Dataset Evaluation for Multi Vehicle Detection using Vision Based Techniques

Aarti Kishor Ajapunje

Technische Universität Chemnitz
M. Sc. Embedded Systems
aarti-kishor. ajapunje@s2016. tu-chemnitz. de

Julkar Nine

Technische Universität Chemnitz
Department of Computer Engineering
julkar. nine@informatik. tu-chemnitz.de

Wolfram Hardt

Technische Universität Chemnitz
Department of Computer Engineering
wolfram.hardt@informatik.tu-chemnitz.de

Abstract[1]— Vehicle detection is one of the primal challenges of modern driver-assistance systems owing to the numerous factors, for instance, complicated surroundings, diverse types of vehicles with varied appearance and magnitude, low-resolution videos, fast-moving vehicles. It is utilized for multitudinous applications including traffic surveillance and collision prevention. This paper suggests a Vehicle Detection algorithm developed on Image Processing and Machine Learning. The presented algorithm is predicated on a Support Vector Machine (SVM) Classifier which employs feature vectors extracted via Histogram of Gradients (HOG) approach conducted on a semi-real time basis. A comparison study is presented stating the performance metrics of the algorithm on different datasets.

Keywords— HOG, SVM, Vehicle Detection, KITTI Dataset, Stanford Dataset

I. MOTIVATION

Road traffic mishaps at present ranks on eight position for the leading source of demise across the globe. A Global Report focusing on Road Safety was published by World Health Organization (WHO) in the year 2018, depicting the rise in crash fatalities accounting to 1.35million each year [2]. An average of 22,800 people passed away inroad mishaps in the European Union (EU) in 2019, with drivers or passengers accounting for 44.2 percent of fatalities and pedestrians accounting for 20.2 percent [3]. Multiple studies have indicated that the ultimate source of traffic collisions is due to human error [4] [16] [31] [11]. A survey was carried out in 2018 by the National High-way Traffic Safety Administration (NHTSA), where hu-man error was the leading cause of the accidents, constituting to be around 94 percent [25]. Other contributing factors were vehicle system malfunctions and road infrastructure. Furthermore, the European New Car Assessment Program (NCAP) have articulated that circa 94 percent of users concurred that safety is a prime concern while driving [12] [1]. Keeping this in mind, several research institutes have been established around the world with the goal of improving road safety, such as NHTSA(US), CARE(EUROPE), BASt and AARU(GERMANY), CzI-DAS(CZECH), TRL(UK) etc.

While driving on the highway, one must be fully aware while changing the lane. As most of the drivers opt to maintain their speed on the highway, changing lanes might turn out to be fatal if one isn't sure if the target lane is empty. Collision while executing a lane change can be caused by a variety of vital factors as per [7] including failure in recognizing a vehicle, unaware of the presence of a dangerous vehicle in its vicinity and delay in implementation of avoidance maneuver. Lane change crashes accounted for an estimated 4% of all the crashes and about 0.5% of the overall crash fatalities occurred in the United States as stated by Knipling [18]. [24] conducted a Driver's Behavior Survey involving 429 participants, where more than half of them (approx. 57%) have acknowledged that the major cause of their sudden lane change was a result of the driver's distraction caused by the interaction with a fellow passenger or due to their mobile usage.

Traffic mishaps can be intercepted by fixating on the following three tactics as per [22]: (1) Transforming human practices; (2) Integration of advanced Automotive safety features; and (3) Road and Highway infrastructure strategies. Initiating public safety by Imposing Law and Order and encouraging people to follow them, ensuring proper training via driving classes and issuing up-to-date guidelines can all help to improve human practices, while Road and Highway infrastructure solutions comprises of building multi-lane roadways employed with efficient Traffic management systems. Furthermore, Modern Driver Assistance Systems (ADAS) withal to Electronic Stability Control(ESC) are some of the numerous Advanced Automotive Safety Systems which acts as an extra pair of eyes and assist the driver, thus ensuring safe and smooth travel.

ADAS is a modern automated vehicle-based electronic safety system that perceives its environment while driving via cutting-edge sensors such as LIDAR, RADAR and vision-based, thus reducing the risk of human-error-related accidents and ensuring a safe and comfortable travel. Some of the ADAS-based features are as follows [8].

