IJCSN - International Journal of Computer Science and Network, Volume 8, Issue 1, February 2019
ISSN (Online): 2277-5420
www.IJCSN.org
Impact Factor: 1.5

Survey on Object Detection and Classification Using Various Methods

Hitesh Kumar Jain Jathanraj

Department of Electronics and Communication Engineering, SRM Institute of Science and Technology,
Kanchipuram dist., Tamil Nadu 603 203, India.

Abstract - Autonomous vehicles have become the recent trends, these vehicles require various functions to perform in parallel. One such function is target object detection and classification of the target object, which would aid the performance of the vehicle at various scenarios. Over the years various object detection and classification algorithms have been proposed. The paper gives overview of different kind of algorithms and methods used for object detection and classification. In Advance Diver Assistance System (ADAS), the Autonomous Emergency Braking System function depends upon the target vehicles on road in real time. So it is important to detect the object and classify them to improve the performance of the ADAS.

Keywords - *ADAS, Deep Learning, CNN, Object Detection, Object Classification.*

1. Introduction

Human can identify and detect objects present in an image easily. The human visual system is fast and accurate to identify the multiple objects and detect obstacles with conscious. Now it is easy for a computer to identify and detect the object present in an image using the better algorithms. The current state of art driverless vehicles has led to main need to object detection and classification of vehicles and objects on road. Various algorithms and modules have been proposed to address this need of the application in the field of Advanced Driver Assistance System (ADAS). Detection of vehicles on road is a significant feature of ADAS where the challenge is extraction of features from image and classification of object from its rear view also. To perform the above concepts Deep Learning concept the current state of art for vehicle detection is considered which depicts similar to neural networks which signal the brain cells in the Human nervous system. They provide significant advantages in the field of image recognition, speech recognition, face recognition, etc. The large amount of training data sets provides higher flexibility and generalization ability in extraction and manipulation stages.

The ADAS features extend to Real-time detection which requires algorithms that process fast and have high accuracy of detection. DARKNET based methods are more adaptable and applicable for Real-time application. Along with vehicles running on read it requires detection of pedestrians on road and false detection of tree as pedestrian needs to be avoided. A convolutional neural network (CNN) is also an artificial neural network which is used for image recognition and processing data using the deep learning method. CNN uses a multilayer perceptron, which consists of input layer, an output layer and a hidden layer. The hidden layers include multiple convolutional layers, pooling layers, normalization layer and fully connected layer.

2. Literature Survey

[1] The main objective explained in the paper was to detect the target object on roads for driverless systems. The authors proposed an algorithm for object detection using deep neural network concepts. The approach of the algorithm is to provide input images in which the target objects need to be detected. From the input images the rear end or front end or side part information are extracted and using these extracted information's as input to the deep learning algorithm the object detection is performed. The framework of the algorithm is firstly to collect input data. Then machine learning concepts require training data set, so the author has used CIFAR-10 data set for training and loaded to system to aid in feature extraction and detection of target object. Then ground information like region of interest is provided to algorithm, so that it helps to detect the object by extracting the required features. Then depending upon the training data set and prediction of the algorithm the object is classified. The pre-trained data set had 50,000 images which tuned the CNN. Thus for verification of algorithm the author has provided 40 on road images and has achieved 100% accuracy for these images. Thus author has proposed a reformative CNN algorithm which also reduced the cost incurred during usage of CNN algorithms. Also by increasing the recognition rate the accuracy of the system was also greatly increased.

[2] Deep learning being main feature of machine learning concept in field of object detection the author has proposed a method to improve the extraction of features from image and improvement in the classification recognition part, finally direction the algorithm to implement deep learning even in field of artificial intelligence. A general deep learning network has input at the bottom stage and final output at top stage with various intermediate layers, using rectified linear unit or s-shaped functions as the activation functions. The author chose s shaped function in the activation layers and having thousands of neurons in each hidden layers with some pre-processing of images and iteration them to some few time to some multiple of ten times. Then various images representing different road type scenarios were given for object detection and classification. By this method the author was able to achieve an error rate of 3.34% compared to neural network which provides error rate of 6.67%. Thus deep learning can be used to solve wide range of classification issues.

