International Conference on New Horizons in Science Engineering Technology (NHSET-2018)
International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 5 | ISSN : 2456-3307

Real Time Object Classifier

Hemaanand M, Sanjaykumar V
B.Tech, ECE, Amrita University, Coimbatore, Tamil Nadu, India

ABSTRACT:

In this paper, the new surveillance system has been presented using the mobile net Single Shot multibox Detector - SSD and You Only Look Once -YOLO, a modern technique to object detection and classification and feature analysis. Nowadays there are many security breaches happening around us, which causes a great loss to the victims. In order to avoid this discrepancy, we came up with a smart solution that is to develop a security system which can avoid these kinds of security breaches and intrusion or trespassing. Using the YOLO, it is possible to detect to identity of the objects and classify them based on the inserted modules and trained data-sets. Our main objective is to create a smart surveillance system which can predict the intruders or even vehicles, objects and create a way more powerful security system.In this process we have considered two methods for the surveillance, both YOLO and mobile net SSD.The mobile net SSD is used to define and train the classes and set the range of accuracy and YOLO for additional features thus making the YOLO learn general representations of objects. It outshines other detection algorithms, including Deformable Part Models (DPM) and Regional-wise convolutional Neural network (R-CNN), when generalizing from natural images to other domains like surveillance and security.

Keywords: Classification, Detection, Neural networks, Surveillance.

INTRODUCTION:

In today's world, Object identification is required to be fast, accurate and be able to identify a wide variety of objects. With the upcoming of neutral networks, detection frameworks have become increasingly fast and precise. However, most detection techniques are confined to a small circle of objects.

The entire process works on the probability of classes to which the objects belong. ASingle neural network has the ability to predict bounding boxes succeeded by class probabilities directly from the images with a single evaluation. While these results are captivating, Object classification and identification is far simpler than the complexity[1][2].

The working of this system is similar to that of normal surveillance systems but it also offers a wide range of additional features that includes the object recognition and feature classification that includes vehicles, objects and even animals

It can be used both indoor and outdoor areas whereas the classification occurs based on the predefined data that has been already set to the algorithm the functionality requires data from the

users as it trains the data more than twice the mean average precision of other real-time systems. First of all the mobile net SSD performs the overall convolutional scheme on the live-stream video obtained Secondly, YOLO works on global reasoning about the image during predictions.

Unlike, the sliding window and the regional-based proposal techniques, YOLO sees the entire image during training and test time so it implicitly encodes contextual information about the classes as well as their appearances.
When the program in encoded in any of the live-streaming devices such as cameras or a pi-cam connected to a raspberry pi which acts as an individual operating system Then the system undergoes the critical phase of deploying the models that has been trained by the user Enables the live video stream through the camera obtains the data and process a set of convolutions that has been given through the code Initializes the open cl run time

FUNCTIONAL BLOCK DIAGRAM OF ALGORITHM:

This result accords with the that tasks with long execution time are harmful to the schedulability of data short processing data must me faster since the output is functioning with live-stream data classification .The above program, we can expect that the test conditions proposed in this paper, especially the improved test conditions, have a significant performance improvement.

MOBILE NET SSD:
Processing layers of the mobile net SSD

And diversity of visual understanding for computers which leads the breakthrough for futuristic technologies. The Computer vision gets complicated with multiple overlapping objects and different backgrounds and we humans not only classify these different objects but also identity the relations to one another. Hence it depends on how we train them with the modules. YOLOproves to be a promising object detection technique since we frame detect as a regression problem replacing thecomplex pipeline[3].

Surveillance is not just about securing a confined or a particular area. This idea of object detection can be programmed into robots to make them secure not just a building but a perimeter around it. This technology can also be embedded in CCTV cameras and vehicle collision technology. The object detection technique is generally confined to specific modules but can be widely used if a huge variety of modules are trained using the Mobile net SSD[4][5].

EXPERIMENT:

We need to first either choose an image or to be more applicative can also live stream a video to detect the objects inside, where they are, and how much percentage they match with the trained modules.Unlike the other traditional classifier, the combination of YOLO and Mobile SSD divides

the image into 13 by 13 cells. Each of these cells can predict 5 bounding boxes.

After the image is captured, the process of convolution occurs which is followed by max pooling. During convolution, the image is convolved with a mask. The mask is also a signal. In order to perform the process of convolution on an image, different values of masks are used by sliding or superimposing the mask on the image and then multiplying the corresponding elements and then max pooling is carried out. Max pooling is a process that uses the maximum number of value of a pixel in the default dimensions used in the convolution process. This process is repeated again and again in order to find the features.

In Mobile net SSD, we usebatch normalization which performs feature scaling in neural networks for the data between the different layers. It normally occurs after the convolutional layer and basically computes the features for the image.