(i) Vehicle Platoon: When a set of vehicles travel in a close proximity on a highway by communicating with each other resulting in an effective traffic management without compromising on safety and comfort of the passenger. In [17], various approaches related to Vehicle Platooning have been discussed.

(ii) Stop-and-Go Traffic: The vehicle alerts the driver to restore control in case if its speed increases above an acceptable threshold during a traffic jam in Conditional Automation. Using a recursive least-square algorithm to analyze human driving characteristics, a driver adaptive control technique for stop-and-go systems has been presented in [33].

(iii) Blind Spot Detection and Overtaking Maneuver: Here, the vehicle must localize and monitor other vehicles in its vicinity, alerting the driver in case of a potential collision.

II. INTRODUCTION

Vision-based on-road vehicle detection and its identification has been extensively researched on during the last decade. RADAR accurately determines the location, speed, and direction of the target vehicle. It is also insensitive to various illumination and weather conditions, however, its inability to determine the exact shape of the target as well as its limitation to detect vehicles on a wider Field of View (FOV) are some of its major drawback. On the other hand, even though LIDAR gives high 3D precision of the targets located in its surrounding and is unaffected in different weather conditions, it is the most expensive out of the three due to its high computing power need which also makes it prone to system malfunctions and software bugs. Cameras can create an autonomous driving experience that closely resembles that of a human driver. It provides detailed description of the surrounding visual data which is fed as an input for training models and executing complex predictions from neural networks for target detection. Also, with the availability of compact, high-quality cameras at a much cheaper price as compared to that of LIDAR or RADAR in addition to the advancements in hardware such

as Graphical Processing Units (GPUs) and multi-core processors have made real-time implementation possible. Lane Departure Warning (LDW), Front/Rear Collision Avoidance, Blind Spot Detection, Surround View, Pedestrian and Traffic Sign Detection are few of the camera based ADAS components.

This paper is further branched into multiple segments. Segment III shows a glimpse of various Vehicle detection algorithm formulated on image processing. Segment IV gives an in-depth view of our suggested algorithm, subsequently followed by its Implementation in Segment V. In the end, Segment VI analyses the results and collate the performance of the algorithm on different datasets.

III. RESEARCH FOCUS

Vehicle detection due to its broad range of applications, including modern driver assistance systems (ADAS) [21] and traffic surveillance systems [5] [14],vehicle counting, and rescue, has a pivotal role in Intelligent Transportation Systems(ITS). Object detection can be further sub-categorized into two steps: Hypothesis Generation and Hypothesis Verification. Hypothesis Generation deals with determining the region in an image where the object might exist, whereas Hypothesis Verification deals with confirmation of vehicle in an image in a specific Region of Interest (ROI) [27]. Feature-based, motion-based, and classifier-based are the three different types of monocular vehicle-based detection methods as described in [26].

Feature-based or appearance-based features focus on the image characteristics such as symmetry, vertical and horizontal edges, color, texture, corners etc. to determine if a vehicle is present in an image. In symmetry-based method, the front and rear-view of a car appears to be symmetrical, which helps in distinguishing itself from the background [28] [20]. Color-based vehicle segmentation is another approach which helps to discern itself from its background [23] [34]. However, its sensitivity to weather and illumination changes makes color-based vehicle detection risky.

Histogram of Oriented Gradients (HOG) features are obtained by evaluating edge operators over the image, then discretizing and binning the edge intensities into a histogram. After that, the histogram is employed as a feature vector. Initially HOG features were executed for pedestrian detection [10], however it can now be utilized over a vast variety of applications, including vehicle detection. HOG are picture attributes that are descriptive and show strong detection performance in a range of computer vision tasks, including as vehicle recognition detection, but they take a long time to compute, affecting the real-time execution. This drawback has been overcome by employing HOG feature extraction on a GPU. Haar-like features were originally used in real-time face identification [30] and are made up of the difference of the sum of pixels of areas inside a certain image patch. These differences are then evaluated against a certain threshold for object classification. The difference of Gaussians is computed in the Scale Invariant Feature Transform (SIFT), and the eigen values of the generated Hessian matrix are attained by thresholding to ensure that recognized features are inside edge boundaries. On the

contrary, the Speeded Up Robust Features (SURF) feature detector is a more effective and faster version of SIFT. Wavelet responses are used by SURF in both the horizontal and vertical dimensions. The dominant orientation is found by using Gaussian weights along with implementation of 60 degree sliding window orientation to determine the sum of all responses.

