JOURNAL OF CRITICAL REVIEWS

ISSN- 2394-5125 VOL 08, ISSUE 05, 2021

OpenCV Based Object Detection and Tracking

Chithaluri Saidulu, Madhavi Kantedi

Department of Computer Science and Engineering

Sree Dattha Group of Institutions, Hyderabad, Telangana, India.

ABSTRACT

Deep learning has gained a tremendous influence on how the world is adapting to Artificial Intelligence since past few years. Some of the popular object detection algorithms are Region-based Convolutional Neural Networks (RCNN), Faster-RCNN, Single Shot Detector (SSD) and You Only Look Once (YOLO). Amongst these, Faster-RCNN and SSD have better accuracy, while YOLO performs better when speed is given preference over accuracy. Deep learning combines SSD and MobileNets to perform efficient implementation of detection and tracking. This algorithm performs efficient object detection while not compromising on the performance.

Object detection

Frame differencing: Frames are captured from camera at regular intervals of time. Difference is estimated from the consecutive frames.

Optical flow: This technique estimates and calculates the optical flow field with algorithm used for optical flow. A local mean algorithm is used then to enhance it. To filter noise a self-adaptive algorithm takes place. It contains a wide adaptation to the number and size of the objects and helpful in avoiding time consuming and complicated preprocessing methods.

Background subtraction: It is a rapid method of localizing objects in motion from a video captured by a stationary camera. This forms the primary step of a multi-stage vision system. This type of process separates out background from the foreground object in sequence in images.

Object tracking

It is done in video sequences like security cameras and CCTV surveillance feed; the objective is to track the path followed, speed of an object. The rate of real time detection can be increased by employing object tracking and running classification in few frames captured in a fixed interval of time. Object detection can run on a slow frame rate looking for objects to lock onto and once those objects are detected and locked, then object tracking, can run in faster frame speed.

Keywords: Object detection, object tracking, OpenCV.

1. INTRODUCTION

Deep learning has gained a tremendous influence on how the world is adapting to Artificial Intelligence since past few years. Some of the popular object detection algorithms are Region-based Convolutional Neural Networks (RCNN), Faster-RCNN, Single Shot Detector (SSD) and You Only Look Once (YOLO). Amongst these, Faster-RCNN and SSD have better accuracy, while YOLO performs better when speed is given preference over accuracy. Deep learning combines SSD and MobileNets to perform efficient implementation of detection and tracking. This algorithm performs efficient object detection while not compromising on the performance.

Since AlexNet has stormed the research world in 2012 ImageNet on a large-scale visual recognition challenge, for detection in-depth learning, far exceeding the most traditional methods of artificial vision used in literature. In artificial vision, the neural convolution networks are distinguished in the classification of images.

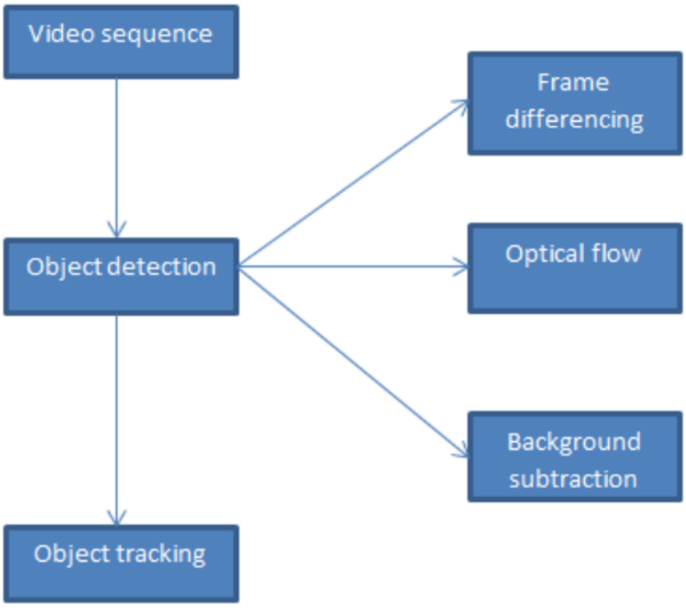


Fig. 1: Basic block diagram of detection and Tracking.

Object detection involves detecting region of interest of object from given class of image. Different methods are –Frame differencing, Optical flow, Background subtraction. This is a method of detecting and locating an object which is in motion with the help of a camera. Detection and tracking algorithms are described by extracting the features of image and video for security applications. Features are extracted using CNN and deep learning. Classifiers are used for image classification and counting. YOLO based algorithm with GMM model by using the concepts of deep learning will give good accuracy for feature extraction and classification.

