© June 2022 | IJIRT | Volume 9 Issue 1 | ISSN: 2349-6002

Enhanced Object Detection with Deep CNN for Advance Driving Assistance

SYED TALHA AMAAN[1], S UDAYA HINDU[2], RAHULKNAIK BARMAVATH[3], DR. SREEDHAR BHUKYA[4]

[1, 2, 3] *Student, Dept of Computer Science and Engineering Sreenidhi Institute of Science and Technology*

[4] *Professor, Dept of Computer Science and Engineering Sreenidhi Institute of Science and Technology*

Abstract— The intelligent vehicle, as a crucial technology in the intelligent transportation system, is the bearer of a complete integration of several technologies. Although vision-based autonomous driving has showed great promise, there is still the issue of analysing the complex traffic scenario using the data acquired. Recently, self driving has been broken down into many tasks utilising various models, such as object detection and intention identification. In this work, a vision-based system was built to recognise and identify numerous items in the traffic scene, as well as forecast pedestrian intentions. The key contributions of this study are: (1) A fine-tuned Part-Affinity-Field's approach to know pedestrian pose was proposed; (3) Explainable-AI (XAI) technology was used to explain and assist the estimation results in the risk assessment phase; and (4) a large autonomous driving dataset with several subsets for each task was brought in; and (5) an endtoend system having multiple models with high accuracy was developed. The overall parameters of the modified Faster RCNN were decreased by 74%, demonstrating that it meets the real-time capabilities. Furthermore, when compared to the state-of-the-art, the detection precision of the enhanced Faster RCNN improved by 2.6 percent.

Indexed Terms— Self-driving, Autonomous, Explainable-AI, Convolutional neural networks

I. INTRODUCTION

Rapid urbanisation has revealed a slew of problems, particularly in the area of transportation, which severely restricts travel and poses security risks. Even though existing self-driving object detection technologies have progressed, there are still potential collision risk factors because automobiles are surrounded by a variety of objects in everyday life, including some uncontrollable moving objects (pedestrians and vehicles) as well as static objects (traffic lights and signs). As a result, it's necessary to swiftly recognise a variety of static elements and accurately analyse the intents of moving objects.

The two types of deep learning methodologies utilised in object detection tasks are one-stage detection algorithms and two-stage detection algorithms. One-stage detection methods like YOLO [1] and SSD [20] convert the detection problem into a unified regression problem right away. Because of the structure's qualities, one-stage procedures are quicker than two-stage methods. Faster R-CNN [3] is a two-stage network that builds a series of candidate bounding boxes before classifying each item using a Convolutional Neural Network (CNN). The two-stage approaches outperform the majority of one-stage methods in terms of detection and localization precision. The suggested model with numerous tasks in this work is built on one-stage procedures to decrease the time spent on object detection.

II. METHODS

A. Existing Method

Even if present object recognition technologies in self-driving cars have progressed, there are still possible accident risk factors since motor cars are surrounded by numerous items in daily life, including some uncontrollable moving objects (pedestrians and automobiles) and static objects (traffic lights and signs).

Despite the rapid advancement of CNN in object recognition over datasets with a high number of object classes, real-time visual object detection in a driving scenario remains a difficult task. The popular CNN detectors perform poorly on the benchmark datasets when it comes to object detection.

- Feature maps from feature output scales are processed individually in current multi-scale CNN models to predict the existence of objects at fixed scales.
- The non-maximal suppression (NMS) approach is used to suppress overlapping object suggestions in the majority of current CNN detectors. There is virtually little probability of properly detecting obstructed items with such a procedure. Occluded objects, on the other hand, are common in driving settings and can be dangerous.

Default anchor boxes of various sizes are used to produce object suggestions in current CNN detectors. In a driving environment, the items of interest have distinct shapes; for example, the width of a car should not exceed the width of a lane.

