Object Detection And Image Recognition

Arun Kumar Dubey[1]
Anita Kumari[2]
Palak Garg[3]
Anuradha R Pai[4]
[1,2,3,4]Bharati Vidyapeeth's College of Engineering, New Delhi
arudubey@gmail.com

Abstract

With the tremendous increase in the amount of images that are being produced daily , there was a need for the development of an robust and efficient object detection system.Object recognition is the process by which objects are detected within images and videos.

Object detection can be used for various purposes including retrieval and surveillance. In this study, various basic concepts used in object detection are described while making use of OpenCV library of python 2.7, improving the efficiency and accuracy of object detection are presented.

Keywords: Object Detection,Python, OpenCV,Numpy,Haar Cascade

1. Introduction

The modern world is encircled with heavy masses of digital visual information, may it be images,videos and so on. It is a need of hour to analyse and organise these plundering ocean of visual information and for these techniques are required. Particularly,it will be more useful to analyse semantic information of the images or videos.One imperative part of image content is the objects in the image. So there is a need for object detection and image recognition techniques.

Object detection is a vital, yet difficult vision task. It is a basic part in numerous applications, for example, image search, image auto-annotation and scene understanding;

be that as it may it is as yet an open issue because of the intricacy of object classes and images. It is being widely used in industries to ease user, save time and to

achieve parallelism. Object detection is a part of computer vision which aims at having a human like vision,which can locate and differentiate various objects such as, numbers,location,size, position etc. The common object detection method is the color-based approach, detecting objects based on their color values.

Object detection is one of the most challenging applications of the image processing. It is a branch of computer vision and artificial intelligence. The aim of this project is to identify and locate the objects in the images or videos and also naming the specific objects detected.

OpenCv(python) is used for the implementation of the project. Numpy library is also required for the same. Also, an OpenCv algorithm i.e Haar Cascade is used and Haar like features are emphasized.

Computer Vision

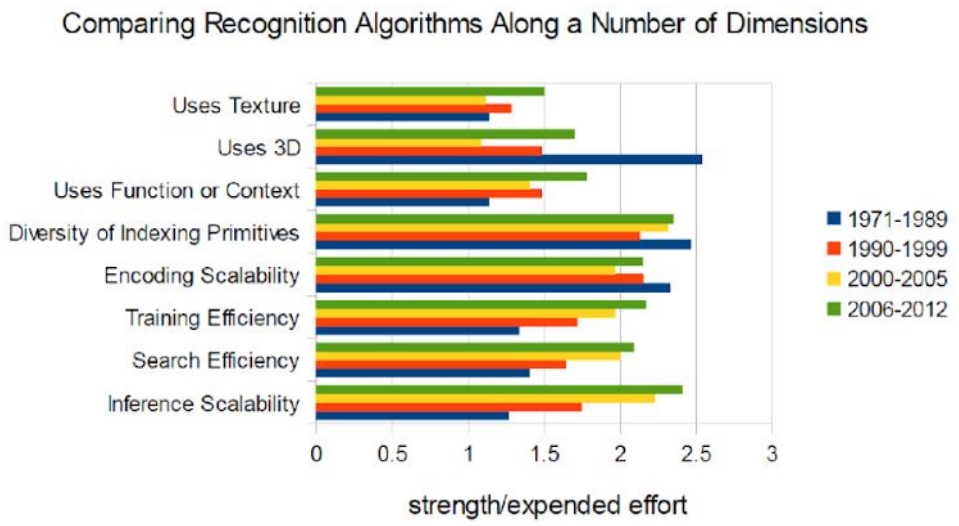
Humans use their eyes and brain to see and spot the objects around.Computer vision is the science giving similar functionality and capability to a machine or a computer. It aims at enabling computers to see,identify and process the images in the same way as human vision does, and then provide appropriate output. It resembles giving human intelligence and instincts to a computer.

2. Background

Modern computer vision approach has its starting point in early 1960s. First and foremost application was pattern recognition systems used in offices. While performing a task,some practical difficulties such as scene complexities increased as some illumination variability increased. Also time, cost and sensor noise constraints became for common.

