International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

An Intelligent Lane and Obstacle Detection using YOLO algorithm

Shalu[1], Sonia Rathee*[2], Amita Yadav[3], Pooja Kherwa*[4], Rashmi Gandhi[5]

Submitted: 17/08/2023 **Revised**: 08/10/2023 **Accepted**: 22/10/2023

Abstract: It is highly challenging to detect lanes quickly and accurately due to a variety of complex noise, so the main aim goal is to develop a collection of image processing techniques and it produce results quickly and precisely under less-than-ideal circumstances. This paper suggests an intelligent lane recognition technique that utilizes a collection of distinct photos and applies the results to a video stream. The Hough transform is chosen as the most effective beeline detection technique, and the Canny algorithm is chosen as the edge detection technique. The ROI is defined to decrease noise for accurate rise and to increase processing speed to satisfy the real-time need.

For detecting the obstacle, to provide fast implementation and smooth real-time update of the obstacles nearby we are implementing the YOLO algorithm.

Keywords: *Intelligent Lane detection, canny edge detector, driver support system, obstacle detection, YOLO algorithm.*

1. Introduction

The safety of everybody on this globe is a fear. humans expect to arrive properly after they tour from one region to another. barring any unforeseen occasions that would occur even as visiting, which includes visitors' injuries. through utilizing better riding help, we can prevent traffic injuries. In most nations with heavy visitors, automobile injuries remain the major reason of accidental dying and injuries. maximum of the transportation injuries arises at the country's highways. A device that can alert the motorist to the threat consequently has the capacity to save a significant wide variety of lives.

Humans are making an investment a variety of cash within the advancement of safe riding practices so that it will growth protection and decrease traffic injuries. Technology forces human beings to suppose more severely about the way to boom safety and shop lives. The safety features that motor provide, which include seat belts, air luggage, and robust frame systems, provide passive safety and may lessen the results of an accident.

The complex and difficult tasks of future road vehicles, such as lane detection[2][3], obstacle detection[1] (cars, pedestrians, animals, etc.), and moving object detection, are essential for achieving the desired level of road safety. Moving object detection is a crucial part of collision avoidance in driving assistance systems.

[1] *Maharaja Surajmal Institute of Technology, New Delhi*
[2] *Maharaja Surajmal Institute of Technology, New Delhi*
[3] *Maharaja Surajmal Institute of Technology, New Delhi*
[4] *Maharaja Surajmal Institute of Technology, New Delhi*
[5] *Amity University , Noida*
** Corresponding Author Email: Sonia.rathee@msit.in*
** Corresponding Author Email: poojakherwa@msit.in*

One method of doing this is by using lane and obstacle detection systems, which function to identify the lane borders on the road and then alert the driver if they change to an incorrect lane or if an obstruction is in their path.

2. Related Work

A. Lane Detection

In the last few centuries, quite many researches have been made and published on lane detection and prediction topic. Some of the techniques that are utilized for detection of lanes are discussed below.

1.Hough Transform

Even in the presence of noise and occlusion ,straight lines in photos can be identified by applying the Hough Transform[9]. By computing different equations for each potential line passing through an image location, it may determine the establishing lines in a picture. An object class can be built from edge vertices. by selecting pixels from an image object set.

An image must first be turned into a binary image by applying a threshold in order to detect lines in it. The dataset is then expanded with the necessary instances. Hough space serves as the foundation of the Hough Transform. A line at angle T and distance d from the origin is matched to each point (d, T) in a Hough Space. The position of a vertex in respect to a straight line is given by using the Hough . Take into account every line that traverses that point at a specified discrete set of angles. Using an array known as accumulator, lines are found using the Hough transform. The number of unidentified Hough transform parameters is the same as the dimension of the accumulator. In the beginning, lines are created that can pass through each spot. The vote for those (d,T) parameters is boosted when a line meets with additional lines that originate from different places. (d,T) parameters pair is determined with the maximum value and then is finally selected as the dominant way visible based on the positions that constitute up this line, on the visual planes.

1. Bilateral Filter

It is an uncomplicated, non-iterative method [12] of smoothing the image while keeping the edges sharp. The two pixels should be close to one another in order for the bilateral filter operation to work. The filter's separates a picture in big- and small- elements, such as visualization and texture.