B. Proposed Method

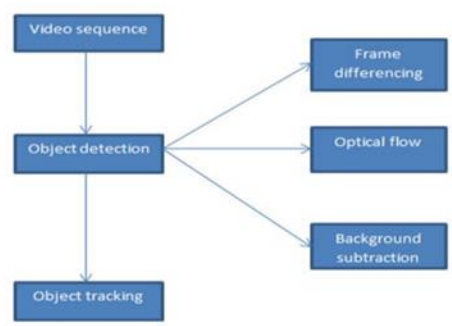
We show in this study that combining an algorithmic change computing proposal with a deep convolutional neural network results in an elegant and practical solution where proposal computation is almost cost-free when compared to the detection network's calculation. To achieve this, we present innovative Region Proposal Networks (RPNs), which share convolutional layers with state-of-the-art object identification networks. The marginal cost of calculating suggestions is low since convolutions are shared at test time (e.g., 10ms per image). The convolutional feature maps employed by region-based detectors, such as Fast R-CNN, may also be used to generate region proposals, according to our findings. We build an RPN on top of these neural features by layering on a few more convolutional layers that regress region limits and objectness scores at each position on a regular grid. As a result, the RPN is a kind of fully convolutional network (FCN) that can be trained end-to-end for the purpose of producing detection recommendations.

C. Our Enhancement

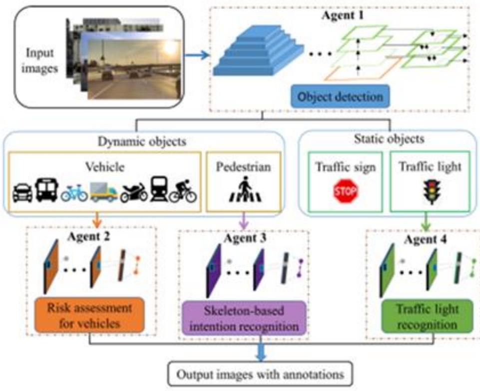
Faster R-CNN, our object detection system, is made up of two parts. A deep fully convolutional network suggests areas in the first module, and the Fast R-CNN detector employs the suggested regions in the second module. Using the increasingly popular nomenclature of neural networks with 'attention' processes, the whole system is a single, unified network for object identification; the RPN module directs the Fast R-CNN module where to look.

- At lower feature output scales, deep convolution of CNN features is performed, which is then merged with features at bigger feature output scales to offer richer context for object recognition at the individual feature output scale. This type of upgrade can successfully handle the problem of big object size variation.
- Soft-NMS is applied to object suggestions from different feature output scales to strike a compromise on the amount and quality of object proposals in order to handle the object occlusion difficulty.
- Anchor box settings might be based on the object aspect ratio distributions. We use the statistics to establish optimal anchor box settings for better item localization and prediction by measuring the aspect ratio statistics of objects from training samples.