ALGORITHMS USED:

The most common detection techniques are the DPM (Deformable parts models) which uses a sliding window approach, where the classifier is run at evenly spaced locations over the entire image, then the R-CNN (Regional-wise Convolutional neural network) which uses regional proposal technique to first generate potential bounding boxes in an image and then run a classifier on these proposed boxes. These techniques exhibit the post processing delay i.e. the post-processing is used to refine bounding boxes and eliminate the duplicate detection's and re score the boxes based on the other objects in the scene. These complex pipelines exhibit significant delay and are difficult to optimize since the components need to be trained individually[6].

Whereas, YOLO (You only look once) proves to be a better alternative as it mainly focuses on reducing the delay i.e. fast and accurate. It's relatively a simpler model than the other detection techniques. A Single Convolutional network simultaneously predicts multiple bounding boxes and class probabilities for those boxes. YOLO trains on full images and directly optimizes detection performance. This combined model as several merits over the other techniques.

Basically, YOLO is all about speed, it is fast and highly accurate. As prescribed earlier, it doesn't have complex pipelines. Furthermore, YOLO achieves. Speaking of applications, Image recognition and classification is a problem that has been around for a long time and many pragmatic applications. Our aim is to use the YOLO technique for surveillance and security[7].

OBJECT DETECTION:

MobileNet can also be deployed as an effective base network in modern object detection systems .In this SSD is evaluated with 300 input resolutions (SSD 300) and Faster-RCNN is compared

with both 300 and 600 input resolution (Faster RCNN
300, Faster-RCNN 600). The Faster-RCNN model evaluates 300 RPN proposal boxes per image. The models are trained on caffe model. When the convolution has been successfully finished then the detected output get bounded by a box[8].

FUNCTIONAL BLOCK DIAGRAM:

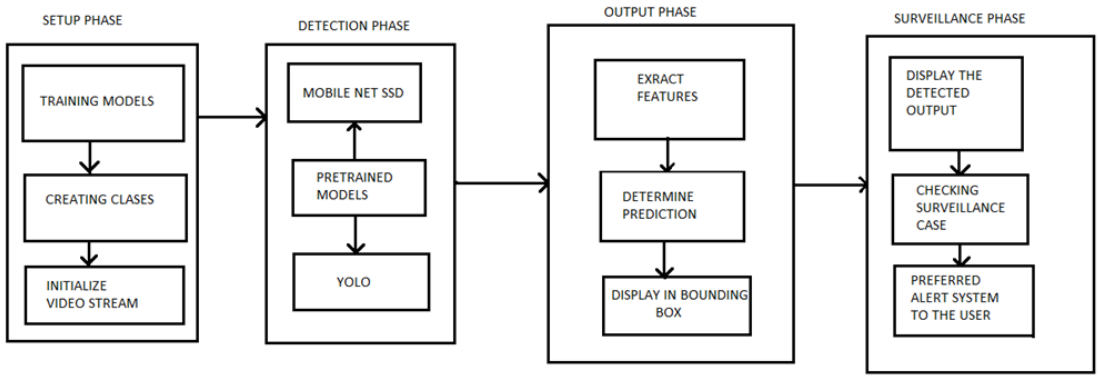


Fig. 1 Flow of Work WORKING:

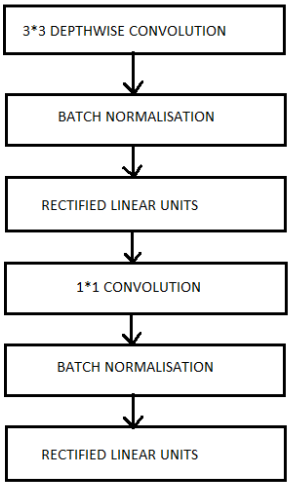


Fig. 2 Processing Layers of mobile net SSD

YOLO:

Different layers of the processing parts YOLO and the convolution layers

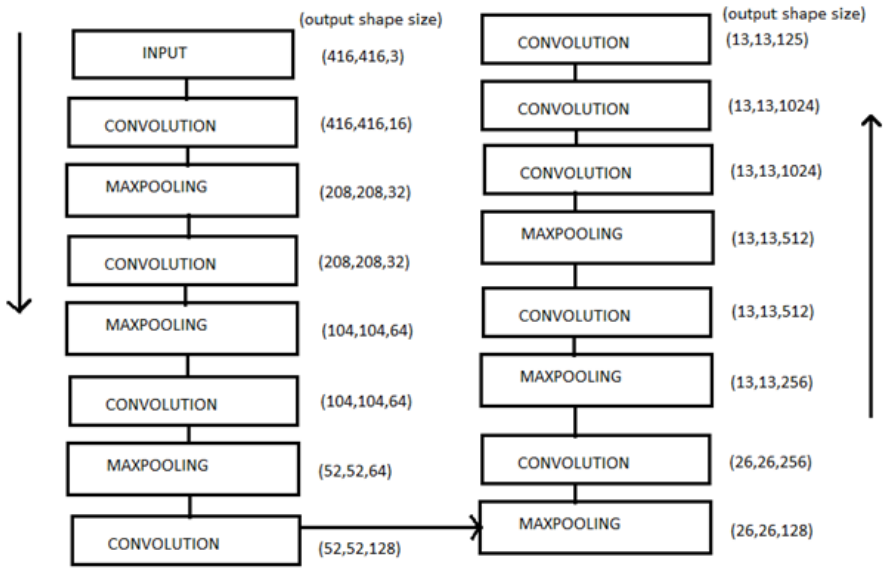


Fig. 3 Layers of YOLO and Convolution Layers

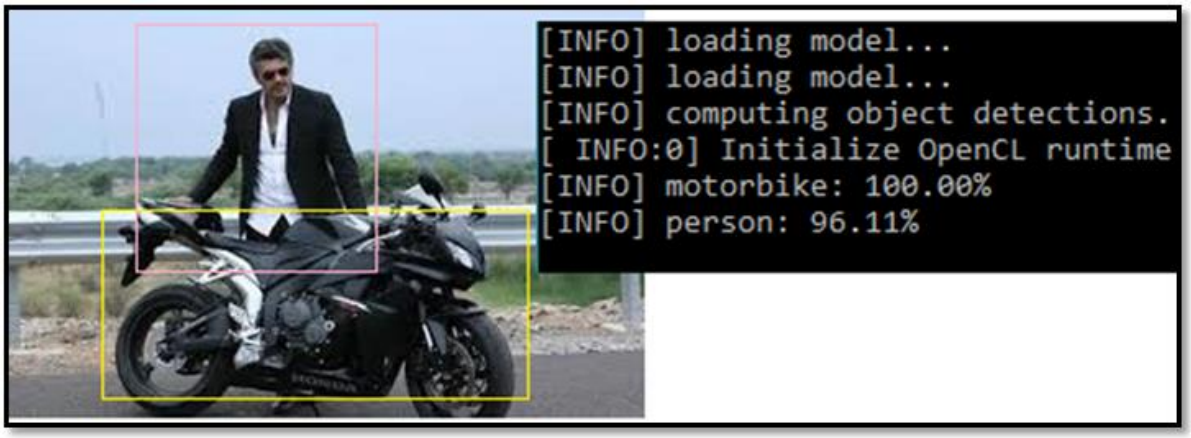
RESULTS:

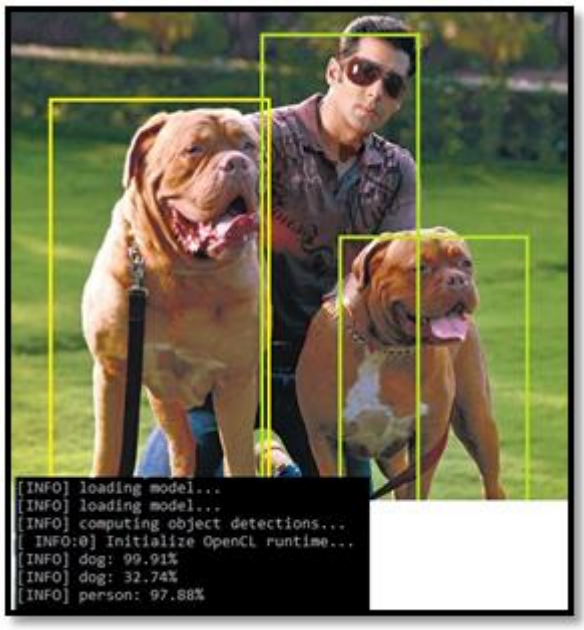
The object detection of YOLO is a quick and accurate technique for ideal computer vision applications. We can verify its real-time performance by connecting it to the webcam and analysing its fetching and displaying time. YOLO generally bounds images individually but when connected to the webcam, it can track down multiple images.

Now comparing the YOLO with the other object detection techniques, YOLO proves to be a promising method as it shows a 53.3% average precision while the DPM and RCNN show 37.8% and 10.4% precision respectively. YOLO has a recall efficiency of 54% while the DPM and RCNN have a recall efficiency of 47% and 18% respectively. This analysis is made with respect to the Picasso Dataset precision-recall curves. The comparison of these techniques with respect to other data sets is also been done:

Table 1. Comparison of the Techniques

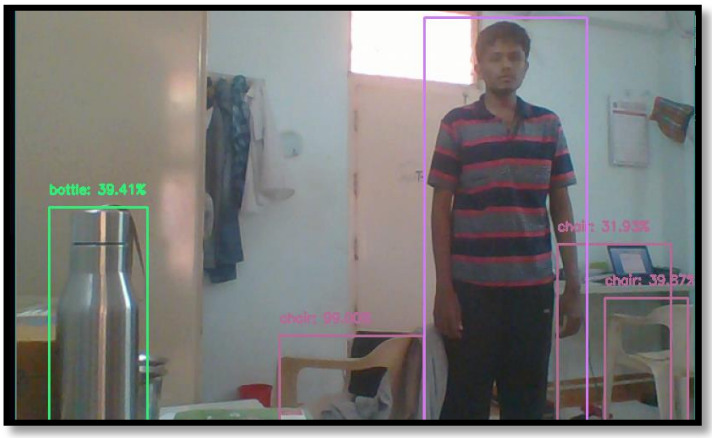
TECHNIQUES	**Visual Object classes 200**	**Picasso**	**People-art**
	Average precision	**Average precisi**	**Average precision**
D&T (Detection and tracking)	-	1.9	-
Poselets	36.5	17.8	-
DPM (Deformable parts models)	43.2	37.8	32
R-CNN (Regional-wise Convolutional Neural network)	54.2	10.4	26
YOLO (You only look once)	**59.2**	**53.3**	**45**


[INFO] loading model...
[INFO] loading model...
[INFO] computing object detections.
[INFO:0] Initialize OpenCL runtime
[INFO] motorbike: 100.00%
[INFO] person: 96.11%


[INFO] loading model...
[INFO] loading model...
[INFO] computing object detections...
[INFO:0] Initialize OpenCL runtime...
[INFO] dog: 99.91%
[INFO] dog: 32.74%
[INFO] person: 97.88%


person: 99.55%
chair: 99.69%


[INFO] loading model...
[INFO] loading model...
[INFO] computing object detections...
[INFO:0] Initialize OpenCL runtime...
[INFO] chair: 31.98%
[INFO] person: 71.49%
[INFO] sofa: 94.60%


bottle: 39.41%
chair: 31.93%
chair: 39.87%


[INFO] loading model...
[INFO] loading model...
[INFO] computing object detections...
[INFO:0] Initialize OpenCL runtime...
[INFO] person: 93.44%
[INFO] person: 89.10%
[INFO] person: 85.94%
[INFO] person: 83.83%
[INFO] person: 75.71%
[INFO] person: 72.48%
[INFO] person: 43.69%
[INFO] person: 35.83%
[INFO] person: 33.04%
[INFO] person: 32.96%
[INFO] person: 30.27%
[INFO] person: 29.91%
[INFO] person: 29.57%
[INFO] person: 28.56%


[INFO] loading model...
[INFO] loading model...
[INFO] computing object detections...
[INFO:0] Initialize OpenCL runtime...
[INFO] chair: 60.00%
[INFO] tvmonitor: 35.34%

CONCLUSION:

Hence, the combination of Mobile net SSD and YOLO is a unified model of object detection and identification. Our objective is to construct a Mobile SSD algorithm along with YOLO to exhibit extra features mainly for the surveillance system. These algorithms can be extended to various domains based on the applications. Unlike the other techniques, Mobile SSD along with YOLO shows extreme performance in object and detection process.

With various data-sets like the VOC (Visual object classes), Picasso and the People-art, our technique shows a relatively higher performance both in average precision and recall. As of today, YOLO is one of the fastest artificial intelligence algorithms to run object classification. The Mobile SSD and YOLO algorithm have been trained with many modules on the python platform for the surveillance purposes and its overall performance can be enhanced if more and more modules are trained so as to detect a wide range of objects

REFERENCES:

[1] J. Dong, Q. Chen, S. Yan, and A. Yuille. Towards unifiedObject detection and semantic segmentation. In Computer Vision–ECCV 2014, pages 299–314. Springer, 2014.
[2] S. Ren, K. He, R. B. Girshick, X. Zhang, and J. Sun. Object detection networks on convolutional feature maps. CoRR, abs/1504.06066, 2015.
[3] C. P. Papageorgiou, M. Oren, and T. Poggio. A general framework for object detection. In Computer vision, 1998. Sixth international conference on, pages 555–562. IEEE, 1998.
[4] D. Erhan, C. Szegedy, A. Toshev, and D. Anguelov. Scalable object detection using deep neural networks. In Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on, pages 2155–2162. IEEE, 2014.
[5] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan. Object detection with discriminatively trained part based models. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(9):1627–1645, 2010. 1
[6] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun. Overfeat: Integrated recognition, localization and detection using convolutional networks. CoRR, Abs/1312.6229, 2013.
[7] P. Viola and M. Jones. Robust real-time object detection. International Journal of Computer Vision, 4:34–47, 2001.
[8] S. Gould, T. Gao, and D. Koller. Region-based segmentation and object detection. In Advances in neural informationprocessing systems, pages 655–663, 2009.