International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 05 | May 2020 www.irjet.net p-ISSN: 2395-0072

Study of Object Detection using Machine Learning

Ms. Bindu rani[1], Shubhashish Dangwal[2], Vikash Chaurasia[3], Tarun Pal[4], Virupaksh[5]

[1]Professor, [2]Student
IT Department, Inderprastha Engineering College

Abstract: *An object detection system finds objects of the real world present either in a digital image or video, where the object can belong to any class of objects namely humans, cars, etc. In order to detect an object in an image or a video the system needs to have a few components in order to complete the task of detecting an object, they are a model database, a feature detector, a hypothesiser and a hypothesiser verifier. This paper presents a review of the various techniques that are used to detect an object, localise an object, categorise an object, extract features, appearance information, and many more, in images and videos.*

I.Introduction

Object detection are divide into three parts.

1. Informative region selection

As different objects may appear in any positions of the image and have different aspect ratios or sizes, it is a natural choice to scan the whole image with a multi-scale sliding window. Although this exhaustive strategy can find out all possible positions of the objects, its shortcomings are also obvious. Due to a large number of candidate windows, it is computationally expensive and produces too many redundant windows. However, if only a fixed number of sliding window templates are applied, unsatisfactory regions may be produced.

2. Feature extraction

To recognize different objects, we need to extract visual features which can provide a semantic and robust representation. Features are the representative ones. This is due to the fact that these features can produce representations associated with complex cells in human brain . However, due to the diversity of appearances, illumination conditions and backgrounds, it's difficult to manually design a robust feature descriptor to perfectly describe all kinds of objects.

3 .Classification.

Besides, a classifier is needed to distinguish a target object from all the other categories and to make the representations more hierarchical, semantic and informative for visual recognition. Usually, the Supported Vector Machine (SVM) .

II. Literature Survey

Object detection is a very popular research direction in the vision field. Launched in 70s, object detection began to be on track until 90s when computers became powerful and application plentiful. It is easy for us as human to recognize objects in the images, however, things become difficult for computers. Adding the different posture of objects and the complex environment around, object detection is more ambiguity. As we know, the evolution of detection algorithm is divided into two stages. Stage one is based on the traditional features of the solution, and the second stage is the deep learning algorithm. Before 2013, most of the researches was based on the traditional feature optimization detection method. After that, both academia and industry turned to deep learning algorithm. With the increasing amount of detection data, the traditional detection method performance will become saturated.

III. Diagram of Object Detection Model

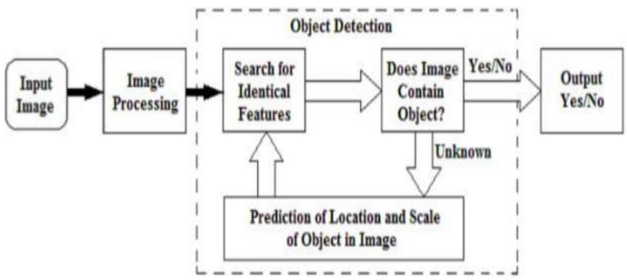

Fig 1. Basic Object Detection model

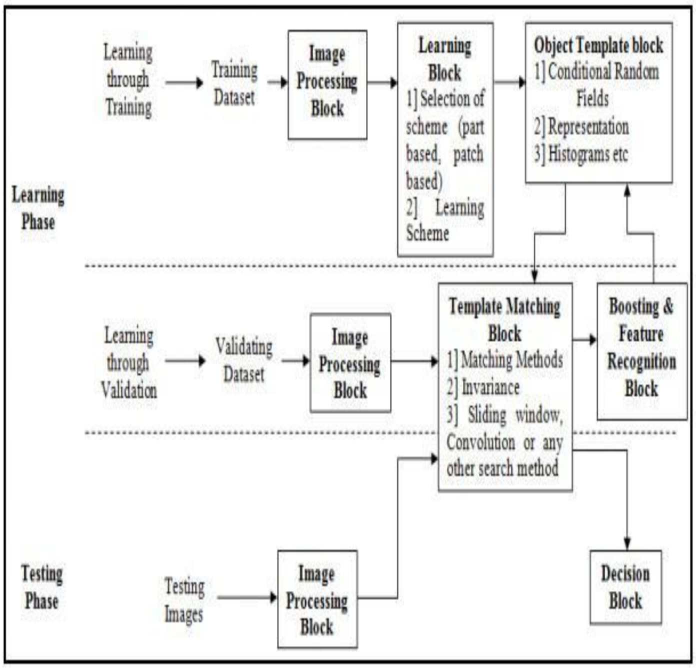

Fig 2 Working of Object Detection system

III. METHODOLOGY

A. Methods of Object Detection

1. Faster Region-based CNN

We trained the Faster Region-based Convolutional Neural Network (Faster R-CNN) model on Caffe deep learning framework by Python language. The Faster R-CNN is a region based detection method. It firstly used a region proposal network (RPN) to generate detection proposals, then used the same network structure as Fast R-CNN to classify object and modify the bounding box.

