AmbujaNeotia

THE NEOTIA UNIVERSITY

ज्ञानम् आत्म प्रदीपाय

Synopsis of Mini Project
on
Real Time Object Detection

Submitted By

Subhadip Samanta (TNU2020021100001 | Data Analytics)
Wrishav Sett (TNU2020021100004 | Data Analytics)
Sudipta Saha (TNU2020021100005 | Data Analytics)
Subarna Das (TNU2020053100011 | AI&ML)

Bachelor of Technology

In

Computer Science and Engineering

Guided By

Mr. Jaydeb Mondal

Introduction

What is Object recognition?

Object recognition is a computer vision technique for identifying objects in images or videos.

Having consulted with our project guide we have decided to delve into the scope of **SSD Single Shot MultiBox Detector.**

SSD is designed for object detection in real-time.
The SSD object detection composes of 2 parts:

- Extract feature maps, and
- Apply convolution filters to detect objects.

SSD uses **VGG16** to extract feature maps.
Then it detects objects using the **Conv4_3** layer.

Limitations of SSD:

- The SSD, while boosting the performance significantly, suffers from decreasing the resolution of the images to a lower quality.
- The SSD architecture will typically perform worse than the Faster R-CNN for small-scale objects.

When To Use SSD? – The single-shot detector is often the preferred method. The main reason for using the single-shot detector is because we mainly prefer faster predictions on an image for detecting larger objects, where accuracy is not an extremely important concern.

Example use cases – The Single-shot detector can be trained and experimented on a multitude of datasets, such as PASCAL VOC, COCO, and ILSVRC datasets.

Problem Statement

In today's scenario, the fastest algorithm which uses a single layer of convolutional network to detect the objects from the image is single shot multi-box detector (SSD) algorithm.

Faster R-CNN uses a region proposal network to create boundary boxes and utilizes those boxes to classify objects. While it has the highest accuracy out of all object detection algorithms the whole process runs at 7 frames per second. Far below what real-time processing needs.

SSD speeds up the process by eliminating the need for the region proposal network. To recover the drop in accuracy, SSD applies a few improvements including multi-scale features and default boxes. SSD does not use a delegated region proposal network. Instead, it resolves to a very simple method. It computes both the location and class scores using small convolution filters. These improvements allow SSD to match the Faster R-CNN's accuracy using lower resolution images, which further pushes the speed higher.

In this project we aim to increase the classification accuracy of object detection by improving the SSD algorithm while keeping the speed constant.

We also aim to make it so that objects that have trouble being matched with sufficient success rate can be classified further by the user and be included in the preexisting database for model training and object detection for future cases.

Reason for topic selection

Object recognition is a major output of both deep learning and machine learning algorithms.

The Single Shot MultiBox Detector which is used for object detection from still images used the SSD(Single Shot Detector) algorithm which bases its computation off the CNN(Convolutional Neural Network) algorithm is one that we already have knowledge of.