1964 involved the automation of the wire-bonding process of transistors, with the ultimate goal of replacing human workers. However ,it attained 95% accuracy in test labs but regarded too low to replace human workers.By 1973, fully automated assembly machines had been made , resulting in the world’s first image-based machine for the automatic assembly of semiconductor devices. It is much more evolved in these 50 years and achieved almost human accuracy.

As the experiments went on and on making the systems more efficient and reducing the need of human workers, computer vision has now become an effective and efficient technology. working very much like human vision. It is now being used in various fields like face recognition,autonomous cars,robots and ofcourse object detection and recognition.



3. Description of tools

In this section, tools and methods which are used for the object detection and image recognition process are described.OpenCV and Numpy are the two python libraries which are required and used for the implementation of the same.

3.1 OpenCv

OpenCv(Open Source Computer Vision) is a computer vision and machine learning software library.It was initially built to provide a common infrastructure for applications related to computer vision and to increase the use of machine perception in the commercial products. This library contains around 3000 algorithms inside and each of them is highly optimized. It mounts real-time vision applications.

These algorithms are easily implemented in Java, MATLAB, Python, C, C++ etc. and are well supported by operating system like Window, Mac OS, Linux and Android. It is free for use under open source BSD license. Initially,this library was

written in C and this C interface made OpenCV movable to some specific platforms such as digital signal processors. Wrappers for languages such as C#, Python, Ruby and Java (using JavaCV) have been developed to encourage adoption by a wider audience.
It works well with python 2.7 but first Numpy package has to be installed.

3.2 Numpy

Numpy is an essential python package which is used for scientific computations. It supports multidimensional arrays and matrices. It contains a powerful N-dimensional array object,sophisticated (broadcasting) functions,tools for integrating C/C++ and Fortran code, useful linear algebra, Fourier transform, and random number capabilities. Not only these scientific uses, NumPy can also be used as an efficient multi-dimensional container of generic data. Therefore , it can be used to store the corresponding RGB or HSV values of the image objects in this multi-dimensional container. It is licensed under BSD license and hence, can be reused with few restrictions.

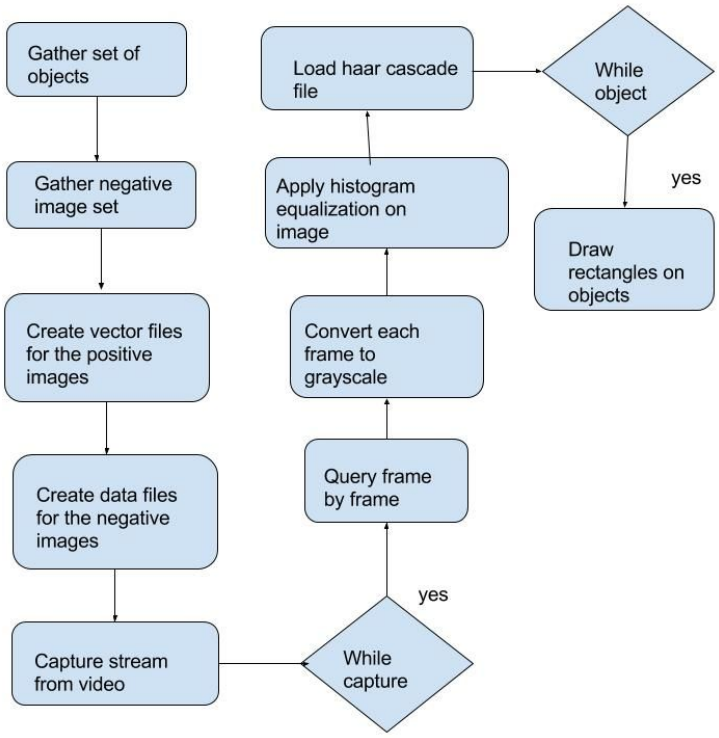
4. OpenCv(Python) Vs MATLAB

Using OpenCv in computer vision has many pros as compared to that of MATLAB.

