GALGOTIAS
UNIVERSITY

SCHOOL OF COMPUTING AND SCIENCE AND ENGINEERING

IMAGE CLASSIFICATION
USING PYTHON

LALIT KUMAR SHARMA
BCA SEM V1

Abstract

Efficient and accurate object detection has been an important topic in the advancement of computer vision systems. With the advent of deep learning techniques, the accuracy for object detection has increased drastically. The project aims to incorporate state-of-the-art technique for object detection with the goal of achieving high accuracy with a real-time performance. A major challenge in many of the object detection systems is the dependency on other computer vision techniques for helping the deep learning based approach, which leads to slow and non-optimal performance. In this project, we use a completely deep learning based approach to solve the problem of object detection in an end-to-end fashion. The network is trained on the most challenging publicly available dataset (PASCAL VOC), on which a object detection challenge is conducted annually. The resulting system is fast and accurate, thus aiding those applications which require object detection.

TABLE OF CONTENTS

1.INTRODUCTION

A few years ago, the creation of the software and hardware image processing systems was mainly limited to the development of the user interface, which most of the programmers of each firm were engaged in. The situation has been significantly changed with the advent of the Windows operating system when the majority of the developers switched to solving the problems of image processing itself. However, this has not yet led to the cardinal progress in solving typical tasks of recognizing faces, car numbers, road signs, analyzing remote and medical images, etc. Each of these "eternal" problems is solved by trial and error by the efforts of numerous groups of the engineers and scientists.

As modern technical solutions are turn out to be excessively expensive, the task of automating the creation of the software tools for solving intellectual problems is formulated and intensively solved abroad. In the field of image processing, the required tool kit should be supporting the analysis and recognition of images of previously unknown content and ensure the effective development of applications by ordinary programmers. Just as the Windows toolkit supports the creation of interfaces for solving various applied problems.

Object recognition is to describe a collection of related computer vision tasks that involve activities like identifying objects in digital photographs. Image classification involves activities such as predicting the class of one object in an image. Object localization is refers to identifying the location of one or more objects in an image and drawing an abounding box around their extent. Object detection does the work of combines these two tasks and localizes and classifies one or more objects in an image. When a user or practitioner refers to the term “object recognition“, they often mean “object detection“. It may be challenging for beginners to distinguish between different related computer vision tasks.

So, we can distinguish between these three computer vision tasks with this example:

Image Classification: This is done by Predict the type or class of an object in an image.

Input: An image which consists of a single object, such as a photograph.

Output: A class label (e.g. one or more integers that are mapped to class labels).

Object Localization: This is done through, Locate the presence of objects in an image and indicate their location with a bounding box.

Input: An image which consists of one or more objects, such as a photograph.

Output: One or more bounding boxes (e.g. defined by a point, width, and height).

Object Detection: This is done through, Locate the presence of objects with a bounding box and typesor classes of the located objects in an image.

2.LITERATURE REVIEW

One of the foremost common Al techniques used for processing big data is machine learning, a self-adaptive algorithm that gets increasingly better analysis and patterns with experience or with newly added data.

YOLO came on the pc vision scene with the seminal 2015 paper by Joseph Redmon et al. "You Only Look Once: Unified, Real Time Object Detection, and immediately got plenty of attention by fellow computer vision researchers. Compared to state-of-the-art detection systems, Compared to state-of the-art detection systems, YOLO makes more localization errors but is much less likely to predict false detections where nothing exists.

Object detection may be a computer vision technique for locating instances of objects in images or videos Object detection algorithms typically leverage machine learning or deep learning to provide meaningful results. When humans have a look at images or video, we will recognize and locate objects of interest within a matter of moments. The goal of object detection is to duplicate this intelligence employing a computer In our context of object detection we are visiting use Pre-train model from tensor flow that are visiting help to detect and determine the item with casually labelling it with what category it belongs. While the model are retrain the algorithm that are we visiting use is Darknet YOLO that was created by

Joseph Redmon and Ali Farhadi. To be accurate to their algorithm utilized in the mode are the image of the tensor flow.