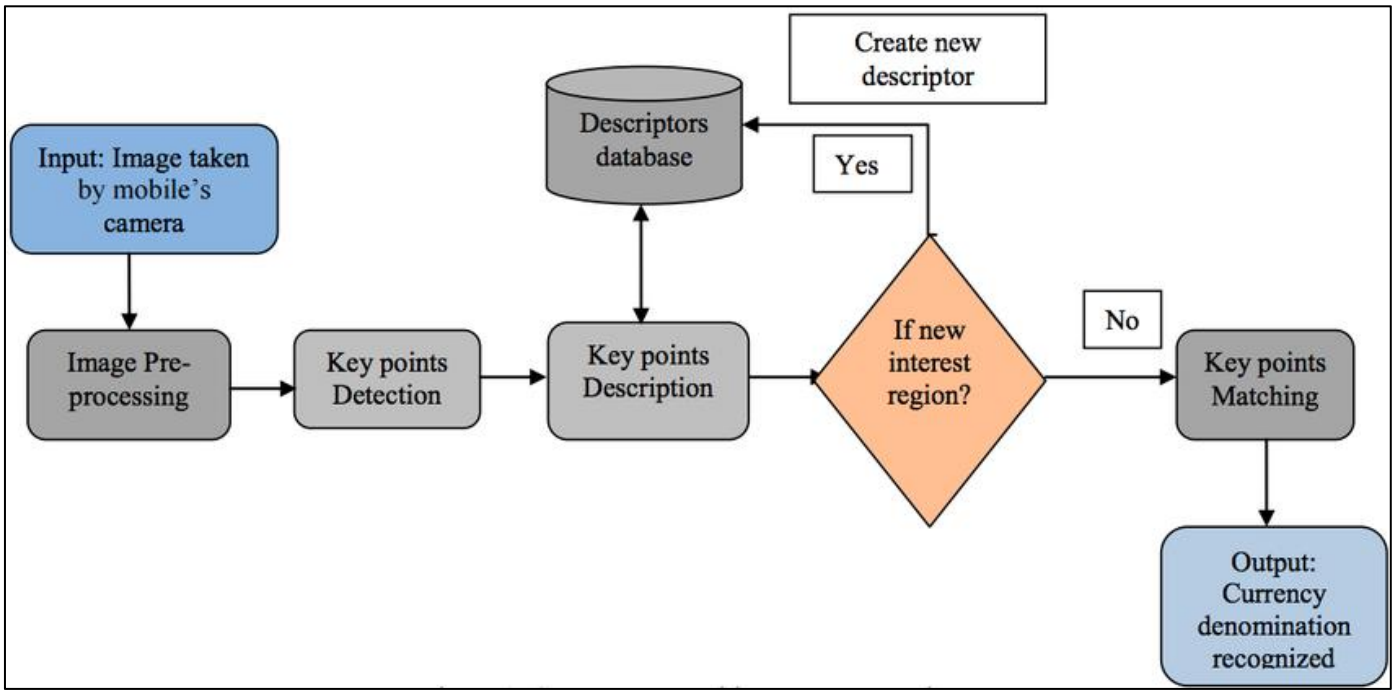
Objective & Scope

The objective of the project that we have decided to undertake is to improve on the already existing algorithm such that its computation is faster by reducing the wide variety of objects to be detected and focusing its use case scenario to particular fields of expertise.

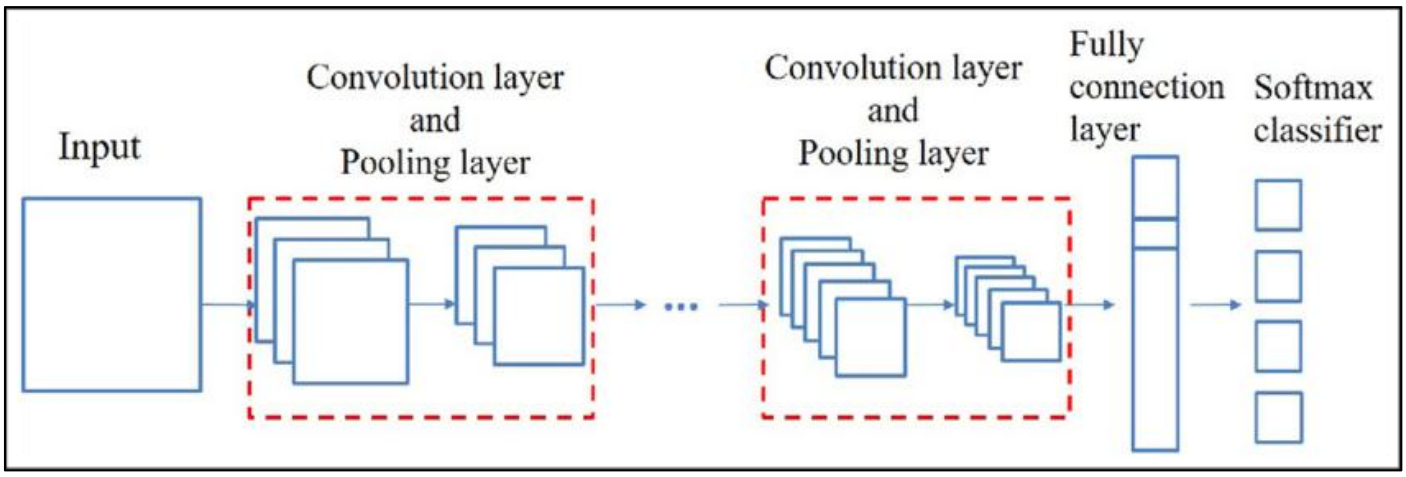
Also to make it so that objects that have trouble being matched with sufficient success rate can be classified further by the user and be included in the preexisting database for model training.

This project can be further developed for object detection through a video instead of an image or also be made so that it can be made to use the camera of a smartphone when made into a smartphone application.
It may be also be designed specifically for face detection through video survillance.