2.Conventional CV features

a) Color Histograms

After we found the bounding box that gave the outline of object, we could construct three histograms of the RGB values of all the pixels within the bounding box. The color of the object could be found combining the most frequent RGB values.

b) Histogram of Oriented Gradients (HoG) Histogram of Oriented Gradients (HoG) is a feature descriptor for object detection. It will build a histogram for localized portions of an image. The histogram is distributed along the angle of orientation (usually 8 directions) and the height is the magnitude of the gradient

c) Scale Invariant Feature Transform (SIFT) Descriptor SIFT Descriptor is also a descriptor built from histogram of gradients, however, the SIFT descriptor has the advantage of being invariant to rotation, translation and resizing. SIFT descriptor was used to extract keypoint from the image

B. Matching Algorithm

In our project, we used K-Nearest Neighbor (KNN) to match the extracted features with the database. The KNN algorithm is a basic unsupervised learning algorithm. It compares the incoming example with all the dataset and output the training example with the smallest distance from the example. Note that here the distance function can be defined by multiple ways. The simplest way would be computing the Euclidean distance of the feature vectors. Another distance definition that is useful here is the Cosine distance, defined by

$$d(x_i, x_j) = 1 - \frac{x_i^T x_j}{|x_i||x_j|}.$$

C. Problem Statement

Many problems in computer vision were saturating on their accuracy before a decade. However, with the rise of deep learning techniques, the accuracy of these problems drastically improved. One of the major problem was that of image classification, which is defined as predicting the class of the image

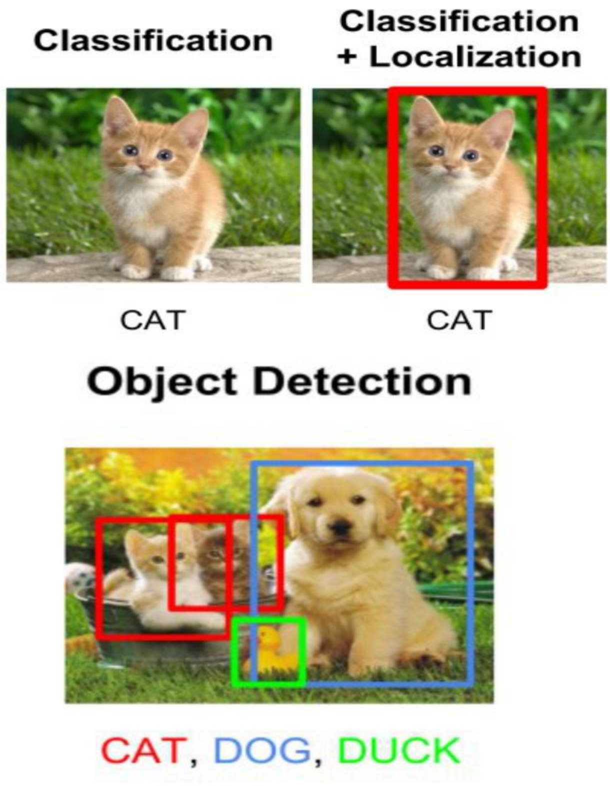

Fig 3. Sample of object Detection

IV. Experimental Results

1.Dataset

For the purpose of this project, the publicly available PASCAL VOC dataset will be used. It consists of 10k annotated images with 20 object classes with 25k object annotations (xml format). These images are downloaded from flickr.

2. Implementation Details

The project is implemented in python 3. Tensorflow was used for training the deep network and OpenCV was used for image pre-processing.

3. Pre-processing

The annotated data is provided in xml format, which is read and stored into a pickle file along with the images so that reading can be faster. Also the images are resized to a fixed size.

4.Quantitative Analysis

The detections were assigned to ground truth objects and judged to be true/false positives by measuring bounding box overlap. To be considered a correct detection, the area of overlap between the predicted bounding box and ground truth bounding box must exceed a threshold.

Fig 4. Detection results on PASCAL VOC dataset.

V. Conclusion

This article mainly established a small daily items detection data set. The data set is then trained on different object detection models and has achieved good results in daily object detection. These welltrained models can be used in the mobile platform, robot platform or other intelligent devices, the daily items to achieve accurate detection. In the future, we can get a better model by increasing the capacity of the data set, the optimization of the model structure and the fine tuning of the parameters.

VI. REFRENCES

1. Amine, K. and Farida, M.H. (2012) 'An active contour for range image segmentation', Signal & Image Processing: .

2.Girshick R. Fast R-CNN[J]. Computer Science,2015.

3.Object Detection wikipedia https://en.wikipedia.org/wiki/Object_detectionhttps://en.wikipedia.org/wiki/Object_detection

4.Object Detection : Simplified - Towards Data Science https://towardsdatascience.com/object-detection-simplified-e07aa3830954

5.A Gentle Introduction to Object Recognition With Deep Learning https://machinelearningmastery.com/object-recognition-with-deep-learning/

[6]Color histogram Wikipedia
https://en.wikipedia.org/wiki/Color_histogram

[7] Histogram of Oriented Gradients (HoG) Wikipedia
https://en.wikipedia.org/wiki/Histogram_of_oriented_gradients

[8.] Scale Invariant Feature Transform (SIFT)
https://en.wikipedia.org/wiki/Scale-invariant_feature_transform

[9] Karleigh Moore, Nathan Landman, and Alex Chumbley
https://brilliant.org/wiki/matching-algorithms/