In every kind, the pattern has been switched out for the weighted mean of its friends. 2 forces are expressed via those weights: the similarity between the encircling and internal samples, which ends in the high scoring of examples and proximity between neighbour and centre pattern, such that nearer samples are given a higher weight.

3. Lane Boundaries Projective Model

The limits of the lanes are shown in this model [11] as both sharp circular curves and straight lines. The slope direction feature, the path model, the gate likelihood function, and the gate prior knowledge can all be used to derive the lane posterior probability. The particle swarm optimization approach can then be used to identify the lane with the highest posterior probability. Then, using the lane model, the lane boundary is identified and the lane geometric structure is properly estimated.

4. Edge Detection

The foundation of this approach [10] is the notion of locating regions in an image when the brightness of the image fluctuates noticeably. An edge is a well-organized grouping of curved line segments.

This group of locations represents sharp fluctuations in the image's brightness. In image processing, edge detection is a method for feature extraction and feature detection. With the preservation of crucial image attributes, this method considerably decreases the amount of data to be processed. It may also remove less relevant information.

B. Obstacle Detection

The various techniques that can be used for detecting an obstacle are:

1) RADAR

One of the most used technology for detecting obstacles in medium to long-range. The signals used by RADAR technology are not limited by any weather conditions hence making it a deal for the job. It can also determine the velocity of the target which can be used to predict the time of impact.[6]

2) LIDAR

It is a laser-based obstacle detection system that provides 360 coverage but depends on the round trip of the laser for detection, It can be tampered with or result in false positives due to environmental conditions like fog, and rain, etc. [13]

3) Ultrasound

This technology depends on the medium like air as it is based on sound waves and cannot be used for medium to long-range detection. As one of the cheapest and most accurate solutions to close-range obstacle detection, it is used to detect nearby vehicles and obstructions. [14]

4) Camera

Provides a complete picture of the surroundings and helps us detect other things than obstacles like traffic lights, lanes, turns, etc. Many machine learning projects are based on this technology due to the amount of information available for extraction. It is limited by the visibility conditions like fog, rain, mist, or obstructions on the camera lens. [15]

5) Infrared (IR)

Other than LIDAR, Infrared can also be used at night time making it more versatile for low- visibility environments.

This paper proposes the use of camera-based technologies for easy implementation and accurate solution to our problem.

At the moment there are various machine learning and deep learning algorithms for obstacle detection such as:

1) RCNN

Stands for Regional CNN and proposes 2000 regions from the image called proposals. These proposals are then classified using CNN.

2) Fast RCNN

To create a faster object detection algorithm, the 2000 regions of proposals are not directly fed into CNN rather they are sent to a convolutional feature map and then reshaped to minimize the number of proposals. Faster than normal RCNN.

3) Faster RCNN

Both RCNN and Fast RCNN uses a selective search algorithm to identify region proposals but Faster RCNN uses another network to predict the regional proposals making it even faster than the above two.

4) YOLO (You Only Look Once)

This algorithm works quite differently from the RCNN and their modified versions. It divides the image into SxS grids, within each grid we take m boundary boxes. With each boundary box, the network makes a class probability map, which can be used to locate the object within the image[1].

3. Proposed Method

In this section, we propose Canny edge detector for lane detection and YOLO algorithm for obstacle detection.

Lane Detection

The unusual aspect detector, a boundary detection operator, makes use of a multi-level approach to perceive several edges in photographs. The use of linear filtering and a Gaussian kernel, Canny aspect detection determines the boundary depth offers every of the pixels in the noise-smoothed photo, such as direction. Nominees matching part pixels were diagnosed through the non-maximal suppression thinning procedure that these pixels undergo. This approach sets the vertex capability of each potential If the brink strength of the 2 pixels after it within the path of the gradient are not higher than the only it has, set the brink pixel to zero. The flattened aspect significance image is then adjusted using endurance. two side energy thresholds are applied in hysteresis. each possible facet pixel beneath the decrease threshold is assessed as a non-edge, and every capacity side pixel above the low threshold is one that may be related to any feasible edge pixel above the high restrict by means of a series of edged pics.