Methodology

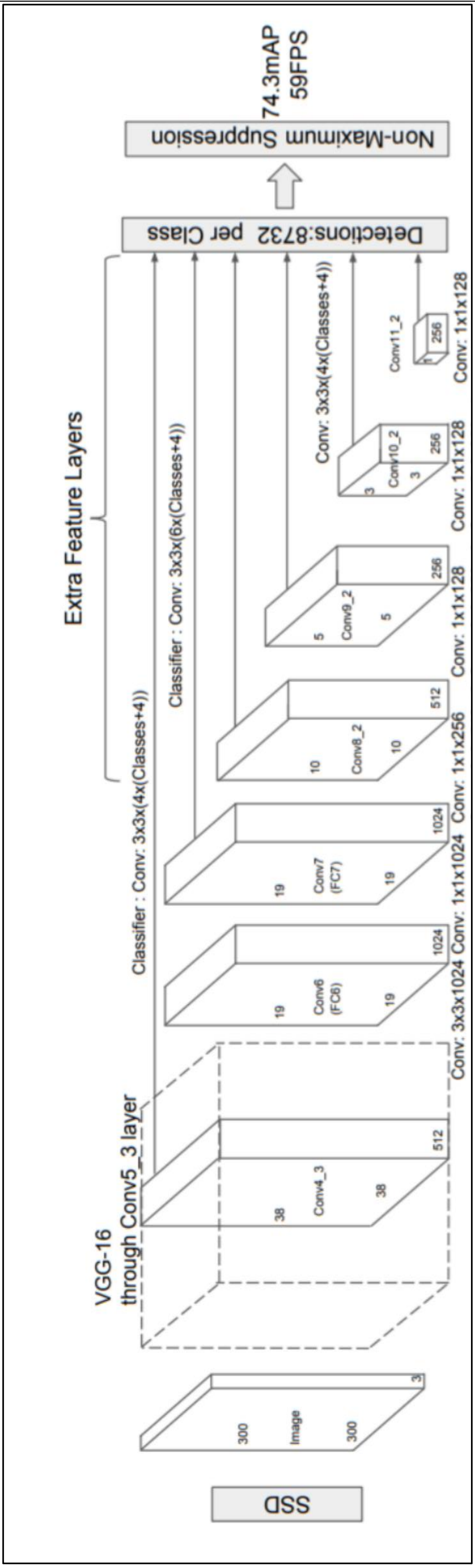


Flow Diagram for Object Detection

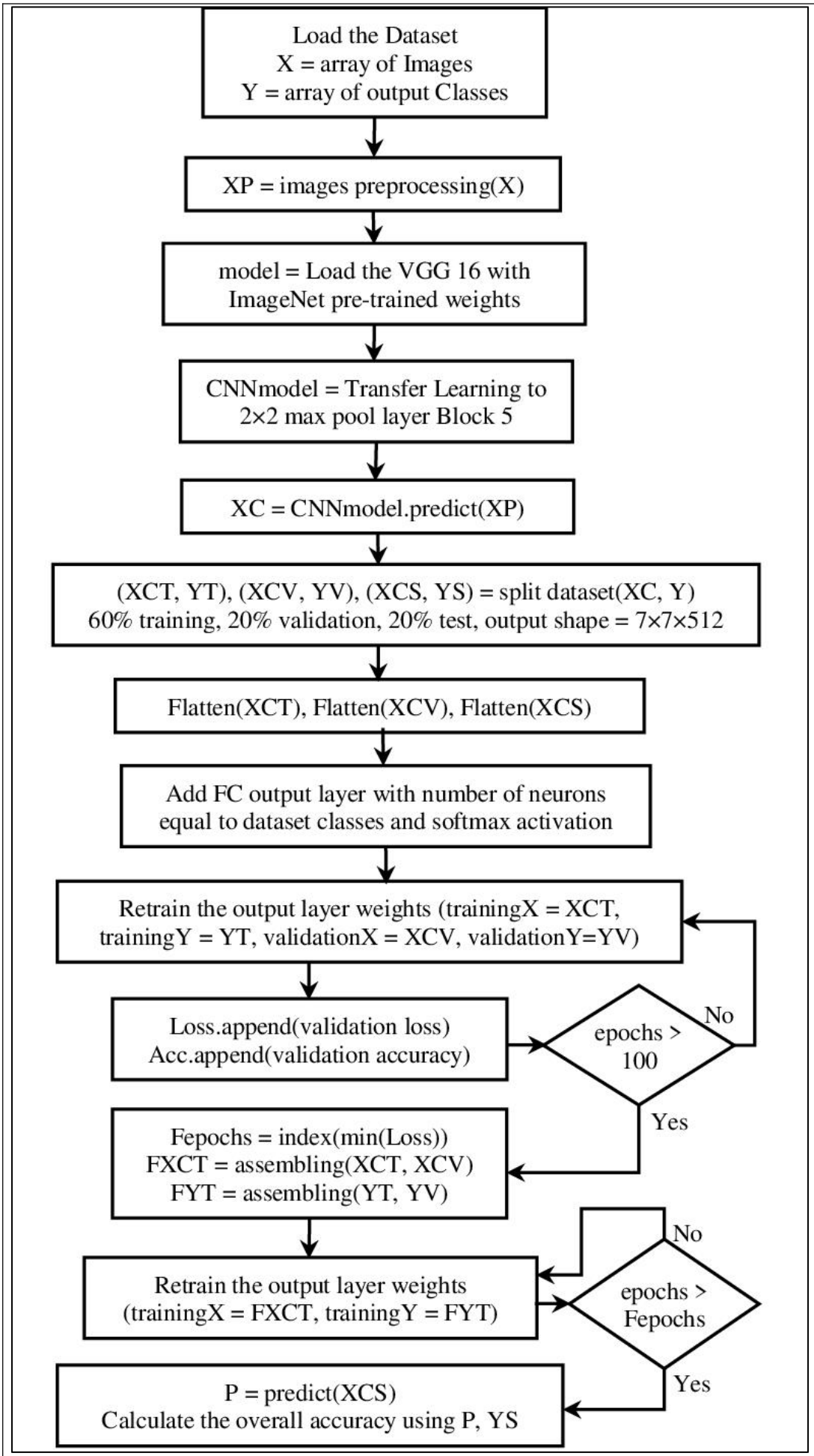


Architecture of CNN

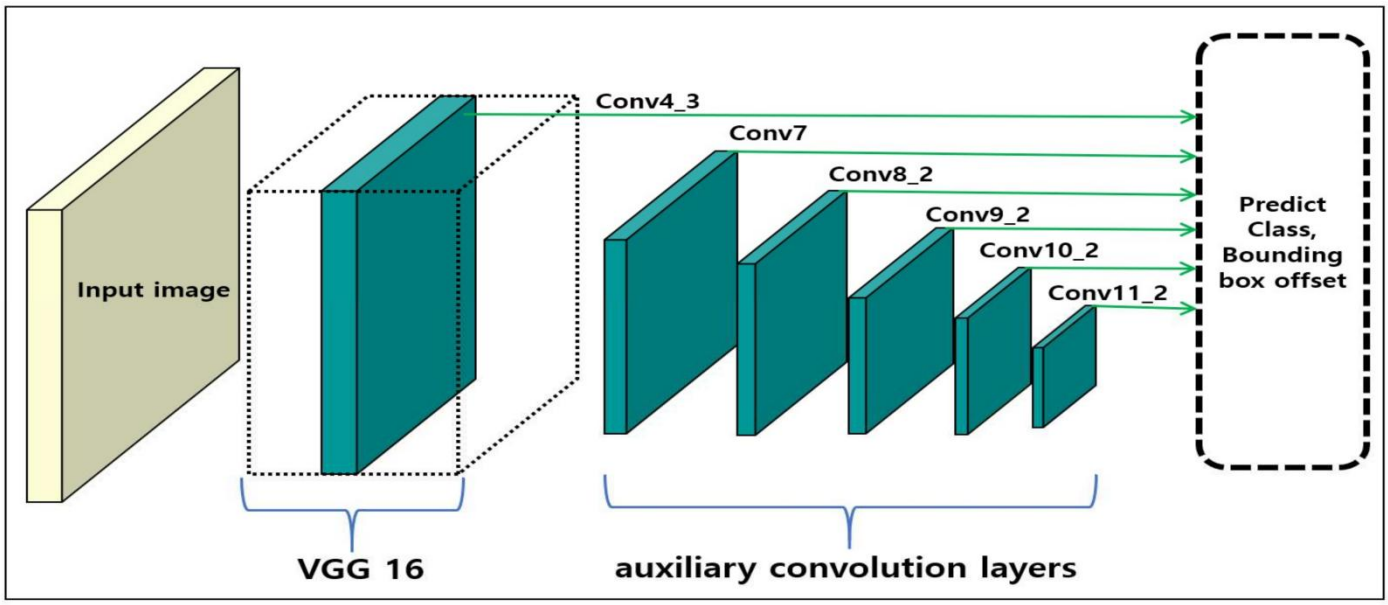
[Note: The Convolutional Neural Network is a subtype of Neural Networks that is mainly used for applications in image and speech recognition. Its built-in convolutional layer reduces the high dimensionality of images without losing its information. That is why CNNs are especially suited for this use case.]



SSD: Single Shot Multi Box Detector



VGG16 network for feature extraction



Conv4_3 layer flow diagram

Hardware and Software

For an Object Detection project development after research we have concluded that the following software are needed:

- Visual Studio
- Anaconda
- CUDA
- cuDNN
- TensorFlow
- Jupyter Notebook
- The Model(to be trained)
- Python Packages

[Note: Software needs subjected to changes.]

The following python libraries used in machine learning for image processing are also needed:

- OpenCV
- NumPy

[Note: There are multiple other libraries for image handling, however we will be using the aforementioned two as these are the ones we have hands on experience with.]

The project will be developed on the following hardware setup:

- 8GB RAM
- Intel Core i3 CPU
- 4GB NVIDIA GeForce GTX 1650 Super GPU