[3] Detection of target vehicles from one region to other region takes long time for detection and has very low recognition rates. To counter that delay author has proposed a region based CNN algorithm which is pre trained for various regions and with increased number of iterations in training which helps in optimizing network parameters. The algorithm was divided into two stages where first stage is the training stage & then second stage is the testing stage. In the training stage appropriate training sample images were taken from ImageNet and the vehicle targets where marked in the training images. They were used as input for region proposal network (RPN) until the network converges and then these converged parameters are fed as input to the R-CNN network. The parameters are varied till R-CNN network converges which in turn obtains the target vehicle detection point. The R-FCN network consists of three parts. First part is FCN which extracts the image features and creates a feature map. This feature maps are given as input to RPN in second stage which helps in deciding the region of interest (ROI). The region of vehicle with highest overlap with the marking region is considered because it has the positive value to detect and classify the target object. To train the entire network method of OHEM (online hard example mining) was used. Back propagation function is also used to calculate the loss functions which help in improving the performance of the network. Thus R-FCN was developed based on deep residual neural network with reducing the effective stride from 32 pixels to 16 pixels which increases the score mapping resolution. By this method the author was able to reach a processing speed of milliseconds for each image. This method avoids dependency of traditional target detection functions. Smaller is the deformation of target vehicle in the images higher is detection rate and with this method the author was able to achieve an accuracy of 87.48%.

[4] Autonomous vehicles require estimation of distance between moving objects. It requires image processing algorithms which can detect and classify objects without restricting the hardware platforms of the autonomous vehicle. Thus author proposes use of Darknet based deep learning algorithm with modified detection stages. Darknet has various advantages like it is the mostly efficient real time detector with adaptability feature. It has fast processing speed with simple network architecture. Kitti database is used to predict 3D bounding box from 2D convolution. To find the distance the bounding box values obtained from the image is used then manipulated with camera intrinsic and extrinsic parameters and finial transformation between coordinate systems give the absolute value.

[5] Main functions of autonomous vehicle are to provide safety, which means the vehicle has to react suddenly when something appears in front. There are times where not only vehicles are present in front but pedestrians too it requires at most detection features to detect pedestrians accurately and reduced false detection. Thus detection of pedestrian can be considered as joint estimation between detection and semantic feature extraction by which the performance of the system can be increased. In the proposed method semantic it aims in assigning object class for each pixel of image. Signet architecture of deep learning was used for pixel wise classification. The encoder requires up-sampling stage for feature classification at lower resolutions. While performing up sampling to make use of pooling indices with corresponding encoders a decoder network was proposed. The architecture was trained using cafe-segnet which has large dataset of urban traffic images. Where various semantic features like tree, bus, bi-cycle, pathways, few animals are labeled pixel wise and rest of the pixel remains blacked out at training stage. In next stage results of semantic features are integrated with detectors which make use of filtered channel features. Before feeding the input to decision forest stage for feature selection the output from these feature extraction channels is combined using sum-pooling method. Higher rate of fast and accurate classification can be achieved by using accumulated hard negative samples in cascade. This is done in booted forest method. Thus integrated feature channels can be directly used to train booted forest as there is no constrain on dimension features. Thus the results of the proposed method resulted in 4% more improvement by using more positive samples. More powerful detection of pedestrians depend upon the learning mechanism which asses the semantic cues.

[6] The type classification of the vehicle must be achieved from any view of the vehicle from image. This considering the view as an attribute author has proposed a method to classify vehicle using regional based convolution neural network framework. To achieve multi-view vehicle classification with CNN it requires training dataset. So author extracted key image frames from installed surveillance cameras and using edge boxes algorithm nearly 100 regions where generated. The input images where sampled as eighty percent for training data set, 10 percent for data validation algorithm and rest ten percent for testing stage. The ultimate aim of multi-label predicting algorithm is to label the unseen instances. The task were defined as classification of vehicle category as first task, secondly predicting the view side as semantic feature and finally in third task discrimination of background objects. The classification was evaluated on the testing dataset and the results obtained where very high. The performance was satisfactory, 83% of accuracy was achieved in vehicle type detection and about 90% accuracy in prediction of attributes. More comprehensive type of vehicle classification can obtained by using more general attributes of semantics.

[7] In this paper the author introduced Sparse Laplacian Filter learning algorithm for the network to obtain the filter from large amount of unlabeled data. The multitask learning is used to train the softmax classifier for small amounts of labeled data. The fig. 1 shows the flow chart for classification of different types of vehicles using sparse laplacian filter algorithm. The CNN architecture composes of two stages low level local features and high level global features. The learning of sparse laplacian filters in convolutional layer provides an effective filter application for the network. The author presented network of five layers as follows convolutional network, the convolutional layers, the absolute value rectification layer, the local contrast normalization layer, the average pooling layer, and the subsampling layer. The author also proposed the filter bank of convolutional layer and the softmax classifier parameter learning to elaborate study of the network. The unsupervised learning of sparse Laplacian filter learning method was introduces with three properties first sparsity of population, second high dispersal, and third sparsity of lifetime. The multitask learning algorithm also introduced for the filtering method. Certain parameters values are calculated and obtained from these two methods. The classification of sparse Laplacian filter learning without Laplacian was obtained an accuracy of 87.18% and with Laplacian the accuracy obtained was 88.11%. In multitask learning also the accuracy obtained are 88.11%. Rich and discriminative information about the vehicle can be captured with the help of sparse laplacian filters. The probability of each type of vehicle was obtained for the vehicle input images.