1) The user must input three variables that allows you to make use of the Canny aspect sensor. the first is sigma, or the pixel-based totally common blunders of the Gauss filter out. The low limit, that's represented as a percentage of the displayed high threshold, is the second one low parameter. The 1/3 parameter high, commonly is given as a percentage point in the variety of gradient significance values for the potential facet pixels, defines excessive threshold to be used inside the repetition. The tiers of Canny aspect detector algorithm, as proven in fig 1, are mentioned as follows:

2) Noise Reduction: The first step is to use a Gaussian filter out to the photo to get rid of any noise considering aspect detection is sensitive to photograph noise. it is fundamental to understand that the detector's overall performance could be impacted by using the selection of the Gaussian kernel's size. The detector's sensitivity to noise decreases with increasing length. additionally, as the size of the increase within the Gaussian clear out's width kernels, the localization error to detect the vertex will also drastically rise.

3) Finding Intensity Gradient of the Image: The Canny algorithms used 4 filtering to pick out the blurry photo's vertical, horizontal, and diagonally edges. An image's facet is probably orientated in some of various instructions. The smoothed image is then filtered employing a Sobel kernel in each directions to yield the first spinoff in each the route that is horizontal and the vertical direction. Those two pix permit us to perceive the brink colour and path for each pixel. The gradient route is usually contrary to edges. it's far rounded to one in all four angles—two diagonals, two sectors and one horizontal—that represent the four instructions. basic concept is straightforward: it looks for pixels with the highest values in the edge directions amongst all the points on the gradient brightness matrices. Thinner edges on the same image are the end result. However, some variance in the intensity of the edges can still be seen: some pixels appear to be brighter than others, this flaw is addressed with the two remaining phases.

4) Double Threshold: After the usage of non-most suppression, the edge pixels that remain intact offer a progressed depiction of real edges in a photograph. But positive side pixels still have noise and variant in color. facet pixels with a low gradient fee ought to be removed to account for these misguided solutions, while it is vital to maintain aspect pixels with a high gradients value. this will be performed through selecting excessive and occasional threshold values. If a pixel's gradient value is more than the high threshold value, it's far deemed to have a strong side. A pixel with a susceptible edge is one whose gradient value is each extra and much less than the low threshold fee. If a facet pixel's gradient price is lower than a low limit cost, it will likely be suppressed [16]. The specification of the two threshold values, which might be derived empirically, is primarily based on the enter picture's content material.

5) Edge Tracking by Hysteresis: At least one (1) of the pixel among the one being processed will be a strong pixel, the hysteresis involves turning weak pixel into a stronger one based on the finding from the threshold.

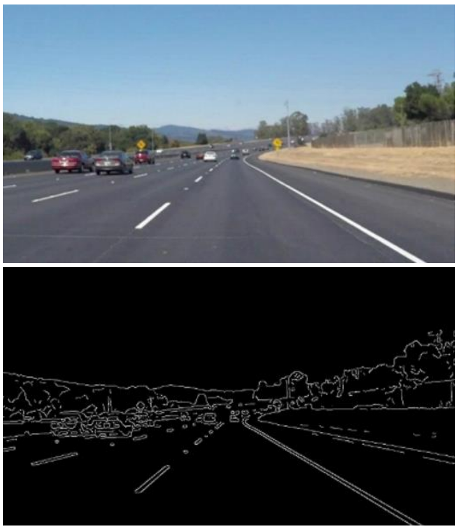

Fig 1. Image before(above) and after(below) Canny edge detection

To isolate characteristics of a specific shape inside an image, the output of clever edge detection is applied using the Hough Transform. used to identify the lanes here. Because each lane line has several lines that have been discovered. To create a single line for each lane line, we must average out all the lines. To span the entire length of the lane lines, we must also extrapolate the lane lines.

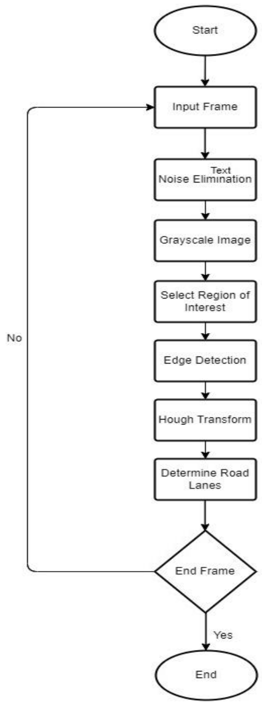

Fig 2. Flowchart for Lane Detection

The flowchart in fig 2 clearly displays the flow of the proposed solution discussed to perform lane detection.

