International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 03 | March - 2023 Impact Factor: 7.185 ISSN: 2582-3930

DNN BASED OBJECT RECOGNITION SYSTEM IN VEHICLE USING IMAGE PROCESSING: A REVIEW

Aditi Soni[1], Mrs. Nidhi Sindhu[2]
Mtech Scholar[1], Assistant Professor[2]
Dept Of Computer Science And Engineering,CSIT,DURG

Abstract:

Object detection and tracking is one of the critical areas of research due to routine change in motion of object and variation in scene size, occlusions, appearance variations, and ego-motion and illumination changes. Specifically, feature selection is the vital role in object tracking. It is related to many real time applications like vehicle perception, video surveillance and so on. In order to overcome the issue of detection, tracking related to object movement and appearance. Most of the algorithm focuses on the tracking algorithm to smoothen the video sequence. On the other hand, few methods use the prior available information about object shape, color, texture and so on. Tracking algorithm which combines above stated parameters of objects is discussed and analyzed in this research. The goal of this paper is to analyze and review the previous approach towards object tracking and detection using video sequences through different phases. Also, identify the gap and suggest a new approach to improve the tracking of object over video frame.

Keywords: Object tracking, object recognition, statistical analysis, object detection, background subtraction, performance analysis, optical flow

I. INTRODUCTION

Recently, there is an advance of miniaturization and lower the cost of cameras have preferred the implementation of large-scale networks of the camera. This increasing numberof cameras could permit novel signal processing applications which employ multiple sensors in extensive areas. Object tracking is the novel procedure for discovering moving objects beyond time by utilizing the camera in videosequences (Kothiya and Mistree, 2015). Their main aim is to relate the target objects as well as the shape or features, location of the objects in successive video sequences. Subsequently, theobject classification and detection are essential for object tracking in computer vision application. Additionally, the tracking is the first step towards locating or detects the moving object in the frame. Followed by this, detected object could be divided as swaying tree, birds, human, and vehicles and so on. Though, in image processing approach object tracking using video sequences, is a challenging task. Furthermore, several issues appear ascribed to occlusion of the object to scene, object to object, complex object motion, real-time processing requirements as well as the improper shape of the object. However, this tracking has a large number of benefits, few of them are traffic monitoring, robot vision, surveillance and security and video communication, public areas like underground stations, airports, mass events and animation (Kim, 2007; Lowe, 2004; Ojha and Sakhare, 2015; Yilmaz et al., 2006). Thus, the particular application needs optimal trade-off among computing, communication, and accuracy over the network. The revenue related to computing and communication relies on the amount and type of cooperation executed among cameras for data collection, dispensing and processing to confirm decisions and to reduce the estimation errors and ambivalence.

The capability of machines to identify the suspicious object and further identify their activities in a specific environment is an important part of permitting a machine to interact with humans in effective and easy manner. The current approach for analyzing and detecting the suspicious object usually needs exceptional markers

connected to the suspicious object that prevents the extensive technology application. In this paper, to study as wll

as analyze the previous approach towards object tracking using video sequences through different phases.

Three key steps in video analysis are discussed as follows:

1. Identification of targeted object in moving sequence.
2. Object tracking based on one frameto another frame.
3. Tracking of the object from camerato camera.

II. Literature Survey:

Bhumika Gupta (2017) et a;. proposed object detection is a well-known computer technology connected with computer vision and image processing that focuses on detecting objects or its instances of a certain class (such as humans, flowers, animals) in digital images and videos. There are various applications of object detectionthat have been well researched including face detection, character recognition, and vehicle calculator. Object detection can be used for various purposes including retrieval and surveillance. In this study, various basic concepts used in object detection while making use of OpenCV library of python 2.7, improving theefficiency and accuracy of object detection are presented.

