Design And Implementation of Object Detection Algorithm For Autonomous Vehicle

Miss.Shalaka Gautam Kamble [1], Miss. Shweta Patil[2]
[2] Assistant Professor
[1, 2] Walchand College of Engineering, Sangli

Abstract- *An autonomous vehicle is a vehicle that guides itself without human conduction and takes decisions concerning for surrounding environment. These days vehicles are equipped with a lot of sensors, actuators, and controllers. Real-time object detection for vehicle detection is a key component in autonomous driving systems. Detecting objects in a video stream frame-by-frame is time-consuming because it is difficult to avoid a collision. For that purpose internet of things combined with different technologies like machine learning, artificial intelligence, local computing provides essential technologies for self-sufficient driving frameworks. Such deep learning is a machine learning technique and it teaches a computer to filter input data through layers to learn how to predict and classify data. For that open CV, computer vision and AI programming library is utilized to create new application. The advancements in pc Vision with Deep Learning has been made and updated with time, over one specific algorithmic program and its Convolutional Neural Network (CNN).*

Keywords- Autonomous Vehicle, Computer Vision, Convolutional Neural Network, deep learning, Object detection.

I. INTRODUCTION

Driverless vehicles are one of the first encouraging possibilities for figuring computing examination; it'd be the best innovative transformation on the very edge of things has come. There are various innovations which are creating an association of independent vehicle with global positioning system, image processing, Open CV, LASER, machine vision, RADAR etc. Contrasted and various sensors like Pi-camera are lower-cost and may give less information out and about (traffic signs, stoplight, passerby, hindrances). Hence, camera like LIDAR is used mostly during the manufacturing of the autonomous vehicle. In any case, the principal drawback of the vision-based approach is that it isn't strong to illumination changes. Identification of object in indoor scenes is the main innovation that machine can finish the assignment of delivery and fetching. Recognition of target is one of the troublesome computer vision undertakings that the foundation of the indoor scenes are progressively unpredictable, the fluctuation of daylight changes and targets are put in numerous postures. Also, the profound learning-based object locator generally requires an outsized number of tags in preparing trained data and interpretation is tedious. Especially medical, optical and remote-sensing images are those images where requirement of professional to annotate is necessary. So object identification is harder in an inside scene with some samples. Hence use of deep learning method is very important for object detection. Profound machine learning is a man-made knowledge that copies the activities of the human mind in the procedure of data training and making samples to be utilized in the higher psychological procedures. This learning strategy could have a lot of AI in computerized reasoning that has appropriate systems for taking in disregarded from data that is untagged. CNN could even be a class of AI, feed-forward neural system that has been utilized to offer great execution in PC vision undertakings, similar to picture order and recognition. CNN is a system of more profound layers with its loads, inclinations, and yields that are utilized for nonlinear actuation. The neurons of the CNN are sorted out during a volumetrically style like tallness, measurement, and profundity.

II. LITERATURE REVIEW

Aliasgar Haji, et al. [1] suggested a way of CNN to acknowledge the objects from image or stream of images. Behavioral cloning may be a method by which sub-cognitive skills like recognizing objects while operating an action which will be acquired and recreated during a computer vision processing. Also while driving the remote controlled car manually the labeled images are captured and used for training the model. The images captured are mapped with the commands of left, right or straight direction.For this, Raspberry Pi uses general purpose input output packages for the remotely controlling of RC car.This makes feasible for the car to form appropriate decisions when driven within the autonomous mode. After testing the Self Driving car in the similar driving conditions the proposed model has given a test accuracy of 84.5%. To extend the data set and improve the model, use of data augmentation was necessary. The model will fail, if there are moving obstacles on the trail or drastic change within the road structure. An application of sensor

fusion techniques, localization, and control features are used to enhance the autonomous behavior of the vehicle.

Reagan L. Galvez, et al. [2] introduced a vision system which was important in manufacturing a portable robot which will finish a specific assignment like surveillance, explosive ordnance disposal and navigation. Such technique was used by robot to control or to operator it within the environment performing subsequent events. In processing of image, categorize and indentifying the particular target was attainable because of the recent development in a deep neural network. To detect objects within the environment mostly CNN was used. Two best in class models were considered for object recognition, one was faster region-based CNN with inceptionv2 and another was single shot multi-box detector with mobilenetv1. The outcome shows that one model was helpful for real-time application on account of processing time and thus second model was often used for object detection.