Classifier-based detection algorithms differentiate vehicles from non-vehicles by understanding the properties of vehicle appearance from a group of training dataset, which are often obtained from an annotated database that includes both images with and without a vehicle, in order to reflect the diversity in the vehicle class. A multitude of feature vectors, including Haar-like features, have been classified using the Sup-port Vector Machine (SVM) classification [9]. HOG properties are also used as the foundation for SVM classification, according to [29].

Many conventional methods for vehicle detection are based on Background Subtraction. [32] have analyzed and made a comparison of eight different methods based on background subtraction. However, in all these methods have overly restrictive assumptions which fails while implementing in realistic complex environment. Deep learning- based approaches has exhibited state-of-the-art, human-competitive, and sometimes better-than-human performance in a variety of computer vision applications, including object recognition, image classification/retrieval, and semantic segmentation. [15] proposed a vehicle detection system for the Hsuehshan tunnel in Taiwan using Background Subtraction and Deep Belief Network (DBN) consisting of three hidden layers architecture which shows an accuracy of 96. 59%. [6] have implemented Single Shot Multibox (SSD) Vehicle Detector based on feed forward convolution network(VGG16). However, it failed to detect distant and occluded vehicles.

A vehicle detection strategy has been propounded on a semi-real time basis which can further be utilized to signal the driver notifying the presence of target vehicle in the adjacent lanes, thus preclude the possibility of a collision. A comparison has been drawn between the performance of the Vehicle Detection model with respect to different datasets.

IV. CONCEPT

In this paper, we put forward the concept and implementation of a Vehicle Detection System predicated on a Linear Support Vector Machine(SVM) Classifier which is trained on features issued via Histogram of Oriented Gradients(HOG) feature extraction approach for the identification of the vehicles present in the vicinity of the source vehicle.

First, the input photos are downsized to 64*64 in or-der to reduce computation time during the training stage. Since every vehicle has its own distinctive edges and contours, HOG helps in determining its change in intensity gradients alongside its magnitude and directions in an image. The image is first split into small regions, also known as cells, where histogram of gradients is computed corresponding to each cell and after combining these histograms, a HOG feature descriptor is obtained. These feature vectors are

utilized to train a Linear SVM Classifier and help it in distinguishing between vehicle and non-vehicle classes.

During the prediction stage, the video clip acquired from the source vehicle's camera is initially split into multiple frames over which sliding window is applied to locate the regions having the possibility of the presence of a vehicle. Vehicles who are away from the source vehicle appear to be smaller in the video stream. Employing a variable sliding window search helps in detecting the vehicle of different sizes. Subsequently, HOG feature ex-traction is carried out over these regions of interests followed by the prediction utilizing the previously trained Linear SVM model which determines whether the detected object belongs to either a vehicle or a non-vehicle category. Finally, a bounding box is generated stating successful vehicle detection. However, multiple conjoining bounding boxes might refer to the same vehicle. This can be resolved via Heat map. A heat map comprising of a null-initialized NumPy array is first generated where for every successful vehicle detection, the pixels within the corresponding bounding box are incremented. Regions depicting high values results in higher chances of the presence of a vehicle. Only the region containing higher value than the threshold is accepted followed by the formation of bounding box around it. The above process is repeated throughout the whole frame. The workflow of Vehicle Detection algorithm is illustrated in Figure 1.
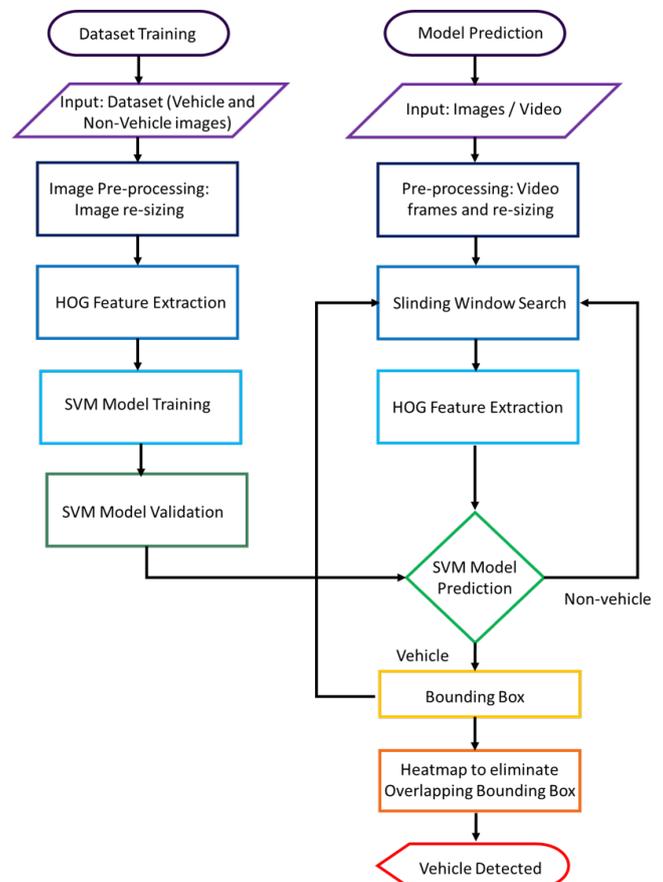