2. LITERATURE SURVEY

Kini et al. proposed Real time moving vehicle congestion detection and tracking using OpenCV. An object Tracking System is used to track the motion trajectory of an object in a video. First, they use the OpenCV's function, select ROI, to select an object on a frame and track its motion using a built-in-tracker. Next, instead of using select ROI, the authors use YOLO to detect an object in each frame and track them by object centroid and size comparison. Then they combine YOLO detection with the OpenCV's built-in tracker by detecting the object in the first frame using YOLO and tracking them using select ROI. Video tracking is widely used for multiple purposes such as human computer interaction, security and surveillance, traffic control, medical imaging, and so on.

Anand, et al. proposed Object detection and position tracking in real time. Initially, Object detection is a computer vision method that enables us to recognize objects in an image or video and locate them. The authors describe an efficient shape-based object identification method and its displacement in real-time using OpenCV library of programming roles mostly targeted at computer vision and Raspberry Pi with camera module.

Abdulgafoor, et al. proposed Real-time moving objects detection and tracking using deep-stream technology and develop a real-time object detection and tracking algorithm embedded in an AI computing device known as Nvidia Jetson TX2. It improves the performance of the proposed algorithm. It brings its work closer to reality. Deep Stream Software Development Kit (DS-SDK) was used to achieve high performance and interact with multiple video sources at the same time as well. Many convolutional neural networks were used inside the proposed algorithm, such as those based on Fast Region-Convolution Neural Network (Fast R-CNN), Single Shot Detector (SSD), and You Only

Look Once (YOLO) network. Its performance in treating different video clips with deep streams of piping compared.

Amitha, et al. proposed Improved Vehicle Detection and Tracking Using YOLO and CSRT." Communication and Intelligent Systems. Initially, CSRT is mainly used for face prediction and moving object detection. The proposed system uses CSRT for vehicle tracking, particularly for cars, buses, and trucks. To perform the vehicle detection task, we have used the YOLO v3 pre-trained model. The accuracy and effectiveness of our vehicle detection and tracking system are tested with 8 different commonly available trackers in various publicly available traffic videos.

Chandhan et al. showed some of the popular object detection algorithms are Region-based Convolutional Neural Networks (RCNN), Faster-RCNN, Single Shot Detector (SSD) and You Only Look Once (YOLO). Amongst these, Faster-RCNN and SSD have better accuracy, while YOLO performs better when speed is given preference over accuracy. Deep learning combines SSD and Mobile Nets to perform efficient implementation of detection and tracking. This algorithm performs efficient object detection while not compromising on the performance.

Saxena et al. focused on the case study of a face detection and object detection like watch detection, pen detection. The goal of the system to be developed is to further ease and augment the everyday part of our lives. It’s an attempt to create own Haar classifier using OpenCV.

Jain et al. encompassed a single deep convolution neural network dividing the input into a cell grid and each cell predicts a boundary box and classifies object directly. The dataset used for multiple object detection is the KITTI dataset. It consists of 80 classes out of which five classes has been considered for this project which are: car, bus, truck, and motorcycle and train. Using the Multiple Object Detection concepts, tracking of vehicles was further implemented. The first frame of the video was taken, and Multiple object detection was performed and in the further frames of the video the object was tracked using its centroid position. This has been developed using OpenCV and Python using YOLOv3 algorithm for the object detection phase.

Ullah et al. described CPU Based YOLO, a real time object detection model to run on non-GPU computers that may facilitate the users of low configuration computer. There are a lot of well improved algorithms for object detection such as YOLO, Faster R-CNN, Fast R-CNN, R-CNN, Mask R-CNN, R-FCN, SSD, RetinaNet etc. In this model, framework optimized YOLO with OpenCV such a way that real time object detection can be possible on CPU based Computers. This model detects object from video in 10.12 - 16.29 FPS and with 80-99% confidence on several non-GPU computers. CPU Based YOLO achieves 31.05% mAP.

Juneja et al. used ssd_v2_inception_coco model as Single Shot Detection models deliver significantly better results. A dataset of more than 100 raw images is used for training and then xml files are generated using labellimg. Tensor flow records generated are passed through training pipelines using the proposed model. OpenCV captured real-time images and CNN performs convolution operations on images. In the proposed model, accuracy of object detection significantly improvises over existing methodologies in practice. There is a substantial dataset to evaluate the accuracy of proposed model. The model may be readily useful for object detection applications including parking lots, human identification, and inventory management.

