IJARSCT
ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

IJARSCT
Impact Factor: 6.252

Volume 2, Issue 1, December 2022

Vehicle Automation System

Prof. B. N. Babar[1], Kunal Bhokare[2], Vijay Maholkar[3], Gaurav Takle[4]
Professor, Department of Information Technology[1]
Students, Department of Information Technology[2,3,4]
Sinhgad Institute of Technology, Lonavala, Maharashtra, India

Abstract: *In recent years, vehicle and pedestrian detection is a hot research topic in the field of computer vision and artificial intelligence, they are widely used in the fields of automatic driving safety and security and pedestrian analysis. Deep learning has made significant breakthroughs in vehicle and pedestrian detection application. With the development of deep learning, the vehicle and pedestrian detection method based on deep learning greatly improves the accuracy of the model. The Efficient Det algorithm is optimized and tailored for real time object detection the detectors with very high accuracy are computationally resource demanding but Efficient Det strikes a good balance between resource constraints and high accuracy. In this paper, we present the implementation and the comparative study followed by the results of tasks such as vehicle and pedestrian detection algorithm based on state of the art object detection algorithm Efficient Det which is based on Efficient Net architecture, especially for small vehicles detection. The validity of the algorithm is verified against Udacity's Self Driving Car dataset.*

Keywords: Efficient Net, Efficient Det, Kalman filter, Object Detection, Object Tracking, Vehicle detection, Pedestrian detection

I. INTRODUCTION

1.1 Motivation

- Vehicle detection in digital image sequences is one of the key technologies of Intelligent Transportation Systems(ITS).
- Vehicle detection is also important in computing traffic congestion on highways.
- Pedestrian Detection in images or videos plays a vital role in ADAS (Advanced Driver Assistance System) and AV (Autonomous Vehicle).
- Automatic pedestrian detection and vehicle detection is important in the development of autonomous vehicles.

1.2 Problem Definition and Objectives

A vehicle and pedestrian detection as well as tracking model that will be built by using deep convolutional neural network. This network will be trained on available dataset to provide reliable decision support for vehicles.

A. Objectives

- To develop a driver assistance system with major functions: Pedestrian and Vehicle detection and tracking by using deep learning.
- To develop a hybrid system using Kalman filter (with background removal) and CNN for detection and tracking of moving objects in videos.
- To provide a comprehensive survey on vehicle detection, pedestrian detection and tracking.
- To study briefly on the available driver assistance system techniques.

IJARSCT
ISSN (Online) 2581-9429
IJARSCT
Impact Factor: 6.252
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, December 2022

II. LITERATURE SURVEY

Table 2.1 Comparison table

Ref. No	Category	Author	Technique	Strengths	Limitations	Results	Accuracy/Precision
1.	Traffic Sign Detection Technique	Karan Singh, Nikita Malik	Convolution al neural network (CNN)	Simple and effective procedure.	model can be further improved by using random cropping for training and validation data	Precision is 97.8% and recall is 98.06%.	95%
2.	Traffic Sign Detection Technique	Ameur Zaibi , Anis Ladgham, a nd Anis Sakly	Enhanced LeNet-5	High accuracy.	Computationally expensive.	Prediction is very Accurate	99.84 %
3.	Traffic Sign Detection Technique	Njayou Youssouf	FasterRCNN and YOLOV4	High accuracy in predictions.	Slightly less Accurate	Algorithm is lighter and faster	99.20 %
4.	Pedestrian Detection Technique	Lang Wang , Jiaqi Gui, Zhe-Ming Lu, and Cong Liu	HOG-SVM	Can achieve pedestrian detection and tracking in real time.	Accuracy relies in extraction zone.	Outperforms the HOG+SVM and GMM+HOG+SVM algorithms in recognition accuracy	90.84 %
5.	Pedestrian Detection Technique	Geethapriya. S, P. Kumar	Integration of Resnet and Yolo V2	Less loss functions.	This algorithm can also be improvised by integrating with any other efficient algorithms.	Two networks is to increase the accuracy rate and to minimize the loss function.	87.7 %
6.	Vehicle Detection Technique	Linkai Chen,Feiyue Ye, Yaduan Ruan ,	CNN	fast and accurate	Not so good for Occlusion	Outperforms Faster R-CNN and SSD,specially small vehicles	75.7% mean Averag e Precision
7.	Vehicle Detection Technique	Huanshen g Song, Haoxiang Liang Huaiyu Li, Zhe Dai and Xu Yun	YOLOv3	good performance	Based on limited dataset.	Good Practicibility	87.88% mean Averag e Precision