Figure 1: Vehicle detection algorithm workflow

V.   IMPLEMENTATION

The suggested semi-real time-based Vehicle detection algorithm is assessed on our target hardware i.e., Laptop with Intel Core i5 Processor utilizing multiple image processing and machine-learning based Libraries for instance OpenCV and Skimage for image processing, NumPy for mathematical calculations, Scikit-learn for extracting machine learning based modules using Python programming language. The algorithm can be further categorized into following steps: Dataset Creation, Model Training, Model Validation and Model Prediction as shown in Figure 2.

Figure 2: Vehicle detection pipeline

A.   Dataset Creation

An appropriate and relevant dataset has its influence on the overall functioning of the classifier and aids in finding similarities in pattern in the data thus making useful predictions. The dataset should reflect its purpose with respect to the application for which it has been considered. In addition to a prime quality dataset containing sufficient sample size, a balanced dataset is equally significant. A balanced dataset gives equal significance to each class. A dataset with an unbalanced class i. e., one class comprising of more images than the other, resulting in miscalculations leading to classification models having lower accuracy, unbalanced accuracy, and an unbalanced detection rate. It is equally important to have a dataset with adequate samples for training the model. After sufficient research and careful consideration which suits the requirements, a few standard datasets such as KITTI dataset [13] and STANFORD dataset [19] along with a custom dataset has been considered to assess the performance of our Vehicle detection model.

B.   Dataset Training

The first step of Dataset Training involves pre-processing using OpenCV Library where image resizing is performed. This is shortly followed by HOG feature extraction process employed on the input images which focuses on the distinctions of the corners and contours in relation to the vehicle to distinguish itself from its surrounding. The skimage.feature.hog() function from Skimage Library takes the image as an input along with several other HOG hyperparameters including the total orientation bins, cell and block size and the type of block normalization to provide HOG feature descriptor of the image as an output. Apart from HOG features, color channel histogram features and spatial features are also computed. Color channel histogram points out the color distribution in an image. This is useful as the color profile of the vehicle stands out in comparison with its surrounding environment. Furthermore, spatial features point out the geographical distribution of points in an image. All the cells located in each block undergo feature

normalization to make them insensitive to sudden illumination and changes in edge contrast using StandardScalar class from sklearn.preprocessing module having null mean and unit variance. All these image feature descriptors along with their corresponding image labels are appended into a feature vector array and are provided as an input for training a LinearSVM Classifier model to distinguish between a vehicle and a non-vehicle class. The SVM Classifier used here is extracted from the LinearSVC module from Scikit-learn Library.