3.IMPLEMENTATION

Over two different machine the necessities was different to figure on the project thanks to this the project had different outcome in both of the cases. Algorithm were differ and data process was same yet different success rate. The accommodate TENSERFLOW is pretty simple. It went on the higher cases against R CNN and CNN model. In traditional computer vision approaches, a window was accustomed search for objects at different locations and scales. Because this was such an upscale operation, the ratio of the thing was usually assumed to be fixed. TENSERFLOW on the opposite hand approaches the thing detection problem in an exceedingly completely different way. It forwards the entire image just one occasion through the network. SSD is another object detection algorithm that forwards the image once through a deep learning network, but TENSERFLOW is way faster than SSD while achieving very comparable accuracy.

How Does it work you ask?

First, it divides the image into a 13x13 grid of cells, the dimensions of those 169 cells vary betting on the dimensions of the input. For a 416x416 input size that we utilized in our experiments, the cell size was 32x32 Each cell is then chargeable for predicting variety of boxes within the image.

For each bounding box, the network also predicts the arrogance that the bounding box actually encloses an object, and therefore the probability of the enclosed object being a specific class.

Most of those bounding boxes are eliminated because their confidence is low or because they're enclosing the identical object as another bounding box with very high confidence score. This method is termed non-maximum suppression.

The authors of YOLOv3, Joseph Redmon and Ali Farhadi, have made YOLOv3 faster and more accurate than their previous work YOLOv2. YOLOv3 handles multiple scales better. They have also improved the network by making it bigger and taking it towards residual networks by adding shortcut connections.

***Tools:** Python and OpenCV

Here are a few reasons you may want to use

1. **Easy integration with an OpenCV application**: If your application already uses OpenCV and you simply want to use YOLOv3, you don’t have to worry about compiling and building the extra Darknet code.
2. **Python support**: Darknet is written in C, and it does not officially support Python. In contrast, OpenCV does. There are python ports available for Darknet though.

4.METHODOLOGY

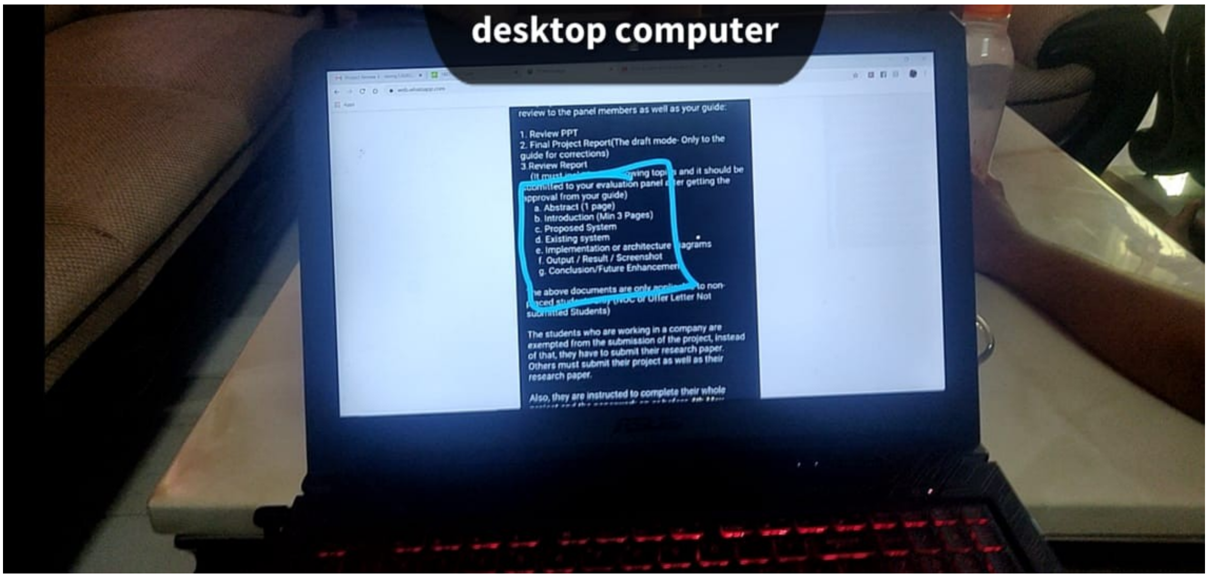
While there are lot of model already skyrocketing in the market, this steady yet better working model have better approach and simplicity to it. The duration of analyzing the different model to seek the desire output in different model are different. As one model may fast but other develop models like YOLO are seemingly good and fast then other classifier based model.