Obstacle Detection

The N grids in the image must each have an equal dimensional SxS region for the YOLO method to function. Each one of these N grids is accountable for discovering and identifying the object it holds.

Those grids consequently predict the object call, the danger that the item might be within the cell, and the B bounding field coordinates in relation to their cell values.

Due to the fact cells from the image are used for both detection and popularity, this method considerably simplifies computation. It nevertheless induces a massive wide variety of replica predictions due to the fact many cells may are expecting the equal object with diverse container limitations guidelines. YOLO uses non-maximal suppression to deal with this trouble.

In non-maximal suppression, YOLO filters all bounding packing containers with fewer probable values. to perform this, YOLO considers the probability ratings associated with each alternative and chooses the very best one. Following that, the bounding bins with the largest Intersection over Union with the existing high possibility bounding box are suppressed. This stage is repeated till the remaining bounding packing containers are obtained.

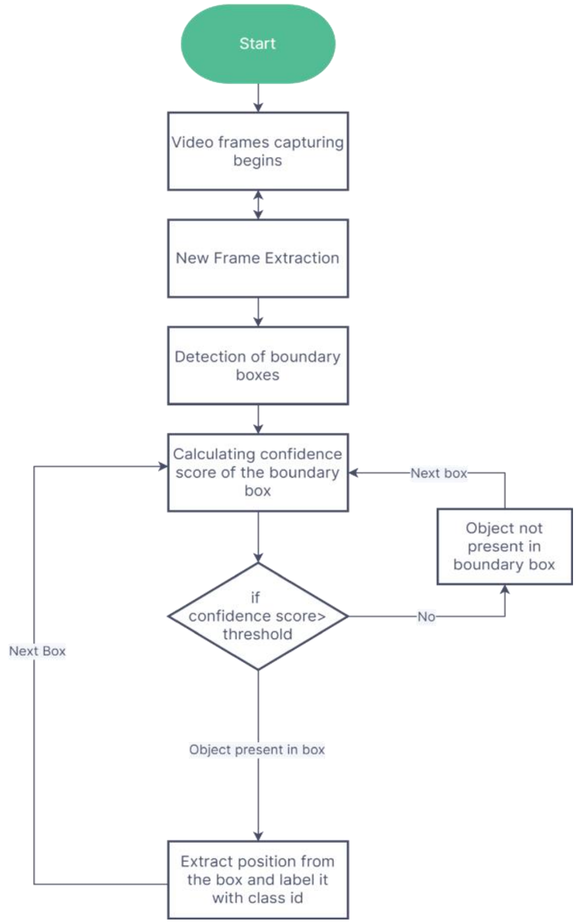

Fig 3. Flowchart for Obstacle Detection

The flowchart in fig 3 shows the flow of the proposed solution for obstacle detection discussed above.

4. Experimental Result

In this part, we conduct experiment to determine the accuracy of the system.

Dataset

Finding a sizable dataset that accounts for the anticipated difficulty of roads is the most difficult aspect of this process. This is a particular challenge due to the different road structures, colors, surroundings, and noises present.

No verified data set was found for Local obstacles so we prepared our dataset from different areas and labeled 6000+ images.

The foreign dataset had random-sized and oriented images of obstacles which did help in making our model better but we need more data from the angles on vehicles where the camera could be placed. We gathered pictures and videos of road lanes taken from 3 heights - car roof, deck and bonnet to stimulate the functioning of our model for a normal car. Train:Validation:Test ratio is 70:20:10.

Model Training and Validation

The model was trained for about 100 epochs to ensure higher accuracy and best possible model fitting. Out of these, the best fit was selected, results obtained by which are discussed in this section below.

To train our data, we need to create a training, validation and testing batch.

Fig 4. Training Batch

The training dataset as shown in fig 4 is the set of data that is fed into the machine learning model to discover and learn patterns.

