Recent Developments in Electronics and Communication Systems
KVS Ramachandra Murthy et al. (Eds.)

doi:10.3233/ATDE221267

A Framework for Object Detection with Distance Metrics in Vehicular Ad hoc Networks

R.V.S. Lalitha[a], Divya Lalitha Sri Jalligampala[a], Kayiram Kavitha[b], Ch.V.Raghavendran[c]
[a]*Department of CSE, Aditya College of Engineering & Technology, Surampalem, Andhra Pradesh, India*
[b]*CSE, Gokaraju Rangaraju Institute of Engineering & Technology, Hyderabad, India*
[c]*Department of I.T, Aditya College of Engineering & Technology, Surampalem, Andhra Pradesh, India*

Abstract. The detection and tracking of objects in autonomous vehicles is essential for operation safety. There are several approaches for computing the distance between static objects. Conventional machine learning methods are using distance metrics to calculate the distance between the objects like Manhattan distance, hamming distance and Euclidean distance based on p-norm measure. But coming to the field of moving objects the focal length is the point of concern. In this paper, the object detection and also tracking of the object is worked out from the moving camera. The detection is performed based on You Only Look Once(YOLO) algorithms and the distance is calculated by finding the focal length between the object and camera. The methods tailored gave accurate results in assessing the spatial distance between the camera and the moving object.

Keywords: YOLO, Deep Learning, Focal length, Object detection, Video streaming. Introduction

1. Introduction

Object detection is used pedestrian detection system in autonomous cars. The recognition of moving objects in autonomous cars is the researcher's interest, as vehicle steering needs to observe surroundings based on Artificial Intelligence. One of the essential requirements of autonomous cars is automatic lane detection. Autonomous cars require extraction of the contents of image and video data which helps driving assistance for lane marking. Computer vision algorithms enable cars to observe surroundings without human intervention. YOLO algorithms compare the object from video database with predefined images from the image database. The accuracy involved in assessing object depends on the population of the database and the objects arrived. The research progress in self-driving cars reduces human errors in navigation. The detection of an object depends upon the distance between the car and object while driving. Hence moving object detection is the point of concern here. Usually distance measurement in real time scenario is done usually through proximity sensors or Ultrasonic sensors. In this paper,

distance measurement is done based on video data using camera. The proposed methods involve object detection using YOLO and the distance measurement from the moving car from video data. In this paper, section 2 deals with the previous work carried out, section discusses the methods there of in detecting object and measuring distance between the object and camera, section 3 gives details about results and finally conclusion summarizes how accurately the object detection is carried out and about the distance measurement.

2. Related Work

Object detection indicates the capability of identification of objects either from image or from video. C.Tang proposed an application framework using deep learning. The deep learning techniques are compared with the traditional classification algorithms[1]. G. Chandan the Faster RCNN, SSD and YOLO are used for better accuracy. The authors proposed a new algorithm by combining SSD and Mobile Nets for efficient detection and tracking [2]. E Dong analyzed the mapping of object detection and tracking using kernel functions. The use of kernel correlation increased the real time performance for video surveillance data[3,10]. Abhishek Gupta exhibited detailed survey on object detection and perception in autonomous vehicles. Using deep learning both linear and non-linear problems can be solved by using largenumber of hyper parameters [4].Hironobu Fujiyoshi discussed image recognition with computer vision algorithms and CNN models. Alexnet does image recognition with 1000 objects for classification. The image classification with Bag of Features and the deep learning- based image classification is well discussed[5,7]. M. T. Islam performed image classification with deep learning for food dataset. The classification algorithms are applied for categorizing food items. About 16,000 images are used for categorization of food items[6]. B. Bamne suggested transfer learning-based object detection using convolution neural networks. The object detection is carried out similar to the human perceptions[8]. P. Malhotra compared various object detection algorithms RCNN, Fast RCNN and YOLO[9]. B. Rajesh analyzed object detection using MATLAB tool box and automated driving tool box [11].

3. Statistical Analysis of Object Detection Algorithms

The data collected from scopus.com database on 17-07-2022 says that there are about 56,559 documents starting from the year 2018 to 2022 using the search key TITLE-ABS-KEY (object AND detection . An intensive research work is going on in this particular area. Several popular authors worked on edge detection and edge tracking. The last five years statistics are given in Table.1.