Object tracking could also be a key advance in PC vision for video reconnaissance, open security, and traffic examination. Two connected parts of video reconnaissance were discovering object and tracking. Target location in recordings is the primary event before playing out an entangled assignment likes tracing. Neural networks type deep learning may be a useful programming in real-time applications. This helps to learn various stages of presentation and reflection of knowledge like pictures, sound, and content. Gaussian mixture model based target localization, deep learning neural network-based identification and tracing of it utilizing a correlation filter was presented by Supreeth HSG, et al. [3] for detecting objects. This method can improve the efficiency by handling false detection. The algorithm was developed to recognize only vehicles and pedestrians when the working was examined with help of false alarm and true positive rate as metrics probability. The test aftereffects of the developed technique were seen as better with an exactness of 88%.

A monocular vision-based autonomous vehicle on raspberry pi operated with deep neural networks was given by Truong-Dong Do, et al. [4]. The main technique behind these was its capacity to comprehend large number of parameters making use of tagged information. The author demonstrated a deep neural network to make a architecture that legitimately maps unprocessed information to an anticipated controlling point which would have been yield. They used the information gathered from the structure of car built with scale of one tenth radio control (RC) vehicle, front-facing camera raspberry pi 3 model B computer and for next process. This information then utilized to train various parameters of the CNN model. The trained information was street pictures joined with the time-synchronized directing point shaped by physically operating. The experimental results show viability and power of the self-sufficient of the autonomous structure in path following events. The problem observed during preparing and checking the system was latency of camera. This camera latency was the period which was the time between the camera sensors gathers the scene to the machine peruses the digitized pattern. Hence, based the working of the raspberry pi and camera this period could be long which was not good for the system.

Celi, et al. [5] discovered an object detection technique depending on deep learning of a few specimens that consists of structures like target identification network, semantic-applicable discovery and synthetic samples generator. Algorithm given produces manufactured pictures and equivalent tags utilizing the manufactured samples generator to understand the issues of certain examples automatically. The architecture like deep supervision is utilized in the detection of target in the network to solve the issue of vanishing-gradient generated by deeply supervised object detector. The objects with different scale were detected by a multi-layer prediction structure. Berkeley 3D Object Dataset (B3DO) was a 3D image dataset that was gathered in real domestic. This contains 848 color images and their related depth images. The author compared the proposed approach with other techniques. First, were you Only Look Once and another was DSOD depending on CNNs that had given relevant pattern analysis, statistical was modeling and Computational Learning (PASCAL). This resulted in getting the highest detection accuracy. The mean normal accuracy was about 18*10^-2 more as compared to DSOD and 0.25 higher than the YOLO method.

Jihun Kim, et al. [6] has used vision and RADAR sensors to upgrade the target detection precision in cars while driving on road. The data related to the RADAR sensor and the vision sensors were shown in various reference framework. Here the author applied the method called pseudo inverse based on point alignment. This type gives specific information regarding the coordinates of the radar object in the image plane. Each point on image plane through the transformation matrix shows the values of reference framework for each sensor. The linear least square (LS) methodology was utilized to calculate the matrix transformation, where a minimum of 4 resulted information's was observed. The author proposed a model, which will give target information expressed by one point in the form of pixel value. Since the concentration was on car or any target, it should represent it not a single dot only but as an area,. This could be reliable if the calculated accuracy was about 95%.

Ankith Manjunath, et al. [7] introduced the task and measurement area in the radar supplication. Depending on this

item recognition using radar and tracing issues was mathematically calculated. Bayesian filter was used as a filter for radar-based applications. Kalman filter was one of the types of it predominantly utilized in tracking. It has two main parts in car applications; first was vehicles were driven non-directly and Constant Turn Rate and Velocity (CTRV) movement model gives the development of the vehicle being followed. The Cartesian kinematic state space was converted into the polar measurement space which represented the transformation was non-linear. To check the non-linearity in the motion and the measurement models, Extended Kalman Filter (EKF) and Unscented Kalman Filter (UKF) can be implemented. The author checked the non-linearity in the movement and then estimation models, i.e. Extended Kalman Filter (EKF) and Unscented Kalman Filter (UKF) can be executed.

The deep CNN for real-time application in detection with classification of on highway road the target was operated by Seyyed Hamed Naghavi, et al. [8]. This network was trained on the KITTI road dataset and identification of on street target consisting of cars, cycles and people were performed. The final system of CNN forms 448x448 information pictures at 47 edges for each second on an NVIDIA GeForce GTX960 GPU. Here working boost of about 5.2% mAP over SSD300 has been accomplished on the KITTI dataset by including the number of highlight maps after the base system and remaking the organization of their framework cell. Better worldwide thinking and less bogus positives than disjoint models like Faster R-CNN have been accomplished with the assistance of a solitary shot discovery approach. This model arrived at 78.4% mAP on the KITTI dataset, which was 11.9% more than the straightforward YOLO model and 5.2% higher than the SSD300 model.