The input data is randomly jumbled up and fragmented into training, validation, and test datasets in the ratio of 75:15:10. Initially SVM Classifier is trained on the training dataset, followed by its performance assessment on the validation set. All the misclassification samples from the validation set are further added to the training dataset and the classifier model is retrained. Also, the hyperparameters are fine-tuned to enhance the performance of the classifier until the best model is achieved. After this, the prediction of the SVM classifier on the test data takes place which finally assess the overall accuracy of the model. The overall stages of Dataset Training are illustrated in the Figure 3.
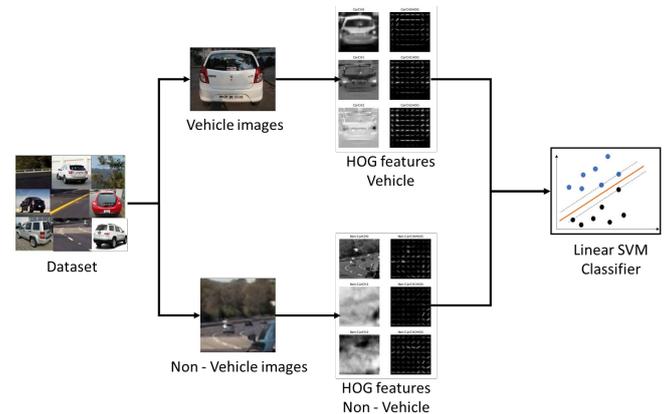

Figure 3: Various Stages of Dataset Training using SVM model

C.   Classifier Prediction

To access the overall performance of the SVM Classifier, the input video is first pre-processed and divided into frames. In every frame, sliding window is employed to detect the presence of vehicle. Sliding window with variable scaling size is employed, having a small scaling size initially and with every step the scaling size of the sliding window increases. This aids in detecting vehicles of different sizes throughout the frame. Moreover, since SVM Classifier has been trained on (64*64) image size, it is crucial to extract an image patch having similar size before extracting features from them. Furthermore, only the lower half of the frame below the skyline is inspected for the presence of a vehicle. This remarkably enhanced the overall detection results. On the other hand, each vehicle might contain multifold conjoining detections which are eliminated with the help of heatmap generation. The various stages of SVM Classifier Prediction are illustrated in the Figure 4.
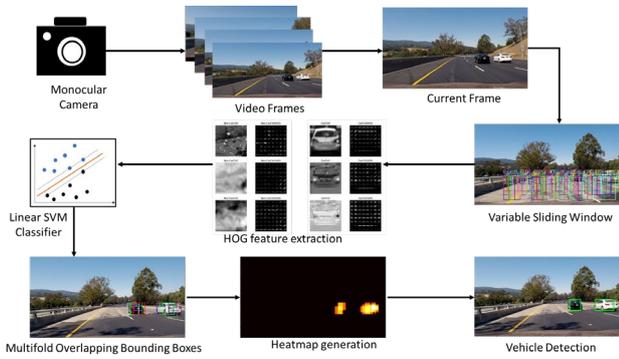
Figure 4: Various Stages of SVM Classifier Prediction

VI. Results and Evaluation

The experimental results achieved by our Vehicle Detection System on different datasets is perused in this segment along with further inspection of the performance metrics of our model including Precision, Recall along with F1 score.

It is of utmost importance to find the best set of HOG Hyperparameters which requires refinement though multiple trial and error. After multiple trials, the most suitable HOG parameters have been finalized as seen in Table 1.

HOG Hyperparameters	Values
Color Space	YCrCb
HOG Orientations	9
HOG no. of pixels per cell	8
HOG no. of cells per block	2
HOG Channels	(0,1,2)
Image patch size	(64,64)
No. of Histogram bins	16
Spatial binning dimensions	(16,16)
Spatial features	On
Color Histogram features	On
HOG features	On

Table 1: Finalized HOG Parameters values

A. KITTI Dataset

KITTI Car Dataset [13] consists of 8792 car images and 8968 non-car images amounting to a total of 17760 input images. The dataset contains balanced images taken from five categories captured from different orientations as well as different distances. Hence, even though it contains low-resolution images, when an SVM model is trained using these images, it achieves an F1 score of 88.1% and 3FPS without any optimization. It even succeeds in recognizing cars which are at a distant location from the source vehicle by fine tuning the scaling and step-size of the sliding window search. This results in the rise in computational complexity which affects both the FPS and Accuracy. Therefore, a

balance must be struck amongst FPS and Accuracy, resulting in around 98.76% F1 score and 7 FPS with optimization on simple highway videos. The output results attained using KITTI Dataset can be seen in Figure 5.

Figure 5: Examples showing Vehicle Detection employing KITTI Dataset.

B. Stanford Dataset

Stanford Car dataset [19] consists of 8145 car training images and 8041 car test images amounting to a total of 16,186 high quality car input images. The dataset suffers from the drawback of low Accuracy of circa 76%, as majority of the images from this dataset have the side-view of the car. Thus, due to this flaw in the dataset, the algorithm sometimes fails to detect the rear-end or front of the car. Furthermore, it also fails to detect some specific type of cars. The output results attained using Stanford Dataset can be seen in Figure 6.