I. **Ease of Use:** Using and learning python programming is very easy. Python is a simple language and can be learnt in a short duration of time even ,if not known before. Whereas, there is a different way of writing codes in MATLAB.
II. **Python is new scientific computing language:** Pastly, MATLAB was known as the scientific computing language but now, with libraries like OpenCV,matplotlib,Numpy it has become more efficient.
III. **Less implementation cost:** As compared to MATLAB, implementation cost of the project in python is much lesser.
IV. **Visualization and Debugging:** Visualisation with matplotlib is as effective as MATLAB . Also, debugging is much easier in python as compared to other languages.

V. **Runtime:** MATLAB has slower runtime that of Python.

5. IMPLEMENTATION OF THE MODEL



An object detection and image recognition model follows broadly three basic steps that are:

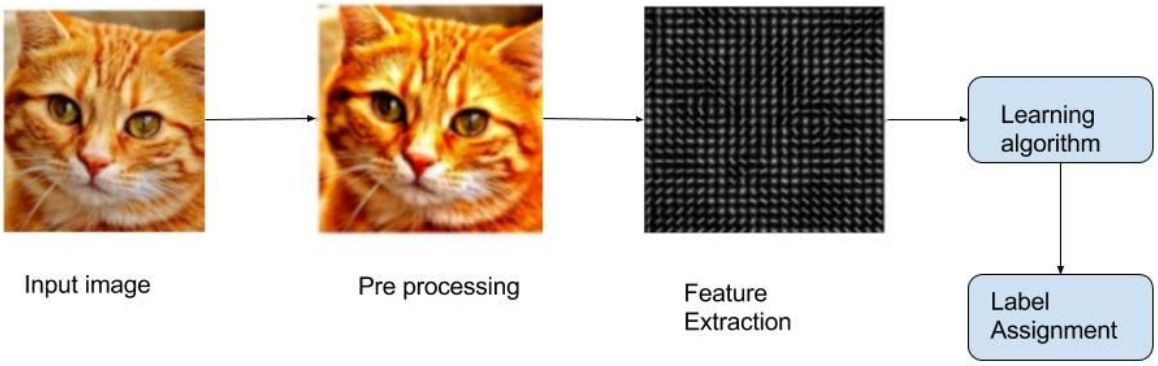
A. Image Classification
B. Detection

A. Image Classification

An image classifier follows an algorithm which takes an image as input and outputs what the image contains. The output is the image label which tells about the class to which the image belongs to(for example, cat,elephant,chair etc). But for recognising the content of the image, image classifier has to be trained well. It is ,therefore, trained by giving different set of positive and negative images to differentiate among different features and to give the desired output.

Anatomy of an Image Classifier

The following diagram shows how an image classifier works:



Step 1: Preprocessing

Firstly an input picture is pre-processed to normalize the contrast and brightness effects.Also, preprocessing step involves subtracting the mean of image intensities and divide by the standard deviation. Methods like gamma correction and colour space transformation can also be used for the same and may result in better results ,sometimes.

After that, the input image is cropped and resized to a fixed size image as it is integral part for the feature extraction.

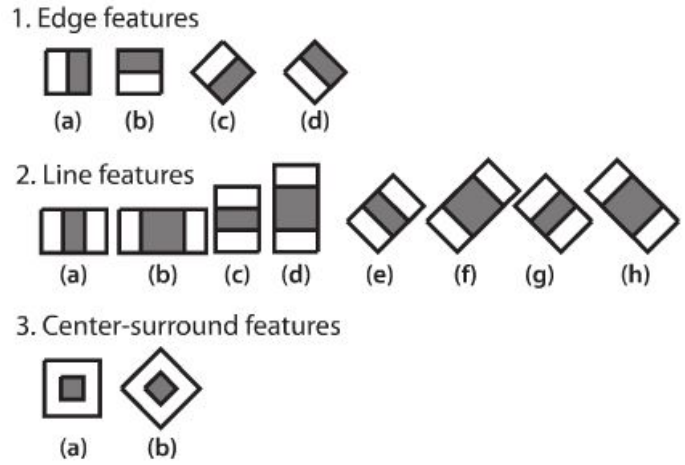
Step 2: Feature Extraction

It is very important to remove the extra information contained in the image that is not necessary for the classification. Therefore,firstly simplification of the image by extracting only important information and leaving the rest is done. It can be done by running an edge detector on the image.Some well-known features used in computer vision are Haar-like features introduced by Viola and Jones, Histogram of Oriented Gradients (HOG), Scale-Invariant Feature Transform (SIFT), Speeded Up Robust Feature (SURF) etc.