Kartik Umesh Sharma (2017) et al, proposed an object detection system finds objects of the real world present either in a digital image or a

video, where the object can belong to any class of objects namely humans, cars, etc. In order to detect an object in an image or a video the system needs to have a few components in order to complete the task of detecting an object, they are a model database, a feature detector, a hypothesiser and a hypothesiser verifier. This paper presents a review of the various techniques that are used to detect an object, localise an object, categorise an object, extract features, appearance information, and many more, in images and videos. The comments are drawn based on the studied literature and key issues are also identified relevant to the object detection. Information about the source codes and online datasets is provided to facilitate the new researcher in object detection area. An idea about the possible solution for the multi class object detection is also presented. This paper is suitable for the researchers who are the beginners in this domain.

Mukesh Tiwari (2017) et al. presented object detection and tracking is one of the critical areas of research due to routine change in motion of object and variation in scene size, occlusions, appearance variations, and ego-motion and illumination changes. Specifically, feature selection is the vital role in object tracking. It is related to many real time applications like vehicle perception, video surveillance and so on. In order to overcome the issue of detection, tracking related to object movement and appearance. Most of the algorithm focuses on the tracking algorithm to smoothen the video sequence. On the other hand, few methods use the prior available information about object shape, color, texture and so on. Tracking algorithm which combines above stated parameters of objects is discussed and analyzed in this research. The goal of this paper is to analyze and review the previous approach towards object tracking and detection using video sequences through different phases. Also, identify the gap and suggest a new approach to improve the tracking of object over video frame.

Aishwarya Sarkale (2018) et al. proposed humans have a great capability to distinguish objects by their vision. But, for machines object detection is an issue. Thus, Neural Networks have been introduced in the field of computer science. Neural Networks are also called as 'Artificial Neural Networks'. Artificial Neural Networks are computational models of the brain which helps in object detection and recognition. This paper describes and demonstrates the different types of Neural Networks such as ANN, KNN, FASTER R-CNN, 3D-CNN, RNN etc.
with their accuracies. From the study of various research papers, the accuracies of different

Neural Networks are discussed and compared and it can be concluded that in the given test cases, the ANN gives the best accuracy for the object detection.

Karanbir Chahal (2018) et al. proposed Object detection is the identification of an object in the image along with its localization and classification. It has wide spread applications and is a critical component for vision based software systems. This paper seeks to perform a rigorous survey of modern object detection algorithms that use deep learning. As part of the survey, the topics explored include various algorithms, quality metrics, speed/size trade-offs and training methodologies. This paper focuses on the two types of object detection algorithms- the SSD class of single step detectors and the Faster R- CNN class of two step detectors. Techniques to construct detectors that are portable and fast on low powered devices are also addressed by exploring new light weight convolutional base architectures. Ultimately, a rigorous review of the strengths and weaknesses of each detector leads us to the present state of the art.

Richard Socher (2018) et al. proposed recent advances in 3D sensing technologies make it possible to easily record color and depth images which together can improve object recognition. Most current methods rely on

very well-designed features for this new 3D modality. We introduce a model based on a combination of convolutional and recursive neural networks (CNN and RNN) for learning features and classifying RGB-D images. The CNN layer learns low-level translationally invariant features which are then given as inputs to multiple, fixed-tree RNNs in order to compose higher order features. RNN scan be seen as combining convolution and pooling into one efficient, hierarchical operation. Our main result is that even RNNs with random weights compose powerful features. Our model obtains state of the art performance on a standard RGB-D object data set while being more accurate and faster during training and testing than comparable architectures such as two-layer CNNs.

Yordanka Karayaneva (2018) et al. presented schools in many parts of the world use robots as social peers in order to interact with children and young students for a rich experience. Such use has shown significant enhancement of children's learning. This project uses the humanoid robot NAO which provides object recognition of colours, shapes, typed words, and handwritten digits and operators. The recognition of typedwords provides performance of the corresponding movements in the sign language. Five classifiers including neural networks are used for the handwritten recognition of digits and operators. The accuracy of the object recognition algorithms are within the range of 82%-92% when tested on images captured by the robot including the movements which represent words in the sign language. The five classifiers for handwritten recognition produce highly accurate results which are within the range of 87%-98%. This project will serve as a promising provision for an affective touch for children and young students.