Fig 5. Validation Batch

In figure 5, the validation dataset is a sample of data withheld from the model's training that is used to measure the model's skill while adjusting its hyperparameters.

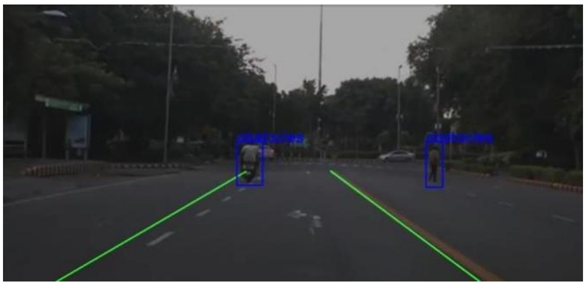

Fig 6. Output

The developed system successfully detects lane and obstacles as seen in fig 6.

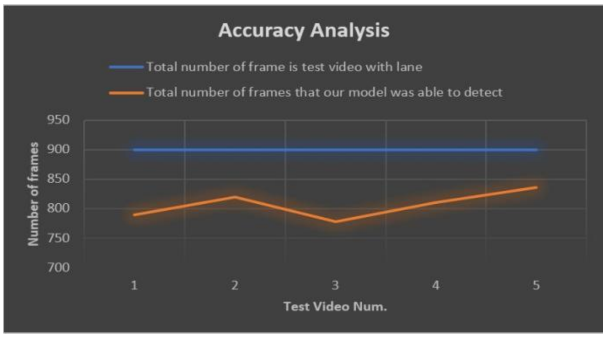

Fig 7. Quantitative frame comparison between true and predicted frames

The blue graph in fig 7 shows the total number of frames in the test video with lanes ,i.e., 900 in each of the five test videos. The orange graph shows the number of frames finely detected by the model in each test video out of 900.

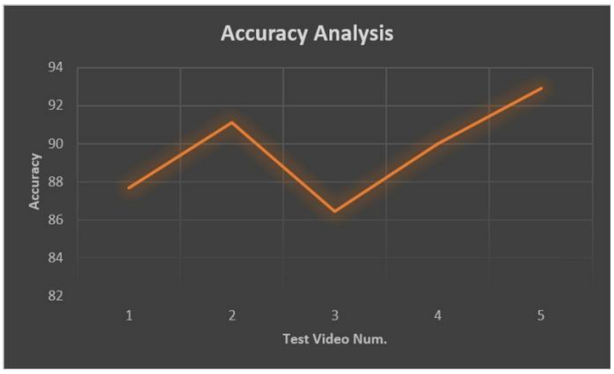

Fig 8. Positive Frame Detection Accuracy

The graph in fig 8 shows the accuracy of the model obtained by calculating the percentage of number of frames finely detected by the model to total number of frames containing lanes in each video.

TABLE I. ACCURACY OF LANE DETECTION

Test Run	*Total frames in test video*	*Number of frames detected by model*	*Accuracy*
1	900	789	87.66
2	900	820	91.11
3	900	778	86.44
4	900	810	90
5	900	836	92.88
		Average Accuracy	89.62

Table 1 represents 5 test videos, each 30 seconds long, from every video 900 frames are extracted. Assuming that each frame has a lane present in it, we are determining how many frames were identified by the system. This gives us a good estimate of the efficiency. Accuracy is calculated separately for each of the five test videos which is then used to find the average accuracy of the mode which comes out to be 89.6%.

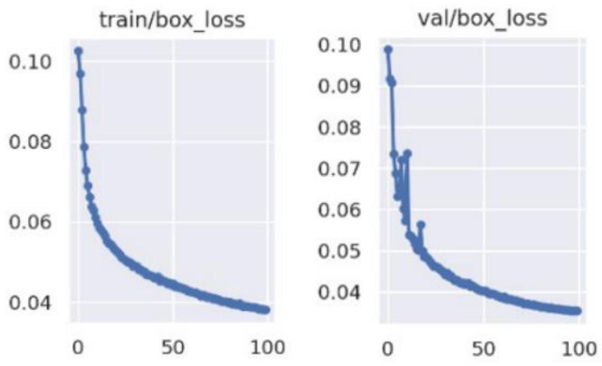

Fig 9. Box Loss Graph

The box loss graph in fig 9 represents how well the algorithm can locate the center of an object and how well the predicted bounding box covers an object. The lower the value the better. Thus it is concluded that the model detects the centre of the obstacle and predicts the boundary box with high precision.