We will be using Haar-like features for the implementation of our model.

Haar-like features

It is an OpenCv algorithm based on machine learning approach where a cascade function is trained by a lot of positive and negative images. After input of positive and negative images, features are being extracted based on Haar features shown in the figure below.



Each feature is a single value obtained by subtracting sum of pixels under white rectangle from sum of pixels under black rectangle. Now all possible sizes and locations of each kernel is used to calculate overflowing features. To make is less cumbersome integral images were introduced which simplifies calculation of sum of pixels, how large may be the number of pixels, to an operation involving just four pixels. It surely makes the process super fast but also, extract many irrelevant features. So for extracting the best and important features from those 160000+ features, *Adaboost* works well.

Each and every feature is applied in all the training images and best threshold which will classify the object is found. There will surely be errors and misclassifications but we choose the one with minimum error rate, which means they are the best features which classifies the object best. The process is continued to achieve new error rates and weights to meet the accuracy.Final classifier is a weighted sum of these weak classifiers.

Now, the final setup has around 6000 features. By *Cascade Classifier* , these 6000 features are applied on different stages. If a window fails the first stage, it is discarded and if not it goes to the next stage. It saves a lot of time as comparing 6000 features at a time.

B. Detection

For the detection purpose,

I. A sliding window is performed over the image,i.e, it moves through the whole image and at different scales.

II. For each windows of the sliding window, features are extracted, computed and classified.

III. A rectangle is drawn and the name is flashed on the object being detected.

6.STEPS INVOLVED IN OBJECT DETECTION IN PYTHON 2.7

6.1 Installation of OpenCV-python

For installation of OpenCv, it is required to install the packages-python 2.7.x, Numpy,matplotlib to their default locations. Python will be installed to c/Python27/. On the terminal, import Numpy and make sure it is working well. After that OpenCv is downloaded Go to OpenCV/build/python/2.7 folder. Copy cv2.pyd to C:/Python27/lib/site-packages.

6.2 Read an Image

To read an image,CV2.imread() function is used. The image should be in the same directory,otherwise, full path of the image should be entered. Arguments specifying how an image should be read are:

1. CV2.IMREAD_COLOR: It is the default flag used to load the colors of the image. It tells about the transparency of the image.

2. CV2.IMREAD_GRAYSCALE: This function is used to present the image in grayscale mode for extraction of the features.

3. CV2.IMREAD_UNCHANGED: Loads image as such including alpha channel.

6.3 Feature Detection and Description

In this step, understanding of the features and extraction of the required features from the images is done. For this,OpenCV provides two techniques, Brute-Force matcher, and FLANN based matcher.

7. Mathematical Description

HAAR LIKE FEATURES

Haar detection Cascade uses the idea of eliminating the negative images with very little processing. Number of classifiers are computed using OpenCV to every sub-region in the image. On the off chance that the sub-region does not pass the majority of the classifiers than that picture is discarded and further calculation is performed. On the other hand if the sub-region passes the first stage then it is passed onto the next stage for further calculations. For the success of object detection, an image sub-region must pass all the classifiers. Classifier can be trained using OpenCV.

Haar-like features are an over entire arrangement of two-dimensional (2D) Haar capacities, which can be utilized to encode nearby appearance of articles.

The feature value F which has k rectangles is obtained as in the following equation:

$$F = \sum_{i=1}^{k} w(i).\mu^{(i)}$$

Where,

μ^i is average intensity of pixels in image which is enclosed by the ith rectangle.

The weights assigned to the rectangles are set to default integer values satisfying the below condition

$$\sum_{i=1}^{k} w^i = 0$$

One of the primary purposes behind the fame of the Haar-like features is that they give an exceptionally appealing trade off between speed of assessment and precision.

8.Applications

8.1 Face detection

Well known applications incorporate face recognition and individuals tallying. Have you at any point seen how facebook recognizes your face when you upload a photograph? This is a straightforward use of object detection that we find in our day by day life.