Abdul Muhsin M (2019) et al. proposed everybody deserve to live independently, especially those who disabled, with the last decades, technology gives attention to disabled to make them control their life as possible. In this work, assistive system for blind is suggested, to let him knows what is around him, by using YOLO for detecting objects within images and video streams quickly based on deep neural network to make accurate detection, and OpenCV under Python using Raspberry Pi3. The obtained results indicated the success of the proposed model in giving blind users the capability to move around in unfamiliar indoor outdoor environment, through a user friendly device by person and object identification model.

Geethapriya. S (2019) et al. proposed the Objective is to detect of objects using You Only Look Once (YOLO) approach. This method has several advantages as compared to other object detection algorithms. In other algorithms like Convolutional Neural Network, Fast Convolutional Neural Network the algorithm will

not look at the image completely but in YOLO the algorithm looks the image completely by predicting the bounding boxes using convolutional network and the class probabilities for these boxes and detects the image faster as compared to other algorithms.

R. Sujeetha (2019) et al. proposed object detection and tracking could be a immense, vivacious however inconclusive and trending area of computer vision. Due to its immense use in official surveillances, tracking modules applied in security and lots of others applications have made researchers to devise a lot of optimized and specialized methods. However, problems are faced in implementing object detection and tracking in real-time; like tracking in real time and giving appropriate optimized results, over dynamic computation to find the efficient performance with respect to time factor, or multiple objects tracking create this task more difficult. Though, several techniques are devised but still lies a lot of scope of improvement, however during this literature review we've seen some illustrious and multiple ways of object detection and

tracking. In this method we will be using Tensor Flow and Open CV library and CNN algorithm will be used and we will be labelling the detected layers with accuracy being checked at the same time .For validation purpose live input video will be taken for the same where objects will be getting detected and it can besimulated same for real-time through external hardware added. In the end we see the proper optimized and efficient algorithm for object tracking and detection.

III. OBJECT DETECTION AND TRACKING PROPOSEDMETHOD:

a) OpenCV

OpenCV is a fairly wide resource for image recognition, deep learning, and image analysis that is becoming increasingly important in contemporary networks. OpenCV can recognize objects, faces, and even human handwriting from photos and videos. Object detection is a subset of machine learning, signal processing, and big data that deals with identifying features in photos and images. The first step entails using a large number of negative and positive labelled images to train a cascade function. After the classifier has been trained, the training images are used to extract identifying features, known as "HAAR Features." HAAR features are basically rectangular features with bright and dark pixels in different areas. The value of each function is determined by subtracting the amount of pixel intensity in the bright region from the pixel intensity in the dark region. These attributes are calculated using all of the image's potential sizes and locations. Many irrelevant features can be present in an image, while only a few ways define the object, a few related features may be used. The classifier is learned to extract useful features from the pre-labeled dataset and add sufficient weights to each feature to obtain the lowest possible errors. Poor function refers to a single feature. The weighted sum of the weak features is the final classifier.

Fig. 1 - Image Recognition Using OpenCV

The context takes up a large portion of the image; The item to be viewed is only a small portion of the picture. Cascaded classifiers are used to speed up the detection process. If even a single negative feature is detected in a region of an image during this step, the algorithm goes on to the next region after ignoring the region for further processing. The requisite object in the image is the only area that contains all of the identifying features. The requisite object in the

image is the only area that contains all of the identifying features.

b) CNN ARCHHITECTURE

A convolution neural network comprises of information and an output layer, just as various hidden layers. The hidden layers of a CNN commonly comprise of a progression ofconvolution layers that ev with an increase or other dot product. The activation function is generally a RELU layer, and is in this way followed by extra convolutions, for example, pooling layers, completely associated layers and normalization layers, referred to as hidden layers on the grounds that their sources of input and output are masked by the activation function and last convolution.