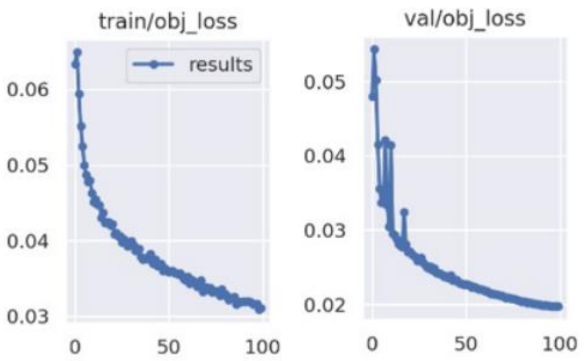

Fig 10. Object Loss Graph

In essence, abjectness' of loss is a measurement of the likelihood that a object exists in inside a suggested zone of liking. If the objectivity is high, this indicates that there may be a barrier in the image window. The graph in fig 10 shows that the model was able to identify the obstacles present in the ROI well.

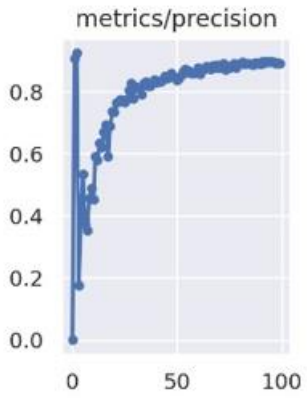

Fig 11 Precision Graph

A model's precision in foreseeing positive labels is measured by its precision. As seen in fig 11, the precision of the model is high.

Precision can be calculated using the following formula:

Precision = True Positives / (True Positives + False Positives)

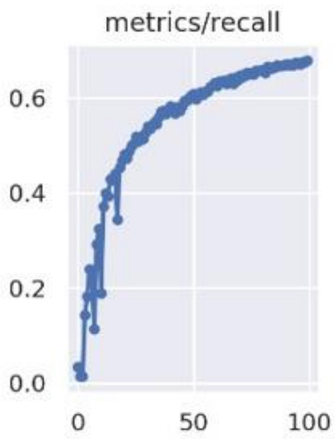

Fig 12. Recall Graph

Recall calculates the percentage of actual positives a model correctly identified (True Positive). As shown in fig 12, the model has identified true positives moderately.

Recall can be calculated using the following formula:

Recall = True Positives / (True Positives + False Negatives)