III. SOFTWARE REQUIREMENTS SPECIFICATION

3.1 Introduction

Project Scope

Input will be video based particularly taken from highways (traffic scenario).Detection and tracking of vehicles will be done which are appearing in the input video

Assumptions and Dependencies

- The video uploaded should be of good quality.
- The system must have enough computation power to work the application.
- User should have all the software and library install to run this application.
- The validity of the algorithm should be verified against Udacity's Self Driving Car dataset.
- To get high accuracy we should have a large and standardize dataset.

Mathematical Modeling

- **Input:** Video
- **Output:** To detect vehicle, pedestrian and traffic sign
- **Algorithms:** Efficient det 0 Algorithm.

Mathematical Formulation:

- **System** = {Train, Test, classification}
- **Train** = {pre-process, feature extract, classification}
- **Test** = {pre-process, feature extract, classification}
- **Failure condition:** If the dataset is biased and bad training can result in reduced accuracy.

3.2 Functional Requirements

- The datasets include a set of images to be trained and then generate output in categorical form (1 of the 11 traffic scene classes) with help of using different algorithms and techniques.
- The dataset of images on which model is being trained to identify traffic scene classes from input provided. The dataset must be relevant to the input video; it should include real time traffic scene images.
- The software system should be able to perform these tasks efficiently.

3.3 Nonfunctional Requirements

Performance Requirement

- Security-The system should be highly secure as it deals with critical tasks of autonomous driving technology and should be reliable for the passengers of the vehicle.
- Maintainability-The system should be maintainable and should be provide required updates and fixes whenever necessary.
- Scalability-The system should also be highly scalable as it should be able to incorporate higher data processing capabilities and new and improved features over time.
- Reliability-It should be reliable as it deals with potentially fatal consequences for the passengers of the vehicle.

Software Quality Attributes

- Adaptability: This application is adaptable and simple to use.
- Availability: All users have free access to this application. Everyone can get the software because it is easily accessible.
- Maintainability: If any error occurs after the project has been deployed, the software developer can immediately fix it.

- Reliability: The software's performance has improved, increasing its stability and ensuring that it functions consistently.
- Testability: The software will be evaluated based on its results and performance.

3.4 System Requirements

Software Requirements (Platform Choice)

- Google Colab
- Tensorflow
- Keras
- Pandas
- Numpy

Hardware Requirements

- Processor : Intel core i3/i5 (10th or higher generation)
- Ram size : 8GB
- Hard disk capacity : 500 GB
- Monitor type : 15 Inch shading screen

Machine Specific: Intel i7+ processer series, Nvidia Quardo or AMD graphics above series with 4 GB graphics memory for best performance