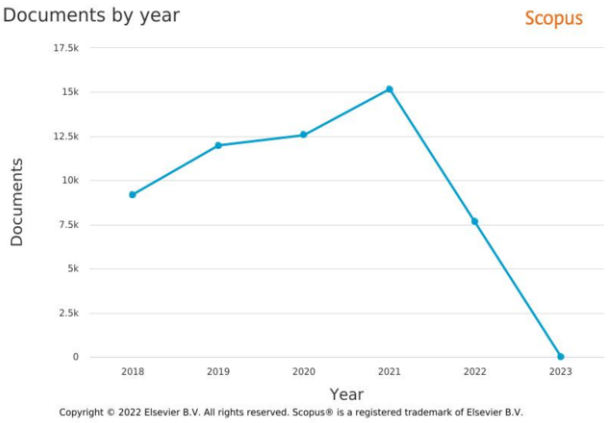

Figure.1. Documents per year by source

Table.1. No.of documents published.

Year ↓	Documents ↑
2023	14
2022	7654
2021	15169
2020	12573
2019	11973
2018	9176

4. Object Detection and Estimating its Distance

The neural network model defined in YOLOv3 model is stored in yolov3.cfg file which contains pre defined weights mentioned in yolov3.weights. For object detection 'coco.names' is used which contains 80 classes. Darknet is used to display information by loading data from config file based on weights. opencv-python, numpy, pandas, torch_nightly, matplotlib, torch and imutils libraries are used in object detection. The proposed framework is subdivided into 3 sub categories:
Object detection from video data
Computing distance between the object and camera
Displaying the label and distance in real time

4.1 Object detection from video data

Image classification involves predicting the class label of an image. Object detection is associated with two major tasks, localization and detection. Localization tells us about the location of the object and detection tells about the type of the object detected. Object localization is performed by drawing bounding boxes from the images taken from the camera. Object detection is carried out using YOLO, which classifies objects based on the confidence by comparing images from the dataset populated. Object detection locates the presence of an object in the bounding box by labeling it. Computer vision tasks are implements using OpenCV. The weights are taken from yolov3 pre-trained model weights from yolov3.weights file and with 80 classes. The expected objects usually we come across are stored in text file, like person, bicycle, car etc. The matched objects are stored in another text file. coco.names contains 80 classes and voc.names contains 20 classes of darknet open source neural network framework. For the object to be detected, we have to load names file, and then we have to prepare an image from video for submitting to neural network model. In this paper, yolov4.cfg file is used and convolution neural network is created with 7 layers. The activation functions used are mish and linear. Activation functions allow back propagation as they allow back propagation based on derivative function for the given inputs. The object will be detected based on pretrained models and assigned weights.

4.2 Computing distance between the object and camera

The object detection process is implemented using Tensor flow with CUDA (Compute Unified Device Architecture) enabled GPU (Graphics Processing Units) programming. Video is captured using VideoCapture() method and is divided into frames. The frames are divided into segments using bounding boxes and then images are identified by comparing names with the names file. For this to decide, confidence is set as 0.6. If the confidence is more than 60% then the object is finalized otherwise ignores. Softmax is used to assign probabilities to multiple classes. Convolution is used to define patterns in the image. Maxpool is used to select max values of selected inputs. Reorg is used to reorganize output feature map.To increase sampling rate upsample function is used. The average of features observed is computed using avgpool function. Connected function is used to connect neurons in one layer to neurons in next layer. Rectified Linear Activation function is used to produce output ifit is +ve otherwise zero. Leaky ReLU is same as ReLU but is has small negative slopes also. By inclusion of all this function in the torch library the detection of object is accurately done.

4.3 Displaying the label and distance in real time

The distance between the object and the camera is calculated by adding w.item() and h.item() where w and h are width and height of the bounding boxes and inline with x and y positions. The distance measurement along with the labels which are compared with names file with the confidence will be displayed.

5. Results and Discussion

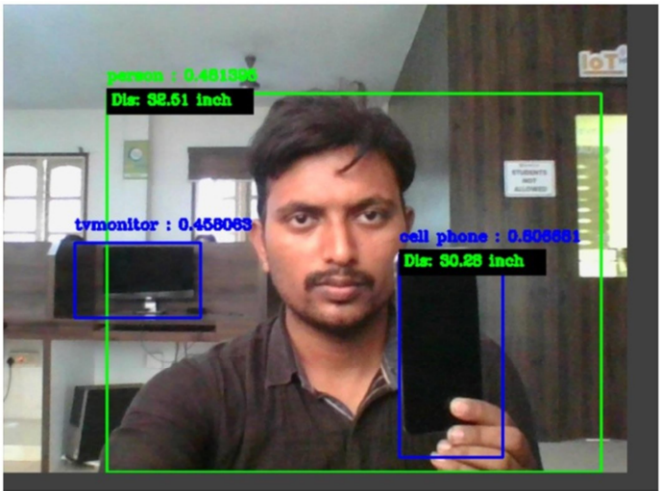

Figure 2. Object detection and its distance from camera

So, the objects are accurately detected. In Fig.2. three objects are detected accurately, tvmonitor, mobile phone and person. Their distances are also displayed.