Pandey, et al. proposed Detect and Track the Motion of Any Moving Object Using OpenCV. The authors found that a video picture is moving and that moving object is detected using OpenCV and the detected picture has been represented in frames with the help of contour by computer vision (CV) in a computer system. Detecting and recognizing an object is the initial stage of image systems in computer vision. Therefore, is a real-time identification of tracking a large moving object system using open computer vision (CV).

3. PROPOSED SYSTEM

Fig. 2: Block diagram of proposed system.

YOLO Algorithm

What is YOLO?

YOLO is an abbreviation for the term 'You Only Look Once'. This is an algorithm that detects and recognizes various objects in a picture (in real-time). Object detection in YOLO is done as a regression problem and provides the class probabilities of the detected images.

YOLO algorithm employs convolutional neural networks (CNN) to detect objects in real-time. As the name suggests, the algorithm requires only a single forward propagation through a neural network to detect objects.

This means that prediction in the entire image is done in a single algorithm run. The CNN is used to predict various class probabilities and bounding boxes simultaneously.

The YOLO algorithm consists of various variants. Some of the common ones include tiny YOLO and YOLOv3.

Why the YOLO algorithm is important

YOLO algorithm is important because of the following reasons:

- Speed: This algorithm improves the speed of detection because it can predict objects in real-time.
- High accuracy: YOLO is a predictive technique that provides accurate results with minimal background errors.
- Learning capabilities: The algorithm has excellent learning capabilities that enable it to learn the representations of objects and apply them in object detection.

How the YOLO algorithm works

YOLO algorithm works using the following three techniques:

- Residual blocks
- Bounding box regression
- Intersection Over Union (IOU)

Residual blocks: First, the image is divided into various grids. Each grid has a dimension of S x S. The following image shows how an input image is divided into grids.

ISSN- 2394-5125 VOL 08, ISSUE 05, 2021

Fig. 3: Residual blocks image source.

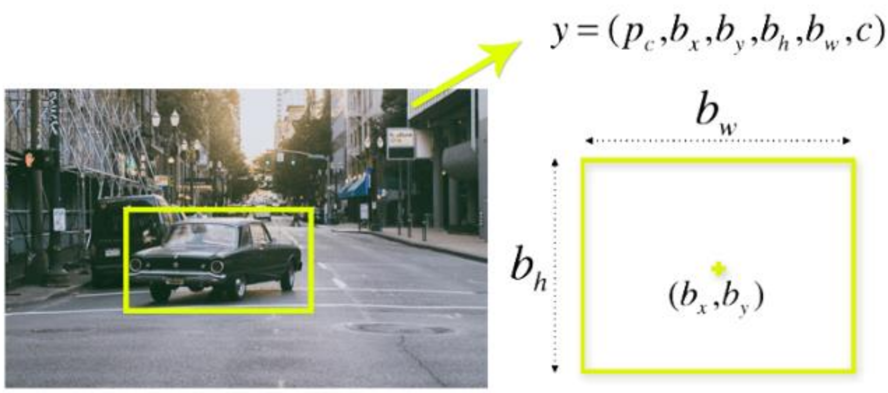
In the image above, there are many grid cells of equal dimension. Every grid cell will detect objects that appear within them. For example, if an object center appears within a certain grid cell, then this cell will be responsible for detecting it.

Bounding box regression: A bounding box is an outline that highlights an object in an image.

Every bounding box in the image consists of the following attributes:

- Width (bw)
- Height (bh)
- Class (for example, person, car, traffic light, etc.)- This is represented by the letter c.
- Bounding box center (bx,by)

The following image shows an example of a bounding box. The bounding box has been represented by a yellow outline.



YOLO uses a single bounding box regression to predict the height, width, center, and class of objects. In the image above, represents the probability of an object appearing in the bounding box.

Intersection over union (IOU): Intersection over union (IOU) is a phenomenon in object detection that describes how boxes overlap. YOLO uses IOU to provide an output box that surrounds the objects perfectly.

Each grid cell is responsible for predicting the bounding boxes and their confidence scores. The IOU is equal to 1 if the predicted bounding box is the same as the real box. This mechanism eliminates bounding boxes that are not equal to the real box.