V. ALGORITHM

5.1 Efficient-Det Algorithm

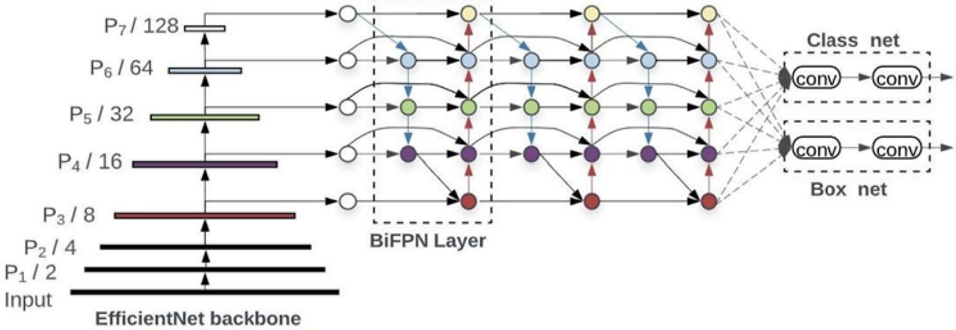
The Efficient-Det model was proposed object detection model by Google Brain team in CVPR 2020 based on Efficient-Net Architecture it was able to achieve state of the art accuracy among the currently available object detection models, The disadvantage of Region proposal methods like RCNN is that it could take up to 47 seconds to perform testing on an image and hence rendering it redundant for real time applications and Fast RCNN taking up to 2.3 seconds for the same, furthermore YOLO algorithm suffers from the drawback of not being able to detect smaller objects due to spatial limitations using which YOLO performs object detection, The Efficient-Det overcomes these limitations using Efficient-Net architecture, it is able to perform best in real time, in our implementation we were able to get the multiple object detection across 11 classes given in the dataset in 0.389 seconds per frame, which is extremely fast compared to Fast RCNN, The Efficient-Det model strikes a good balance between accuracy and computational resource constraints using fewer parameters and lesser number of FLOPs (Floating point operations per second) on the edge all the while being able to perform well in real time environment. The algorithm uses Image-Net pre-trained Efficient-Net architecture, the Bi-FPN (Bi-directional-Feature Pyramid Network) network works as a feature network on top of the Efficient-Net architecture.

5.2 Compound Scaling

To improve the efficiency and accuracy of the network we use compound scaling which is a aggregation of width, depth, baseline and resolution scaling methods into a single compound scaling method, introduced in Efficient-Net paper[]. The first step is to perform grid search to get relation between scaling dimensions of the baseline Efficient- Net network under resource limitations. Compound scaling improves efficiency and accuracy for the scaling up of the existing models like Mobile-Net and Image-Net as compared to the conventional scaling method.

5.3 Efficient-Det Architecture

Efficient-Det uses Efficient-Net backbone as feature extraction network as it reuses same features of Efficient-Net B0 to B6 ImageNet pretrained checkpoints, It uses Bi-FPN as a feature network which is a bi-directional feature network, Bi-FPN takes scales from 3 to 7 levels and do repeated fusion for a number of iterations, the higher scale levels detect larger objects with lower resolutions. These features are then given as an input to a class and box network to give object class and bounding box predicted bounds respectively, the class and box network parameters are shared across all feature levels. The box/class network of Efficient-Det uses softmax classifier to classify objects and uses bounding box regressors to predict bounding box bounds.

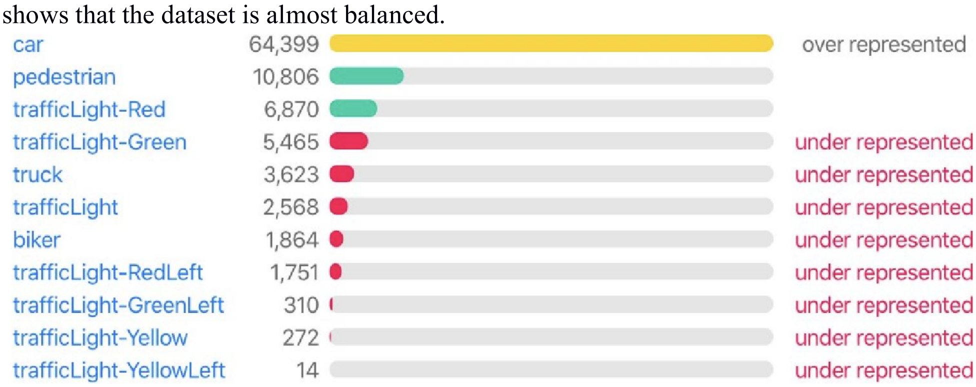


5.4 Kalman Filter

The Kalman Filter is an efficient optimal estimator (a set of mathematical equations) hence can be adapted as an object tracking algorithm that provides a recursive computational methodology for estimating the state of a data controlled process from measurements that are typically noisy, while providing an estimate of the uncertainty of the estimates. It performs recursive data processing and generates optimal estimate of desired quantities given the set of measurements. For linear system and white Gaussian errors, it "best" estimates the following states of a system based on all previous measurements. For non-linear system optimality is qualified, The algorithm doesn't need to store all previous measurements and reprocess all data each time step making it computationally economical and resource efficient and hence can be deployed in edge and real time systems. The algorithm serves three main objectives in our proposed implementation system which are first predict the detected object's next state that includes position within the image frame given the current state which has current object detections, second perform corrections on predictions and measurements and last reduce noise induced by inaccurate detections making the system more robust.