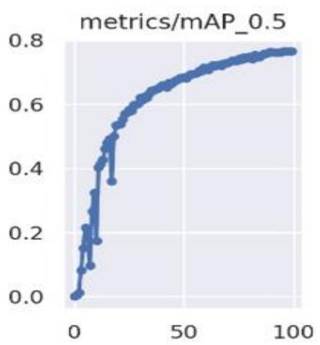

Fig 13. Mean Average Precision Graph for IoU 0.5

In fig 13, m_AP@[.95] corresponds to the mean average AP for IoU at 0.95

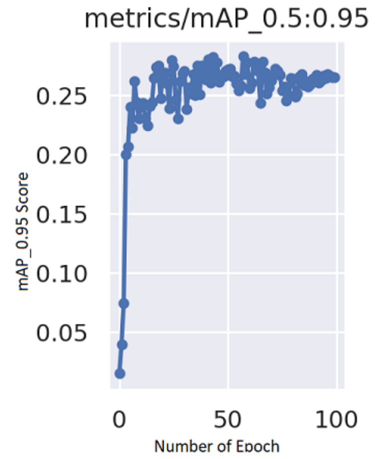

Fig 14. Mean Average Precision Graph for IoU 0.95

In fig 14, m_AP@[.5] corresponds to the mean average AP for IoU at 0.5

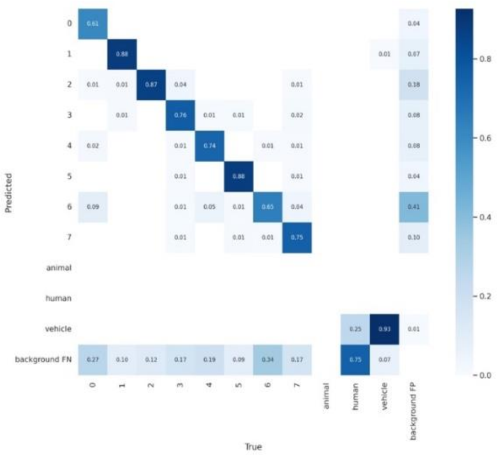

Fig 15. Correlation Matrix

A desk showing correlation coefficients between variables is referred to as a correlation matrix as shown in fig 15. every desk mobile shows the correlation between any two variables. Correlation matrices are used to summarise data, and they're also used as inputs and diagnostics for extra complicated investigations and evaluation. The linear dating between two variables is measured by means of this. It ranges from -1 - 1, with a value among

– a very bad linear correlation among 2 variables is shown via a fee of -1.

– 0 denotes the absence of a linear dating among two variables

– a totally tremendous linear correlation among two variables is indicated by a value of 1.

The association between the 2 variable is stronger the farther the correlation coefficient is from zero.

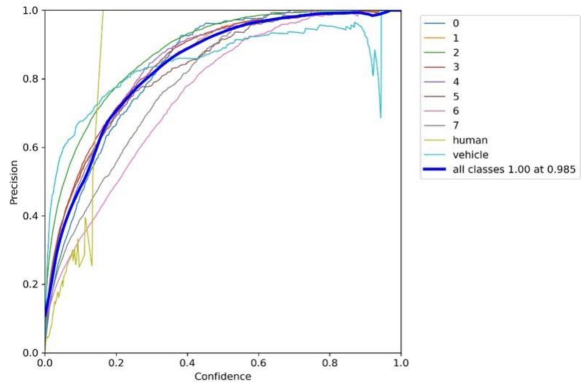

Fig 16. Precision Confidence Graph

Fig 16 represents the relationship of model precision with increasing confidence.

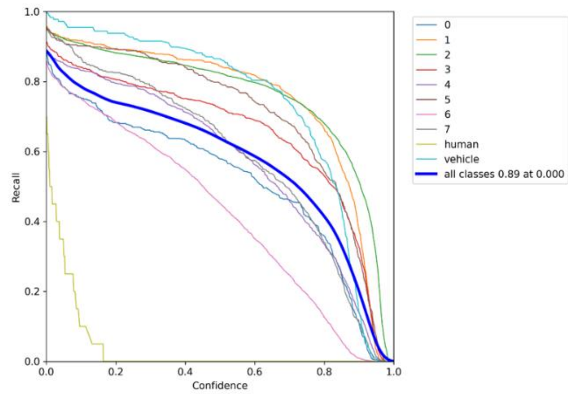

Fig 17. Recall Confidence Graph

The graph in fig 17 represents the relationship of model recall with increasing confidence.

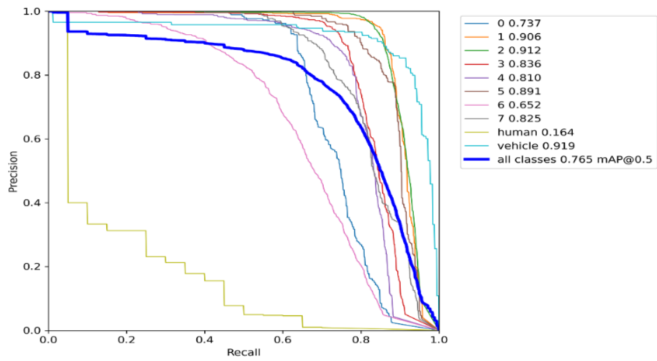

Fig 18. Precision Recall Graph

In fig 18, it is a plot of the false positive rate (x-axis) versus the true positive rate (y-axis) for several different candidate threshold values between 0.0 and 1.0. Put another way, it plots the false alarm rate versus the hit rate.