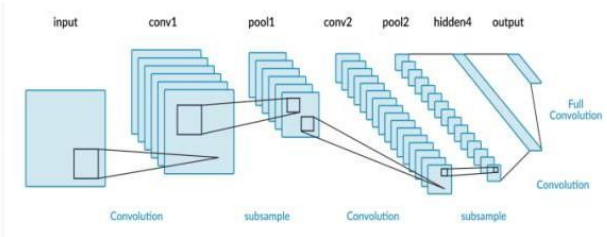

Fig. 2: Convolution Neural Networks (CNN)

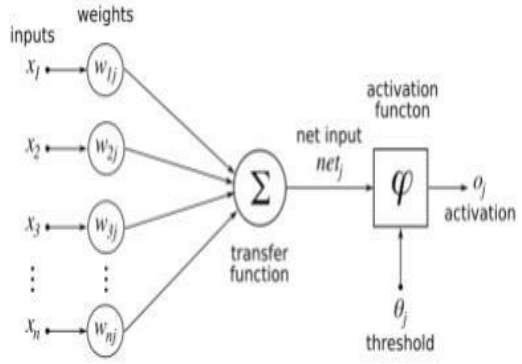

Fig. 3 - The Basic Component of Artificial Neural Networks, The Artificial Neuron, has the Following Structure

c) SVM CLASSIFIER

Support Vector Machines are a kind of supervised machine learning algorithm that gives analysis of knowledge for classification and multivariate analysis. While they will be used forregression, SVM is usually used for classification we feature out plotting within the n-dimensional space. Value of every feature is additionally the worth of the precise coordinate. Then, we discover the perfect hyper-plane that differentiates between the 2 classes. The basic principle behind the working of Support vector machines is straightforward – Create a hyper- plane that separates the data-set into classes allow us to start with a sample problem. Suppose that for a given data-set, you've got to classifyred triangles from blue circles. Your goal is to make a line that classifies the info into two classes, creating a distinction between red triangles and blue circles.

d) Single Shot Detector (SSD) algorithm

SSD is a popular object detection algorithm that was developed in Google Inc. [1]. It is based on the VGG-16 architecture. Hence SSD is simple and easier to implement.

Fig. 4. VGG-16 SSD Model.

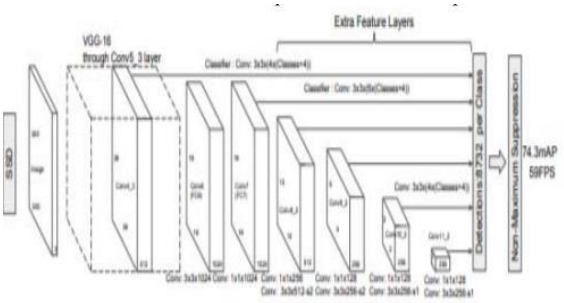

Fig. 4 shows VGG 16 SSD model. A set of default boxes is made to pass over several featuremaps in a convolutional manner. If an object detected is one among the object classifiers during prediction, then a score is generated. The object shape is adjusted to match the localization box. For each box, shape offsets and confidence level are predicted. During training, default boxes are matched to the ground truth boxes. The fully connected layers are discarded by SSD architecture. The model loss is computed as a weighted sum of confidence loss and localization loss. Measure of the deviation of the predicted box from the ground truth box is localization loss. Confidence is a measure of in which manner confidence the system is that a predicted object is the actual object. Elimination of feature resampling and encapsulation of all computation in a single network by SSD makes it simple to train with MobileNets. Compared to YOLO, SSD is faster and a method it performs explicit region proposals and pooling (including Faster R-CNN).

e) MobileNets algorithm

MobileNets uses depth wise separable convolutions that helps in building deep neural networks. The MobileNets model is more appropriate for portable and embedded vision- based applications where there is absence of process control. The main objective of MobileNets is to optimize the latency while building small neural nets at the same time. It concentrates just on size without much focus on speed. MobileNets are constructed from depth wise separable

convolutions. In the normal convolution, the input feature map is fragmented into multiple feature maps after the convolution [2].