5.5 Dataset

Udacity self driving car dataset was built in the view to help developers working in the field of autonomous vehicles.It consists of about 11 classes and overall 15000 images . There are 97,942 labels into this dataset. Also there are about 1720 images which are without labels that is null images .The images are of the size 1920 x 1200 .This dataset has 97,942 labels around 11 classes and 15,000 images. There are about 1720 null Examples that means the images with unlabelled data.The 11 classes present in the dataset are namely Car , Pedestrian , Traffic Light-Red , Traffic Light-Green, Truck, Traffic Light, Biker, Traffic Light-Yellow , Traffic Light-RedLeft, Traffic Light- GreenLeft, Traffic Light-YellowLeft. The Class Car consists of about 64399 car images that is the reason the class balance slightly shifts towards class Car into this dataset. It is the most over represented class of the Udacity Self driving Car Dataset. The most under- represented class is Traffic Light-Yellow Left which consists of about 14 images.eed images present in the dataset which shows that the dataset is almost balanced.



VI. SYSTEM DESIGN

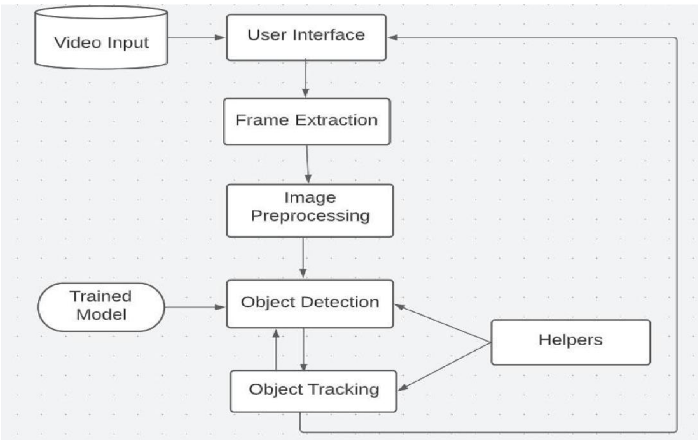


Fig 4.1 System Architecture

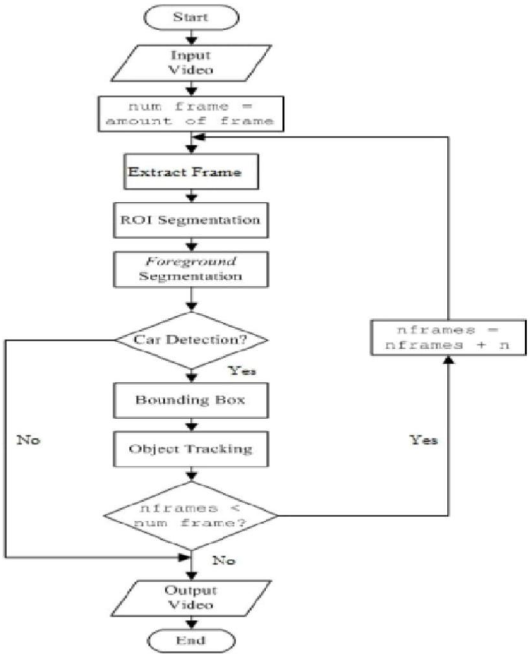


Fig 4.2 Data Flow Diagram

VII. CONCLUSION

Thus, we implemented object detection and tracking methods like Efficient-Det and Kalman filter respectively with training performed on Udacity's self driving dataset provided to us by Roboflow, The Efficient-Det object detection is a state of the art object detection algorithm proposed by Google Brain team based on Efficient-Net architecture, is best choice among currently available object detectors for real time and edge systems under computational resource constraint, the Kalman filter along with Efficient-Det serves as a robust object recognition system that can be used for experimental purposes in vehicle automation system.

REFERENCES

[1]. M. Everingham, et al, The PASCAL Visual Object Classes Challenge: A Retrospective, International Journal of Computer Vision, 2020.