5. Conclusion and Future Works

The project implemented Canny edge detection algorithm for lane detection and YOLO for obstacle detection in images obtained from the video stream taken as an input using a camera placed on the deck of the vehicle. YOLO model has been trained using self-made data set of Indian roadways. In this paper, lanes have been identified with 89.6% accuracy and obstacles have been identified with 89% accuracy. However, detecting lanes on unstructured roads or in case of blur/low resolution videos is beyond the scope of the system. These constraints can be addressed in future works.

References

[1] D. Wu, M. Liao, W. Zhang, X. Wang, X. Bai, W. Cheng, W. Liu, “YOLOP: You Only Look Once for Panoptic Driving Perception”, School of EIC, Huazhong University of Science & Technology, 2022

[2] K.Y. Chiuand, S.F. Lin, “Lane detection using color based segmentation” , IEEE Intelligent Vehicles Symposium ,2005

[3] B.Yu and A.K. Jain, “Lane Line Detection Based On Machine Vision” , IEEE 4th International Conference on Power, Intelligent Computing and Systems (ICPICS),2022

[4] F. Camarda, “Multi-sensor data fusion for lane boundaries detection applied to autonomous vehicle”, Automatic Control Engineering, University of Technology of Compiègne, 2022

[5] F. Xu, L. Chen, J. Lou , “A real-time road detection method based on reorganized lidar data”, PLoS ONE., 2019

[6] M.Bertozzi and A.Broggi, “Gold:A parallel real time stereo vision system for generic obstacle and lane detection,” IEEE Trans. On Image Processing, vol. 7, no.1, pp.62–81, 1998

[7] G. Luo, X. Chen, W. Lin, “An obstacle detection algorithm suitable for complex traffic environment”, World Electric Vehicle Journal, vol. 13, no. 4, p. 69, Apr. 2022

[8] Y.Alon, A.Ferencz and A. Shashua, “Off road path following using region classification and geometric projection constraints,” CVPR, 2006

[9] Mariut, F., C. Fosalau, D. Petrisor., "Lane mark detection using Hough Transform" In Electrical and Power Engineering (EPE), 2012 International Conference and Exposition on, pp. 871-875. IEEE, 2012.

[10] Fani, Hongli, W. Wang., "Edge Detection of Color Road Image Based on Lab Model." In Computational and Information Sciences (ICCIS), 2013 Fifth International Conference on, pp. 298-301. IEEE, 2013.

[11] Chen, Yong, M. He., "Sharp curve lane boundaries projective model and detection." In Industrial Informatics (INDIN), 2012 10th IEEE International Conference on, pp. 1188-1193. IEEE, 2012.

[12] Prakash, J., M. B. Meenavathi, and K. Rajesh., "Linear feature extraction using combined approach of Hough transform, Eigen values and Raster scan algorithms." In Intelligent Sensing and Information Processing, 2006. ICISIP 2006. Fourth International Conference on, pp. 65-70. IEEE, 2006.

[13] The Basics of LiDAR - Light Detection and Ranging - Remote Sensing(https://www.neonscience.org/resources/learning-hub/tutorials/lidar-basics)

[14] A. Francis, G., Arulselvan, M., Elangkumaran, P., Keerthivarman, S., V. Kumar, J. (2020). Object detection using ultrasonic sensor. Int. J. Innov. Technol. Explore. Eng, 8, 207-209.

[15] Zou, X. (2019, August). A Review of object detection techniques. In 2019 International Conference on Smart Grid and Electrical Automation (ICSGEA) (pp. 251-254). IEEE.

[16] L. Lin, D. Wang, S. Zhao, L. Chen, N. Huang. "Power Quality Disturbance Feature Selection and Pattern Recognition Based on Image Enhancement Techniques", IEEE Access, 2019

[17] Kevin Harris, Lee Green, Juan González, Juan Garciam, Carlos Rodríguez. Automated Content Generation for Personalized Learning using Machine Learning. Kuwait Journal of Machine Learning, 2(2). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/180

[18] Mary Mathew, R. ., & Gunasundari, R. . (2023). An Oversampling Mechanism for Multimajority Datasets using SMOTE and Darwinian Particle Swarm Optimisation. International Journal on Recent and Innovation Trends in Computing and Communication, 11(2), 143–153. https://doi.org/10.17762/ijritcc.v11i2.6139