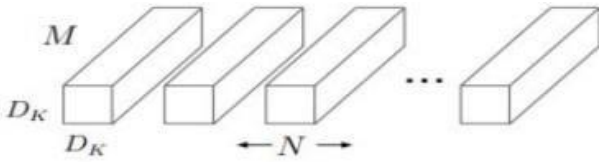

Fig. 5. Normal Convolution [2]

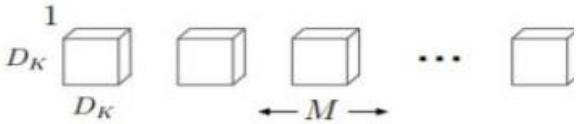

Fig. 6. Depth wise Convolution Filters [2]

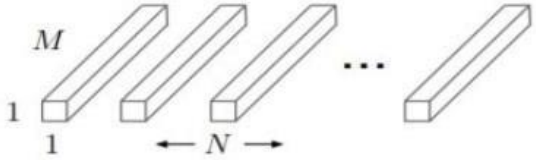

Fig. 7. 1×1 Convolutional Filters called Pointwise Convolution in the context of Depth wise Separable Convolution [2].

The number of parameters is reduced significantly by this model through the use of depth wise separable convolutions, when compared to that done by the network with normal convolutions having the same depth inthe networks. The reduction of parameters results in the formation of light weight neural networkas shown in fig 5 to 7.

IV. Conclusion:

In this paper, review on different object detection, tracking, recognition techniques, feature descriptors and segmentation method which is based on the video frame and various tracking technologies. This approach used towards increase the object detection with new ideas. Furthermore, tracking the object from the video frames with theoretical explanation is provided in bibliography content. The bibliography content is the most significant contribution of research since it will lead to anew area of research. We have identified and discussed the limitation/future scope of various methods. Also, we have noted some methods which give accuracy but have high computational complexity. Specifically, the statistical methods, background subtraction,temporal differencing with the optical flow was discussed. However, these technique needs to concentrate towards handling sudden illumination changes, darker shadows and object occlusions.

References:

1. Abdul Muhsin M, Farah F. Alkhalid, Bashra Kadhim Oleiwi, “Online Blind Assistive System using Object Recognition”, International Research Journal of Innovations in Engineering and Technology (IRJIET) ,Volume 3, Issue 12, pp 47-51, December – 2019.
2. Aishwarya Sarkale, Kaiwant Shah, Anandji Chaudhary, Tatwadarshi P. N.,

   “A Literature Survey: Neural Networks for Object Detection”, VIVA-Tech International Journal for Research and Innovation Volume 1, Issue 1 (2018)
   ISSN(Online): 2581-7280 Article No. 9.
3. Bhumika Gupta, Ashish Chaube, Ashish Negi, Umang Goel, “Study on Object Detection using Open CV Python”, International Journal of Computer Applications Foundation of Computer Science (FCS), NY, USA, Volume 162, Number 8, 2017.
4. Geethapriya. S, N. Duraimurugan, S.P. Chokkalingam, “Real-Time Object Detection with Yolo”, International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958,Volume-8, Issue-3S, February 2019.
5. Karanbir Chahal1 and Kuntal Dey, “A Survey of Modern Object Detection Literature using Deep Learning”, International Research Journal of Engineering and Technology (IRJET), Volume 8, Issue 9, 2018.

6. Kartik Umesh Sharma and Nileshsingh **V.** Thakur, "A Review and an Approach for Object Detection in Images", International Journal of Computational Vision and Robotics, Volume 7, Number 1/2, 2017.
7. Mukesh Tiwari, Dr. Rakesh Singhai, "A Review of Detection and Tracking of Object from Image and Video Sequences", International Journal of Computational Intelligence Research, Volume 13, Number 5 (2017).
8. R. Sujeetha, Vaibhav Mishra , "Object Detection and Tracking using Tensor Flow", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-1, May 2019.
9. Richard Socher, Brody Huval, Bharath Bhat, Christopher D. Manning, Andrew Y. Ng, "Convolutional-Recursive Deep Learning for 3D Object Classification", International Conference on Computational Intelligence and Communication Networks, 2018.
10. Yordanka Karayaneva and Diana Hintea, "Object Recognition in Python and MNIST Dataset Modification and Recognition with Five Machine Learning Classifiers", Journal of Image and Graphics, Volume 6, Number 1, June 2018.