III. DATA FLOW DIAGRAM



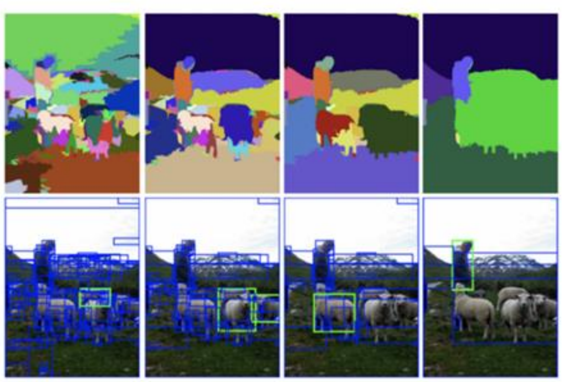
IV. SYSTEM FLOWCHART



V. ALGORITHM USED

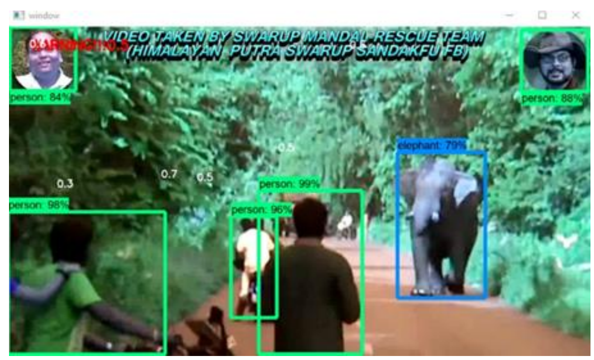
A. Faster-RCNN

The initial stage towards Faster R-CNN is R-CNN (R. Girshick et al., 2014). It finds regions of interest using search selection (J.R.R. Uijlings & al. (2012)) and feeds them to a ConvNet. By grouping comparable pixels and textures into numerous rectangular boxes, it tries to discover areas that could constitute an item. 2,000 recommended regions (rectangular boxes) from search selection are used in the R-CNN article. The 2,000 locations are then fed into a CNN model that has already been trained. Finally, the outputs (feature maps) are classified using an SVM. The regression between predicted and ground-truth bounding boxes (bboxes) is calculated.



B. Deployment

Our System is deployed using Flask. Flask is a python library which is relatively easy and lightweight web framework. It Provides essential tools and so many capabilities for building a web platform. Using this we can easily build web platforms just by using single python file so in that way it's beginner friendly for developers and developers can create web platforms rapidly to test out their model and make it an end-to-end system.



VI. CONCLUSION

A vision-based object detection and recognition framework for autonomous driving was presented in this study. One item detection task and three recognition tasks are included in the proposed system. Various objects are recognised using an improved Faster RCNN algorithm model with fewer parameters, which achieves greater detection accuracy and processing speed than the original. Vehicles, pedestrians, and traffic signals are incredibly significant things in the self-driving subject for recognised objects. As a consequence, the connected objects are divided into three categories. By comparing it to other CNN models, the best suited model with the highest accuracy is picked for each recognition assignment. The RISE algorithm is also used to explain the classification results by constructing saliency maps for each image. In the future, more effort should be made on improving the overall pace of the recommended framework. To improve the system's performance, a separate pipeline that can efficiently handle single-frame-based and multi-frame-based recognition will be implemented in the next study. Given the importance of the distance between autonomous cars and other objects, distance prediction should be included in this framework.

REFERENCES

[1] Wei Liu and Alexander C. Berg, “SSD: Single Shot MultiBox Detector”, Google Inc., Dec 2016.
[2] Andrew G. Howard, and Hartwig Adam, “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications”, Google Inc., 17 Apr 2017.
[3] Justin Lai, Sydney Maples, “Ammunition Detection: Developing a Real Time Gun Detection Classifier”, Stanford University, Feb 2017
[4] ShreyamshKamate, “UAV: Application of Object Detection and Tracking Techniques for Unmanned Aerial Vehicles”, Texas A&M University, 2015.
[5] Adrian Rosebrock, “Object detection with deep learning and OpenCV”, pyimagesearch.
[6] Mohana and H. V. R. Aradhya, "Elegant and efficient algorithms for real time object detection, counting and classification for video surveillance applications from single fixed camera," 2016 International Conference on Circuits, Controls, Communications and Computing (I4C), Bangalore, 2016, pp. 1-7.
[7] AkshayMangawati, Mohana, Mohammed Leesan, H. V. Ravish Aradhya, “Object Tracking Algorithms for video surveillance applications” International conference on communication and signal processing (ICCSP), India, 2018, pp. 0676-0680.
[8] Apoorva Raghunandan, Mohana, Pakala Raghav and H. V. Ravish Aradhya, “Object Detection Algorithms for video surveillance applications” International conference on communication and signal processing (ICCSP), India, 2018, pp. 0570-0575.
[9] Manjunath Jogin, Mohana, “Feature extraction using Convolution Neural Networks (CNN) and Deep Learning” 2018 IEEE International Conference On Recent Trends In Electronics Information Communication Technology, (RTEICT) 2018, India.
[10] Arka Prava Jana, Abhiraj Biswas, Mohana, “YOLO based Detection and aClassification of Objects in video records” 2018 IEEE International Conference On Recent Trends In Electronics Information Communication Technology, (RTEICT) 2018, India.