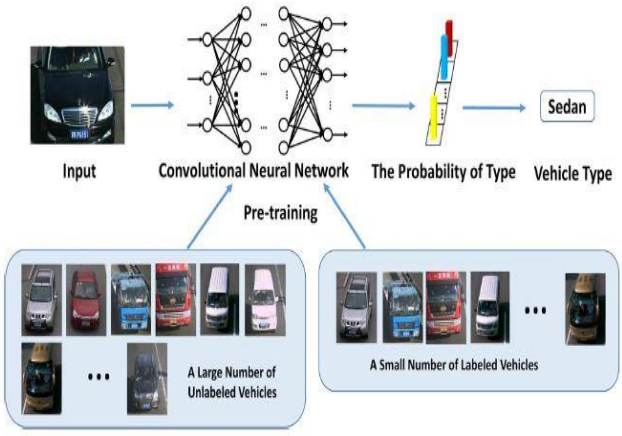

Fig. 1 Classification of different types of vehicle [7]

The below table gives the detail of different types of algorithm used for object detection and classification.

Table. 1 Different algorithms used in survey paper and their results achieved.

Title of Paper	***Authors***	***Algorithm used***	***Results achieved***
Moving vehicle detection using deep neural network	Akhil Soin et al	Deep Neural Network for object detection	100% accuracy for his set of images.
Road vehicle detection and classification based on deep neural network	Zhaojin Zhang et al	Deep Neural Network & Deep learning in Artificial Intelligence	Error rate of just 3.34%
Vehicle target detection based on R-FCN	Zhou Zhigang et al	Region based full Convolutiona l Network	Accuracy of 87.48%
MOD: Multi-camera Based Local Position Estimation for Moving Objects Detection	Jinwoo Kim et al	Darknet based deep learning method	-
Enhanced pedestrian	Tianrui Liu et al	Deep learning	1.4% improved

detection using deep learning based semantic image segmentation		based semantic features segmentation	performance compared to old Semantic segmentation-n methods.
Vehicle type classification and attribute prediction using multi-task RCNN	Zhuoqun Huo et al	Multi-task learning with help of Region based CNN.	Accuracy of 96.13%
Vehicle type classification using a Semisupervised convolutional neural network	Zhen Dong et al	Semi-supervised convolution neural network architecture	Accuracy of 96.1% in daylight & 89.4% in nightlight.

3. Conclusion

This survey paper gives an overview of recent algorithms used for object detection and classifications. To improve the performance of the ADAS real-time detection of object algorithm needs to be robust in nature with high accuracy. Various algorithms with improved accuracy are available for object detection by using Deep Neural Network, Region based full Convolution Network, Darknet Architecture, semantic features extraction, Convolution Neural Network, Pointnet Architecture, etc. Among these Darknet based deep learning method implemented with the help of YOLO provides better object detection in real time application.

References

[1] Soin, Akhil, and Manisha Chahande. "Moving vehicle detection using deep neural network." In Emerging Trends in Computing and Communication Technologies (ICETCCT), International Conference, IEEE (2017), pp. 1-5.

[2] Zhang, Zhaojin, Cunlu Xu, and Wei Feng. "Road vehicle detection and classification based on deep neural network." In Software Engineering and Service Science (ICSESS), 2016 7th IEEE International Conference, IEEE (2016), pp. 675-678.

[3] Zhigang, Zhou, Lei Huan, Ding Pengcheng, Zhou Guangbing, Wang Nan, and Zhou Wei-Kun. "Vehicle target detection based on R-FCN." In 2018 Chinese Control and Decision Conference (CCDC), IEEE (2018), pp. 5739-5743.

[4] Kim, Jinwoo, Yongbon Koo, and SungHoon Kim. "MOD: Multi-camera Based Local Position Estimation for Moving Objects Detection." In Big Data and Smart Computing (BigComp), 2018 IEEE International Conference, IEEE (2018) pp. 642-643.

[5] Liu, Tianrui, and Tania Stathaki. "Enhanced pedestrian detection using deep learning based semantic image segmentation." In Digital Signal Processing (DSP), 2017 22nd International Conference, IEEE (2017), pp. 1-5.

[6] Huo, Zhuoqun, Yizhang Xia, and Bailing Zhang. "Vehicle type classification and attribute prediction using multi-task RCNN." In Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), International Congress, IEEE (2016), pp. 564-569.

[7] Dong, Zhen, Yuwei Wu, Mingtao Pei, and Yunde Jia. "Vehicle type classification using a semisupervised convolutional neural network." IEEE transactions on intelligent transportation systems 16, no. 4 (2015): 2247-2256.