[2]. O. Russakovsky et al, ImageNet Large Scale Visual Recognition Challenge, International Journal of Computer Vision volume, 2019

[3]. C. Crispim-Junior, et al, Semantic event fusion of different visual modality concepts for activity recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence PP (99) (2020).

[4]. Bolc Leonard, "Computer Vision and Graphics",2018 IEEE International Conference, Warsaw, Poland.

[5]. P. Sermanet, Y. LeCun, Traffic sign recognition with multiscale Convolutional Networks in, The 2020 International Joint Conference on Neural Networks (IJCNN), , IEEE, San Jose, CA , USA.

[6]. https://public.roboflow.com/object-detection/self-driving-car [8]https://arxiv.org/abs/1512.02325

[7]. https://towardsdatascience.com/ssd-single-shot-detector-for-object-detection-using- multibox-1818603644ca

[8]. https://www.sciencedirect.com/topics/earth-and-planetary-sciences/kalman-filter

[9]. https://arxiv.org/abs/1911.09070 [12]https://ai.googleblog.com/2020/04/efficientdet-towards-scalable-and.html

[10]. https://ai.googleblog.com/2019/05/efficientnet-improving-accuracy-and.html

[11]. https://medium.com/visionwizard/efficientdet-a-paper-review-21918d9a648d

[12]. https://arxiv.org/pdf/1911.09070.pdf Wang, Lang, et al. "Fast Pedestrian Detection And Tracking Based On Vibe Combined Hog-Svm Scheme". International Journal of Innovative Computing, Information and Control, 2019.

[13]. Dalal, Navneet, and Bill Triggs. "Histograms of Oriented Gradients for Human Detection".International Conference on Computer Vision & Pattern Recognition,2010.

[14]. Iraei, Iman, and Karim Faez. Object Tracking with Occlusion Handling Using Mean Shift,Kalman Filter and Edge Histogram. 2nd International Conference on Pattern Recognition and Image Analysis (IPRIA), 2015.

[15]. D, Nandashri, and Smitha P. An Efficient Tracking of Multi Object Visual Motion usingHungarian Method. International Journal of Engineering Research & Technology (IJERT), 2015.

[16]. Singh, Karan, and Nikita Malik. "CNN Based Approach for Traffic Sign Recognition System."Advanced Journal of Graduate Research, vol. 11, no. 1, 2022.

[17]. Zaibi, Ameur, et al. "A Lightweight Model for Traffic Sign Classification Based on EnhancedLeNet-5 Network." Hindawi Journal of Sensors, vol. 2021, p. 13, paghttps://doi.org/10.1155/2021/8870529.

[18]. Youssouf, Njayou. "Traffic Sign Detection and Recognition with FasterRCNN and YOLOV4."Heliyon, 2022.

[19]. Wang, Lang, et al. "FAST PEDESTRIAN DETECTION AND TRACKING BASED ON VIBE COMBINED HOG-SVM SCHEME." International Journal of Innovative Computing, Information and Control, vol. 15, Dec,2019.

[20]. S, Geethapriya, and P. Kumar. "An Improved Pedestrian Detection Algorithm using Integration of Resnet and Yolo V2." International Journal of Recent Technology and Engineering, vol. 9, no. 1, May,2020.

[21]. Chen, Linkai, et al. "An algorithm for highway vehicle detection based on convolutional neural network." EURASIP Journal on Image and Video Processing, 2018.

[22]. Song, Huansheng, et al. "Vision-based vehicle detection and counting system" , 2019.

BIOGRAPHY

- Prof. B. N. Babar: Professor in the department of Information Technology in Sinhgad Institute of Technology and contributed as guide and author of the Project.
- Kunal Bhokare: An Undergraduate Scholar pursuing Bachelors of Engineering in Information Technology from Sinhgad Institute of Technology. He is working under the guidance of Prof. B. N. Babar.
- Vijay Maholkar: An Undergraduate Scholar pursuing Bachelors of Engineering in Information Technology from Sinhgad Institute of Technology. He is working under the guidance of Prof. B. N. Babar.
- Gaurav Takle: An Undergraduate Scholar pursuing Bachelors of Engineering in Information Technology from Sinhgad Institute of Technology. He is working under the guidance of Prof. B. N. Babar.