Figure 6: Example showing Vehicle Detection employing Stanford Dataset.

C. Custom Dataset

After the evaluation of the classifier performance on KITTI and Stanford datasets, a custom dataset is formed which ameliorate the forementioned drawbacks of Stanford to a certain extent. Custom dataset is a mixture of KITTI, Stanford and random car-images acquired from Google, amounting to around 1650 training images and 330 test images amounting to a total of 2200 car images. The dataset undergoes pre-processing owing to differences in image dimensions before employing feature extraction. Due to the better preparation of the dataset, leads to a fewer number of False Positives and False Negatives as compared to the Stanford dataset, resulting in higher F1 score of around 84.44% after optimizations. The output results attained using Custom Dataset can be seen in Figure 7.

Figure 6: Examples showing Vehicle Detection employing Custom Dataset.

Optimizations includes using a validation dataset to retrain the miss-classifications achieved after the initial SVM training, reduced ROI for implementation of Sliding Window, variable scaling size with around 0.75 overlap,

reduction in the number of sliding windows result in a noteworthy improvement in the Accuracy and FPS achieved. A comparison between the performance metrics obtained by the Classifier model on simple highway environments concerning the above datasets is illustrated in the Table 2. Additionally, the impact of reduction in the resolution of the Video on the overall Accuracy and FPS achieved is presented via line graphs in Fig. 7 and 8 and respectively.

Parameters	KITTI Dataset	Stanford Dataset	Custom Dataset
Kernel	Linear	Linear	Linear
Test set Images/ Videos(HD)	888/14	855/14	330/14
Total Cars	762	762	762
Detected Cars	668	472	569
False Positive	42	130	71
False Negative	52	160	122
Precision	0.94	0.78	0.88
Recall	0.92	0.74	0.82
F1-score	0.92	0.76	0.84
FPS	7	5.25	4.3

Table 2: Comparison table for performance metrics between different datasets
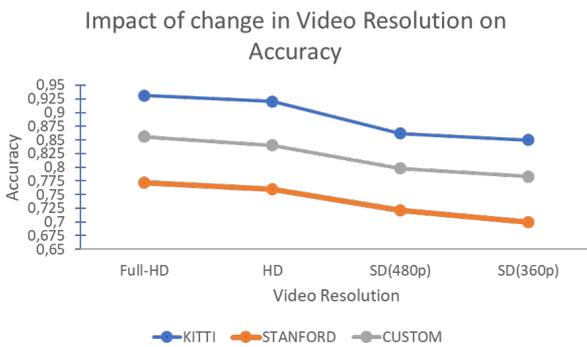

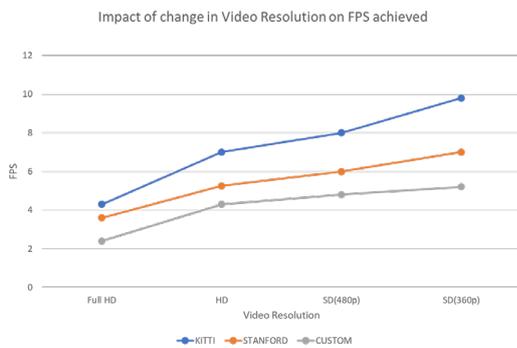Figure 7: Impact of Change in Video resolution on Classifier's Accuracy

Figure 7: Impact of Change in Video resolution on Classifier's Accuracy

The above experimental results illustrate the superior performance of KITTI dataset in comparison with other two datasets. To further scrutinize the robustness of the model with respect to KITTI dataset, the model is assessed on numerous dashcam videos acquired from Google which are based on diverse complex environments including, tunnels and bridges having sudden illumination and contrast changes, transient shadows on roads, heavy-traffic areas consisting of objects having a similar box-like structure to that of a car. This often resulted in numerous misclassifications leading to reduction in Accuracy indicating low robustness of the model due to lack of diversity in the dataset.