Figure 3. Many objects are detected and distance is calculated for nearer objects

In Fig.3. the object is accurately detected, which tells us that object tracking from video stream by dividing it into frames and then into images is successfully performed. In the Figure 2many objects are detected person and mobile. The person object is detected accurately. With reference to confidence computed, it is considered as cell phone and displayed its distance.

6. Conclusion

In this research work, the object detection mechanism in autonomous cars is examined using neural network functions Softmax, Convolution, Maxpool, Reorg, avgpool and Connected functions in convolution layers. The activation functions used are Rectified Linear Activation, Leaky ReLU and ReLU. By setting the confidence 0.6 the detection of object classes are accurately classified. The YOLO bounding boxes predictions is accurately worked in our model. To operate safely, the detection and tracking of object is done and the distance is also displayed. The proposed study relates with how accurately the objects can be detected based on distance measurement. The tracking of objects is helpful for driver assistance and optimization of results is performed in this research in addition to detection alone. Further this research can be extended for effective calculation of inter-vehicular distances.

References

[1] C. Tang, Y. Feng, X. Yang, C. Zheng and Y. Zhou, "The Object Detection Based on Deep Learning," *2017 4th International Conference on Information Science and Control Engineering (ICISCE)*, 2017, pp. 723-728, doi: 10.1109/ICISCE.2017.156.

[2] G. Chandan, A. Jain, H. Jain and Mohana, "Real Time Object Detection and Tracking Using Deep Learning and OpenCV," *2018 International Conference on Inventive Research in Computing Applications (ICIRCA)*, 2018, pp. 1305-1308, doi: 10.1109/ICIRCA.2018.8597266.

[3] E. Dong, Y. Zhang and S. Du, "An Automatic Object Detection and Tracking Method Based on Video Surveillance," 2020 IEEE International Conference on Mechatronics and Automation (ICMA), 2020, pp. 1140-1144, doi: 10.1109/ICMA49215.2020.9233627.

[4] Abhishek Gupta and AlaganAnpalagan and Ling Guan and Ahmed Shaharyar Khwaja, Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues, Array, volume 10, 2021, doi:https://doi.org/10.1016/j.array.2021.100057.

[5] Hironobu Fujiyoshi and TsubasaHirakawa and Takayoshi Yamashita, Deep learning-based image recognition for autonomous driving, IATSS Research, volume 43, 4, 2019.

[6] M. T. Islam, B. M. N. Karim Siddique, S. Rahman and T. Jabid, "Image Recognition with Deep Learning," 2018 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), 2018, pp. 106-110, doi: 10.1109/ICIIBMS.2018.8550021.

[7] S. U. Nisa and M. Imran, "A Critical Review of Object Detection using Convolution Neural Network," 2019 2nd International Conference on Communication, Computing and Digital systems (C-CODE), 2019, pp. 154-159, doi: 10.1109/C-CODE.2019.8681010.

[8] B. Bamne, N. Shrivastava, L. Parashar and U. Singh, "Transfer learning-based Object Detection by using Convolutional Neural Networks," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), 2020, pp. 328-332, doi: 10.1109/ICESC48915.2020.9156060.

[9] P. Malhotra and E. Garg, "Object Detection Techniques: A Comparison," 2020 7th International Conference on Smart Structures and Systems (ICSSS), 2020, pp. 1-4, doi: 10.1109/ICSSS49621.2020.9202254.

[10] A. Rangesh and M. M. Trivedi, "No Blind Spots: Full-Surround Multi-Object Tracking for Autonomous Vehicles Using Cameras and LiDARs," in IEEE Transactions on Intelligent Vehicles, vol. 4, no. 4, pp. 588-599, Dec. 2019, doi: 10.1109/TIV.2019.2938110.

[11] B. Rajesh, D. Ramakrishna, A. Ramakrishna Raju and Ameet Chavan, Object Detection and Classification for Autonomous Vehicle, Journal of Physics: Conference Series, ICCIEA-2020.