How YOLO Algorithm Works

First, the image is divided into grid cells. Each grid cell forecasts B bounding boxes and provides their confidence scores. The cells predict the class probabilities to establish the class of each object.

For example, we can notice at least three classes of objects: a car, a dog, and a bicycle. All the predictions are made simultaneously using a single convolutional neural network.

Intersection over union ensures that the predicted bounding boxes are equal to the real boxes of the objects. This phenomenon eliminates unnecessary bounding boxes that do not meet the characteristics of the objects (like height and width). The final detection will consist of unique bounding boxes that fit the objects perfectly.

For example, the car is surrounded by the pink bounding box while the bicycle is surrounded by the yellow bounding box. The dog has been highlighted using the blue bounding box.

Applications of YOLO

YOLO algorithm can be applied in the following fields:

Autonomous driving: YOLO algorithm can be used in autonomous cars to detect objects around cars such as vehicles, people, and parking signals. Object detection in autonomous cars is done to avoid collision since no human driver is controlling the car.

Wildlife: This algorithm is used to detect various types of animals in forests. This type of detection is used by wildlife rangers and journalists to identify animals in videos (both recorded and real-time) and images. Some of the animals that can be detected include giraffes, elephants, and bears.

Security: YOLO can also be used in security systems to enforce security in an area. Let's assume that people have been restricted from passing through a certain area for security reasons. If someone passes through the restricted area, the YOLO algorithm will detect him/her, which will require the security personnel to take further action.

Pre-processing

Data Pre-processing in Machine learning

Data pre-processing is a process of preparing the raw data and making it suitable for a machine learning model. It is the first and crucial step while creating a machine learning model.

When creating a machine learning project, it is not always a case that we come across the clean and formatted data. And while doing any operation with data, it is mandatory to clean it and put in a formatted way. So, for this, we use data pre-processing task.

Why do we need Data Pre-processing?

A real-world data generally contains noises, missing values, and maybe in an unusable format which cannot be directly used for machine learning models. Data pre-processing is required tasks for cleaning the data and making it suitable for a machine learning model which also increases the accuracy and efficiency of a machine learning model.

- Getting the dataset
- Importing libraries

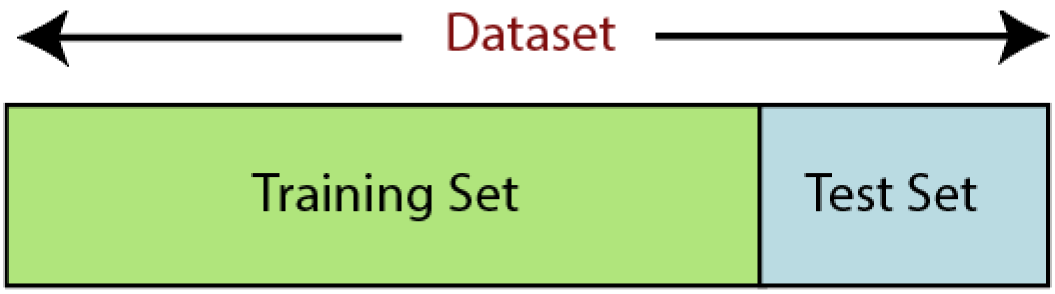
- Importing datasets
- Finding Missing Data
- Encoding Categorical Data
- Splitting dataset into training and test set
- Feature scaling

Splitting the Dataset into the Training set and Test set

In machine learning data pre-processing, we divide our dataset into a training set and test set. This is one of the crucial steps of data pre-processing as by doing this, we can enhance the performance of our machine learning model.

Supposeif we have given training to our machine learning model by a dataset and we test it by a completely different dataset. Then, it will create difficulties for our model to understand the correlations between the models.

If we train our model very well and its training accuracy is also very high, but we provide a new dataset to it, then it will decrease the performance. So we always try to make a machine learning model which performs well with the training set and also with the test dataset. Here, we can define these datasets as:



Training Set: A subset of dataset to train the machine learning model, and we already know the output.

Test set: A subset of dataset to test the machine learning model, and by using the test set, model predicts the output.

Advantages of proposed system

- Process frames at the rate of 45 fps (larger network) to 150 fps (smaller network) which is better than real-time.
- The network can generalize the image better.