VII. CONCLUSION & FUTURE ASPECTS

This paper presented our Vehicle Detection approach formulated on Machine Learning on a semi-real time basis along with a comparison study of its performance on three types of datasets. All datasets have exhibited competency in successful Vehicle detection in simple environments with varying extent of performance capabilities. Experimental outcomes of numerous pre-recorded video segments are illustrated. Evaluation of the custom dataset with respect to the standard open-source datasets has been made. An accurate and robust Vehicle detection have multifold practical applications ranging from traffic flow forecast, collision avoidance, vehicle platooning leading to an effective Transportation system.

The future tasks will incorporate the implementation of our Vehicle detection approach on various edge-devices including Raspberry Pi 3B, 4B+ and Jetson Nano to assess its performance in semi-real and real-time circumstances. Also, due to the vision impairment of the camera, the contrast and glaring of the windshield are easily captured. Additionally, illumination changes and transient shadows of the roads and the lack of diversity in all the datasets in terms of car samples captured during changing seasons, throughout different times of the day results in misclassification of the data in complex surroundings. To address these issues, a much diverse and rich dataset like Berkley DeepDrive (BDD100K) [35] can be employed to make the algorithm more robust. Furthermore, occlusion arises occasionally during the overtaking of target vehicle due to ROI shifting, leading to False Negatives. To rectify this, prioritization must be enforced to focus on a specific car. The authors intend to further optimize and increase the FPS as well as the robustness of the proposed algorithm.

REFERENCES

[1]   The european new car assessment programme vehiclesafety campaigns. https://www. euroncap. com/en/vehicle-safety/safety-campaigns/. Accessed: 2021-12-01.

[2]   Who global status report on road safety 2018. Geneva: World Health Organization, 2018.

[3]   European commission 2019 road safety statistics: what is behind the figures? https://ec. europa. eu/commission/presscorner/detail/en/qanda201004,2020. Accessed: 2021−12−01.

[4] T. Allahyari, G. Saraji, J. Adl, M. Hosseini, M. Iravani,M. Younesian, and S. Kass. Cognitive failures, driving errors and driving accidents. *International journal of occupational safety and ergonomics : JOSE*, 14:149–58, 02 2008.

[5] N. E. Buch, S. A. Velast́ın, and J. Orwell. A review of computer vision techniques for the analysis of urban traffic. *IEEE Transactions on Intelligent Transportation Systems*, 12:920–939, 2011.

[6] K. -H. Chen, T. D. Shou, J. K. -H. Li, and C. -M. Tsai. Vehicles detection on expressway via deep learning: Single shot multi-box object detector. *2018 International Conference on Machine Learning and Cybernetics (ICMLC)*, 2:467–473, 2018.

[7] J. D. Chovan, L. Tijerina, G. Alexander, and D. L. Hendricks. Examination of lane change crashes and potential ivhs counter measures. Technical report, National Highway Traffic Safety Administration. U. S. Department of Transportation,1994.

[8] J. Collado, C. Hilario, A. de la Escalera, and J. Armingol. Model based vehicle detection for intelligent vehicles. pages572 – 577, 07 2004.

[9] C. Cortes and V. Vapnik. Support-vector networks. Chem. Biol. Drug Des. , 297:273–297, 01 2009.

[10] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2005)*, 2, 06 2005.

[11] T. A. Dingus, F. Guo, S. Lee, J. F. Antin, M. A. Perez, M. Buchanan-King, and J. M. Hankey. Driver crash risk factors and prevalence evaluation using naturalistic driving data. *Proceedings of the National Academy of Sciences*,113:2636 – 2641, 2016.

[12] C. Eising, M. Glavin, E. Jones, and P. Denny. Automotive blind-zones: A review of legislation and the use of close-range camera systems. 2008.

[13] A. Geiger, P. Lenz, and R. Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. Pages 3354–3361, 05 2012.

[14] J. -W. Hsieh, S. -H. Yu, Y. -S. Chen, and W. -F. Hu. An automatic traffic surveillance system for vehicle tracking and classification. *Intelligent Transportation Systems, IEEE Trans-actions on*, 7:175 – 187, 07 2006.

[15] B. Huang, J. -W. Hsieh, and C. Tsai. Vehicle detection in hsuehshan tunnel using background subtraction and deep belief network. *In Intelligent Information and Database Sys-tems - 9th Asian Conference, ACIIDS 2017, Proceedings*, Lecture Notes in Computer Science (including sub series Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pages 217–226. Springer Verlag, Jan. 2017.