8.2 People Counting

People counting can be done with the help of Object detection. Analysing store performance or crowd statistics during festivals can also be done using object detection. These have a tendency to be more troublesome as individuals move out of the frame quickly.

8.3 Vehicle detection

Object detection is useful in detecting vehicles such as bicycle or truck. It is also effective in estimation of the object's speed. Object detection can also help in determining the type of ship entering a port. Some of the European countries are using this method for detecting ships.

8.4 Manufacturing Industry

Object Detection is additionally utilized as a part of modern procedures to distinguish items. Let's assume you need your machine to just recognize round articles. Hough circle discovery change can be utilized for identification.

8.5 Online images

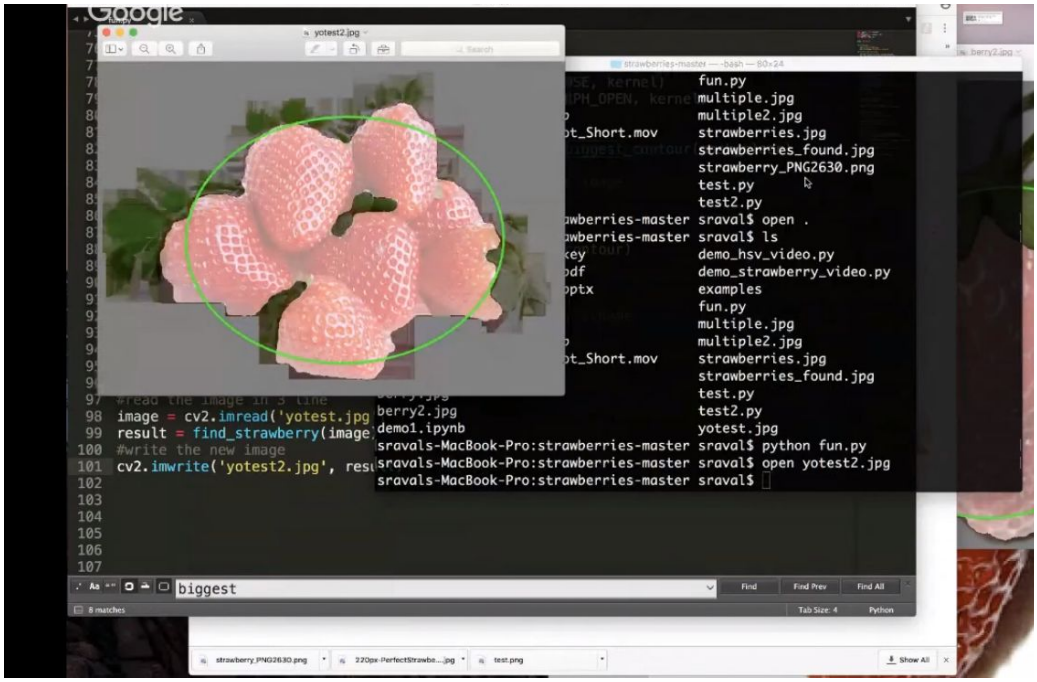
Aside from these object detection can be utilized for grouping pictures discovered on the web. Vulgar pictures are typically filtered out utilizing object detection.

8.6 Security

Later on we may have the capacity to utilize object detection to recognize abnormalities in a scene, for example, bombs or explosives (by making utilization of a quadcopter).

9. Result and Analysis

Our project detect different objects including strawberries,bike,car, duck and many more. The figure below shows the result generated after performing search for strawberries in a query image. The circled part is the final result.



10. Conclusion

There is overgrown possibilities that computer vision can be used to solve the real time problems. It has given machines an artificial intelligence and human like vision. The essentials of object detection along with different methods for accomplishing it and its extension has been
examined. Using python for the implementation ,over MATLAB is preferred because of its easiness and correctness.
Two major steps in object detection are feature understanding and matching and it is necessary that they are performed with high accuracy. Haar cascade algorithm of openCv is being used for completion of the project.
OpenCv will surely acquire an important place in the IT companies and will be a prime requirement for the coders,in coming years.