III. PROPOSED MODEL

Before identifying any object or facial identity from input images or real-time data, it is first essential to detect and extract the target region from the original picture frames. There are several extracting algorithms for these sectors of image processing. Also the collection of dataset needs a pre-processing to normalize the dimensions for better training afterwards that result in improved accuracy. Below resides the basic workflow that we are maintaining for object detection,

Video Acquisition: In this video camera is employed for electronic picture acquisition (as hostile a motion-picture, camera which captures images on film) firstly created for the television but now common in other real time applications as well.

Fig.1 Block daigram of proposed system

Pre-processing: The main target of pre-processing is corresponding development of the image information that restrains unwanted distortions or enhances some image options vital for more processes.

Feature Selection: It is the method where you mechanically or manually choose those options that contribute most to your prediction variable or output within which you are curious about. Having non-related properties can decrease the accuracy of models and build your model so that it can also understands non-related properties.

Feature Extraction: It is a necessary processing part in target recognition and CNN methods. Hence purpose is to extract a collection of options from the dataset of interest. These features must be informatory with respect to the required properties of the original data.

Object Identification: Object detection appertains the capability of computer and software systems to identify objects in an image or scene. This has been utilized for face detection, vehicle detection, pedestrian numeration, web images, security systems and driverless cars.

IV. CONCLUSION

In this paper the review of various approaches is presented to highlight the work done to detect the object using CNN. Later on critical evaluation is done on the basis of technique used, challenges and limitation. From the study it is very clear that the system can show some problems and needs proper solution for it. These problems are low resolution, occlusion and clutter with various levels of changes on R-CNN. This review begins with the basic concept of detecting object using CNN which provides basic architecture for proposed system and solution to the problems. Hence it is also useful for the reconstruction in CNN which gives beneficial perception and suggestions for future progress.

REFERENCES

[1] Aliasgar Haji, Priyam Shah, Srinivas Bijoor ,“Self Driving RC Car using Behavioral Cloning”,Computer

Vision and Pattern Recognition; Systems and Control , 10 Oct 2019.

[2] Reagan L. Galvez ; Argel A. Bandala ; Elmer P. Dadios ; Ryan Rhay P. Vicerra ; Jose Martin Z. Maningo, "Object Detection Using Convolutional Neural Networks", TENCON 2018 - 2018 IEEE Region 10 Conference, Jeju, Korea (South), 2018.

[3] H. S. G. Supreeth and C. M. Patil, "Moving object detection and tracking using deep learning neural network and correlation filter," 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), Coimbatore, 2018.

[4] T. Do, M. Duong, Q. Dang and M. Le, "Real-Time Self-Driving Car Navigation Using Deep Neural Network," 4th International Conference on Green Technology and Sustainable Development (GTSD), 2018.

[5] C. Li, Y. Zhang and Y. Qu, "Object detection based on deep learning of small samples," 2018 Tenth International Conference on Advanced Computational Intelligence (ICACI).

[6] J. Kim, D. S. Han and B. Senouci, "Radar and Vision Sensor Fusion for Object Detection in Autonomous Vehicle Surroundings," Tenth International Conference on Ubiquitous and Future Networks (ICUFN), 2018.

[7] A. Manjunath, Y. Liu, B. Henriques and A. Engstle, "Radar Based Object Detection and Tracking for Autonomous Driving," 2018 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM).

[8] S. H. Naghavi and H. Pourreza, "Real-Time Object Detection and Classification for Autonomous Driving," 8th International Conference on Computer and Knowledge Engineering (ICCKE), 2018.

[9] P. F. Felzenszwalb, R. B. Girshick, D. Mcallester, and D. Ramanan, "Object detection with discriminatively trained part-based models," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 9, p. 1627, 2010.

[10] C. Chen, A. Seff, A. L. Kornhauser, and J. Xiao, "Deep driving: Learning affordance for direct perception in autonomous driving," in ICCV, 2015.

[11] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in CVPR, 2016.

[12] A. Stuhlsatz, J. Lippel, and T. Zielke, "Feature extraction with deep neural networks by a generalized discriminant analysis."IEEE Trans. Neural Netw.& Learning Syst., vol. 23, no. 4, pp. 596–608, 2012.

[13] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, pp. 1627–1645, 2010.

[14] I. Croitoru, S.-V. Bogolin, and M. Leordeanu, "Unsupervised learning from video to detect foreground objects in single images," in ICCV, 2017.

[15] K. Kang, H. Li, T. Xiao, W. Ouyang, J. Yan, X. Liu, and X. Wang, "Object detection in videos with tubelet proposal networks," in CVPR, 2017.