4. RESULTS AND DISCUSSION

Modules

- Importing libraries
- Image/video acquisition
- Apply YOLO algorithm
- Object detection
- Object tracking

ISSN- 2394-5125 VOL 08, ISSUE 05, 2021

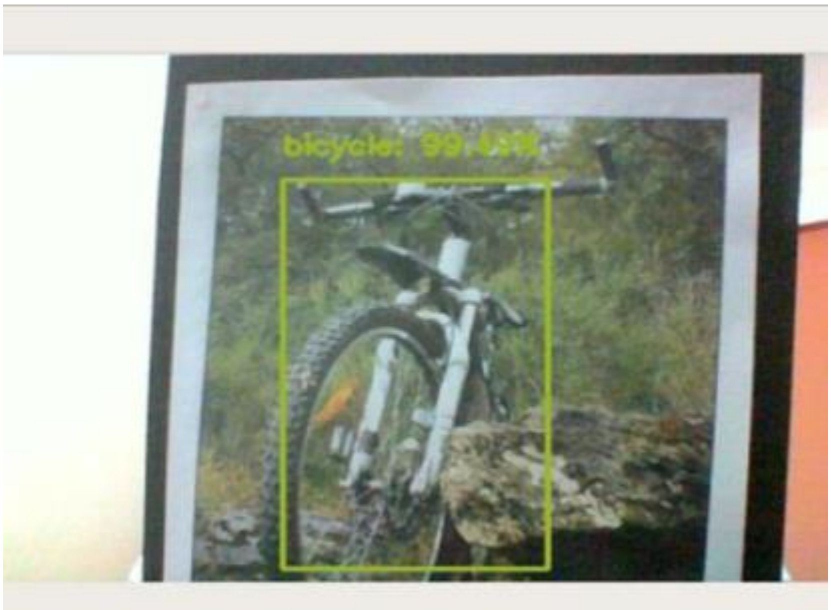


Fig. 3: Detection of bicycle.

Fig. 4: Detection of bus.

Fig. 5: Detection of train.

ISSN- 2394-5125 VOL 08, ISSUE 05, 2021

Fig. 6: Detection of dog.

5. CONCLUSION

Objects are detected using YOLO algorithm in real time scenarios. Additionally, YOLO have shown results with considerable confidence level. Main Objective of YOLO algorithm to detect various objects in real time video sequence and track them in real time. This model showed excellent detection and tracking results on the object trained and can further utilized in specific scenarios to detect, track, and respond to the targeted objects in the video surveillance. This real time analysis of the ecosystem can yield great results by enabling security, order, and utility for any enterprise.

REFERENCES

[1] Kini, Sowmya. "Real time moving vehicle congestion detection and tracking using OpenCV." Turkish Journal of Computer and Mathematics Education (TURCOMAT) 12.10 (2021): 273-279.

[2] Anand, Gokulnath, and Ashok Kumar Kumawat. "Object detection and position tracking in real time using Raspberry Pi." Materials Today: Proceedings 47 (2021): 3221-3226.

[3] ABDULGHAFOOR, NUHA H., and HADEEL N. ABDULLAH. "Real-time moving objects detection and tracking using deep-stream technology." Journal of Engineering Science and Technology 16.1 (2021): 194-208.

[4] Amitha, I. C., and N. K. Narayanan. "Improved Vehicle Detection and Tracking Using YOLO and CSRT." Communication and Intelligent Systems. Springer, Singapore, 2021. 435-446.

[5] 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.

[6] Saxena, Meghna Raj, et al. "Real-time object detection using machine learning and opencv." Int J Inform Sci Appl (IJISA) 11.1 (2019): 0974-225.

[7] N. Jain, S. Yerragolla, T. Guha and Mohana, "Performance Analysis of Object Detection and Tracking Algorithms for Traffic Surveillance Applications using Neural Networks," 2019 Third International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2019, pp. 690-696, doi: 10.1109/I-SMAC47947.2019.9032502.

[8] M. B. Ullah, "CPU Based YOLO: A Real Time Object Detection Algorithm," 2020 IEEE Region 10 Symposium (TENSYMP), 2020, pp. 552-555, doi: 10.1109/TENSYMP50017.2020.9230778.

[9] Juneja, Abhinav, et al. "Real time object detection using CNN based single shot detector model." Journal of Information Technology Management 13.1 (2021): 62-80.

[10] Pandey, Harikesh, Pushpa Choudhary, and Arjun Singh. "Detect and Track the Motion of Any Moving Object Using OpenCV." Cyber Security in Intelligent Computing and Communications. Springer, Singapore, 2022. 355-363