[16] M. Karacasu and A. Er. An analysis on distribution of trafficfaults in accidents, based on driver's age and gender: Eskisehir case. *Procedia - Social and Behavioral Sciences*, 20:776–785, 12 2011.

[17] P. Kavathekar and Y. Chen. Vehicle platooning: A brief survey and categorization. volume 3, 01 2011.

[18] R. Knipling. Ivhs technologies applied to collision avoidance: Perspectives on six target crash types and counter measures. 1993.

[19] J. Krause, M. Stark, J. Deng, and L. Fei-Fei. 3d object representations for fine-grained categorization. InProceedings- *2013 IEEE International Conference on Computer Vision Workshops, ICCVW 2013*, Proceedings of the IEEE International Conference on Computer Vision, pages 554–561,United States, Jan. 2013. Institute of Electrical and Electronics Engineers Inc. 2013 14th IEEE International Conferenceon Computer Vision Workshops, ICCVW 2013; Conference date: 01-12-2013 to 08-12-2013.

[20] A. Kuehnle. Symmetry-based recognition of vehicle rears. *Pattern Recognition Letters*, 12(4):249–258, 1991.

[21] B. -F. Lin, Y. -M. Lin, L. -C. Fu, P. -Y. Hsiao, L. -A. Chuang, S. -S. Huang, and M. -F. Lo. Integrating appearance and edge features for sedan vehicle detection in the blind-spot area. *IEEE Transactions on Intelligent Transportation Systems -TITS*, 13:737–747, 06 2012.

[22] M. Lu, K. Wevers, and R. Heijden. Technical feasibility of advanced driver assistance systems (adas) for road safety. *Transportation Planning and Technology*, 28:167–181, 012005.

[23] R. Mejia-Iigo, M. Barilla-Perez, and H. Montes-Venegas. Color-based texture image segmentation for vehicle detection. pages 1 – 6, 02 2009.

[24] M. Shawky. Factors affecting lane change crashes. IATSSResearch, 44(2):155–161, 2020.

[25] S. Singh. Critical reasons for crashes investigated in the national motor vehicle crash causation survey. Traffic Safety Facts Crash Stats, National Highway Traffic Safety Administration, 2015.

[26] S. Sivaraman and M. M. Trivedi. Looking at vehicles on the road: A survey of vision-based vehicle detection, tracking, and behavior analysis. *IEEE Transactions on Intelligent Transportation Systems*, 14(4):1773–1795, 2013.

[27] Z. Sun, G. Bebis, and R. Miller. On-road vehicle detection: A review. *IEEE transactions on pattern analysis and machine intelligence*, 28:694–711, 06 2006.

[28] S. S. Teoh and T. Braunl. Symmetry-based monocular vehicle detection system. *Machine Vision and Applications -MVA*, 23, 09 2011.

[29] Q. Tian, L. Zhang, Y. Wei, W. wei Fei, and W. hua Zhao. A rapid method based on vehicle video for multi objects detection. *Advances in Mechanical Engineering*, 5:546752, 2013.

[30] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. volume 1, pages I–511,02 2001.

[31] S. Weber, K. Tschech, K. Blum, A. Ernstberger, andV. Labenski. Different types of distraction causing accidents. *DDI2018 Gothenburg Book of Abstracts*. 6th International Conference on Driver Distraction and Inattention, October2018.

[32] Y. Xu, J. Dong, B. Zhang, and D. Xu. Background modeling methods in video analysis: A review and comparative evaluation. *CAAI Transactions on Intelligence Technology*, 1, 06 2016.

[33] K. Yi, J. Hong, and Y. D. Kwon. A vehicle control algorithmfor stop-and-go cruise control. Proceedings of *The Institution of Mechanical Engineers Part D-journal of Automobile Engineering - PROC INST MECH ENG D-J AUTO*, 215:1099–1115, 10 2001.

[34] Z. Zheng and B. Wang. On-road vehicle detection based oncolor segmentation and tracking using harris-sift. *Advanced Materials Research*, 433-440:5334 – 5338, 2012.

[35] Yu, F., Chen, H., Wang, X., Xian, W., Chen, Y., Liu, F., Madhavan, V., & Darrell, T. (2020, June). BDD100K: A Diverse Driving Dataset for Heterogeneous Multitask Learning. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*