The output are 85% accurate in some cases but its batter than previous based model. So the model may need to increase the data trained into it.

5. OUTPUT/RESULTS:

So in our process of finding a better solution, we compared the output below:

Fig:5:output

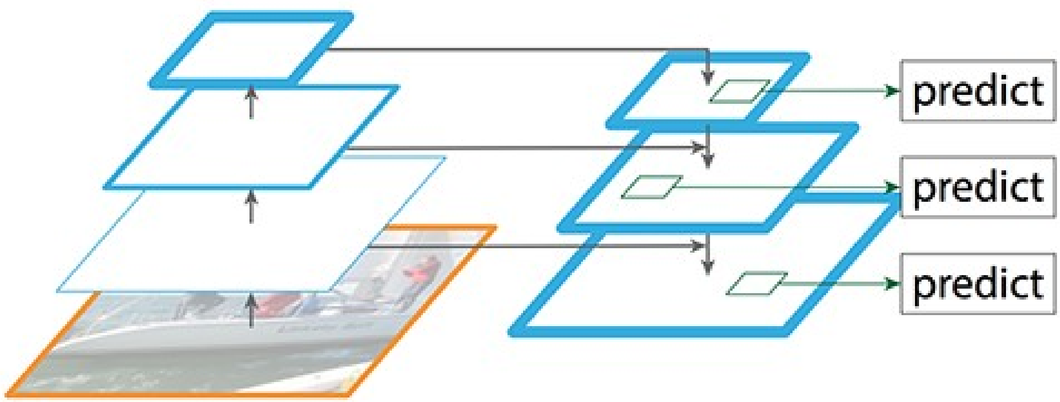


6. DISCUSSION

After going through a lot of model in past and testing and building those into an Object detection machine has gave a good idea on models use in object detection. Some basic support was driven by the site Dark net for understanding the TENSERFLOW system with the weight produced and data collected the project is success.

But meanwhile when the project was in development it was found that most of the resources were somehow connected to the model weight and openCV. A recent trail on the model shows the TENSERFLOWSS weight that we use sometime gave bad results and were not according to the desire output.

The model detect better with any data given and proceeds with same accuracy that it promise. In case study of paper REMOTE sense similar data was taken into account for the work. Our Research shows good results comparatively.



7. CONCLUSION

The project went on the good hand and delivered successive output. May end up on a worse case scenario of updates given by change in code and execution, may end up god or bad for the system. While the product seems to be a good output vs a lack of time taken the model seems to be working well against other model.

Basically all object detection framework still struggle with small objects, especially those bunched along with partial occlusions. Real-time detection with top-level classification and localization accuracy remains challenging, and practitioners must often prioritize one or the opposite when making design decisions. Video tracking may even see improvements within the future if some continuity between frames is assumed instead of processing each frame individually. Furthermore, a noteworthy enhancement which will see more exploration would extend this two-dimensional bounding boxes into three-dimensional bounding cubes.

although many object detection obstacles have seen creative solutions, these additional considerations-and plenty more signal that object detection research is in no way done!

References

[1] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014.

[2] Ross Girshick. Fast R-CNN. In International Conference on Computer Vision (ICCV), 2015.

[3] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster R-CNN: Towards realtime object detection with region proposal networks. In Advances in Neural Information Processing Systems (NIPS), 2015.

[4] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[5] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, ChengYang Fu, and Alexander C. Berg. SSD: Single shot multibox detector. In ECCV, 2016.

[6] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.