle license plate location based on new model YOLO-L and plate pre-identification. IET Image Processing, 2019. 13(7): p. 1041-1049.

[25] Cao, C.-Y., et al., Investigation of a promoted you only look once algorithm and its application in traffic flow monitoring. Applied Sciences, 2019. 9(17): p. 3619.

[26] Junos, M.H., et al., An optimized YOLO-based object detection model for crop harvesting system. IET Image Processing, 2021. 15(9): p. 2112-2125.

[27] Huang, Y., et al., Lane detection based on inverse perspective transformation and Kalman filter. KSII Transactions on Internet and Information Systems (TIIS), 2018. 12(2): p. 643-661.

[28] Bonin-Font, F., et al., Concurrent visual navigation and localisation using inverse perspective transformation. Electronics letters, 2012. 48(5): p. 264-266.

[29] Neven, D., et al. Towards end-to-end lane detection: an instance segmentation approach. in 2018 IEEE intelligent vehicles symposium (IV). 2018. IEEE.

[30] Ma, N., et al., An all-weather lane detection system based on simulation interaction platform. IEEE Access, 2018. 8: p. 46121-46130.

[31] Park, K.-Y. and S.-Y. Hwang, An improved Haar-like feature for efficient object detection. Pattern Recognition Letters, 2014. 42: p. 148-153.

[32] Arróspide, J., L. Salgado, and M. Camplani, Image-based on-road vehicle detection using cost-effective histograms of oriented gradients. Journal of Visual

Communication and Image Representation, 2013. 24(7): p. 1182-1190.

[33] Aziz, L., et al., Exploring deep learning-based architecture, strategies, applications and current trends in generic object detection: A comprehensive review. IEEE Access, 2020. 8: p. 170461-170495.

[34] Li, J., J. Qian, and Y. Zheng, Ensemble R-FCN for Object Detection, in Advances in Computer Science and Ubiquitous Computing. 2017, Springer. p. 400-406.

[35] Xu, B., et al., Automated cattle counting using Mask R-CNN in quadcopter vision system. Computers and Electronics in Agriculture, 2020. 171: p. 105300.

[36] Redmon, J., et al. You only look once: Unified, real-time object detection. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[37] Zhou, F., H. Zhao, and Z. Nie. Safety helmet detection based on YOLOv5. in 2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA). 2021. IEEE.

[38] Zou, Q., et al., Robust lane detection from continuous driving scenes using deep neural networks. IEEE transactions on vehicular technology, 2019. 69(1): p. 41-54.

[39] Zheng, F., et al., Improved lane line detection algorithm based on Hough transform. Pattern Recognition and Image Analysis, 2018. 28(2): p. 254-260.

[40] Kim, K.B. and D.H. Song, Real time road lane detection with RANSAC and HSV Color transformation. Journal of information and communication convergence engineering, 2017. 15(3): p. 187-192.

[41] Philion, J. Fastdraw: Addressing the long tail of lane detection by adapting a sequential prediction network. in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019.

[42] Kaplan, Ö. and E. Saykol. Comparison of Support Vector Machines and Deep Learning for Vehicle Detection. in RTA-CSIT. 2018.

[43] Lin, C.-T., et al. Fast vehicle detector for autonomous driving. in Proceedings of the IEEE International Conference on Computer Vision Workshops. 2017.

[44] Dong, X., S. Yan, and C. Duan, A lightweight vehicles detection network model based on YOLOv5. Engineering Applications of Artificial Intelligence, 2022. 113: p. 104914.