OBJECT DETECTION

END TERM REPORT

by

KARAN AGRAWAL,FARMAN ALAM,SARVESH KUMAR,SARTHAK PANWAR

Section: K18JF

Roll no: 65,35,63,46

Department of Intelligent Systems

Lovely professional University, Jalandhar

04-2020

Student Declaration

This is to declare that this report has been written by us. No part of the report is copied from other sources. All information included from other sources have been duly acknowledged. We ever that if any part of the report is found to be copied, we are shall take full responsibility for it.

Name: Karan Agrawal

Reg no: 11808255

Roll no: 65

Name: Farman Alam

Reg no: 11803407

Roll no: 35

Name: Sarvesh Kumar

Reg no: 11808189

Roll no: 63

Name: Sarthak Panwar

Reg no: 11808233

Roll no: 46

Place: Lovely professional university.

Date: 08-april-2020.

TABLE OF CONTENTS

1.Background

The aim of object detection is to detect all instances of objects from a known class, such as people, cars or faces in an image. Generally, only a small number of instances of the object are present in the image, but there is a very large number of possible locations and scales at which they can occur and that need to somehow be explored. Each detection of the image is reported with some form of pose information. This is as simple as the location of the object, a location and scale, or the extent of the object defined in terms of a bounding box. In some other situations, the pose information is more detailed and contains the parameters of a linear or non-linear transformation. For example for face detection in a face detector may compute the locations of the eyes, nose and mouth, in addition to the bounding box of the face. An example of a bicycle detection in an image that specifies the locations of certain parts . The pose can also be defined by a three-dimensional transformation specifying the location of the object relative to the camera. Object detection systems always construct a model for an object class from a set of training examples. In the case of a fixed rigid object in an image, only one example may be needed, but more generally multiple training examples are necessary to capture certain aspects of class variability. Convolutional implementation of the sliding windows Before we discuss the implementation of the sliding window using convents, let us analyze how we can convert the fully connected layers of the network into convolutional layers. Fig. 2 shows a simple convolutional network with two fully connected layers each of shape .

2.OBJECT DETECTION

Object detection is an important task, yet challenging vision task. It is a critical part of many applications such as image search, image auto-annotation and scene understanding, object tracking. Moving object tracking of video image sequences was one of the most important subjects in computer vision. It had already been applied in many computer vision fields, such as smart video surveillance (Arun Hampapur 2005), artificial intelligence, military guidance, safety detection and robot navigation, medical and biological application. In recent years, a number of successful single-object tracking system appeared, but in the presence of several objects, object detection becomes difficult and when objects are fully or partially occluded, they are obtruded from the human vision which further increases the problem of detection. Decreasing illumination and acquisition angle. The proposed MLP based object tracking system is made robust by an optimum selection of unique features and also by implementing the Adaboost strong classification method.

2.1 YOLO – You Only Look Once

All the previous object detection algorithms have used regions to localize the object within the image. The network does not look at the complete image. Instead, parts of the image which has high probabilities of containing the object. YOLO or You Only Look Once is an object detection algorithm much is different from the region based algorithms which seen above. In YOLO a single convolutional network predicts the bounding boxes and the class probabilities for these boxes.

YOLO works by taking an image and split it into an SxS grid, within each of the grid we take m bounding boxes. For each of the bounding box, the network gives an output a class probability and offset values for the bounding box. The bounding boxes have the class probability above a threshold value is selected and used to locate the object within the image.

YOLO is orders of magnitude faster(45 frames per second) than any other object detection algorithms. The limitation of YOLO algorithm is that it struggles with the small objects within the image, for example, it might have difficulties in identifying a flock of birds. This is due to the spatial constraints of the algorithm.

2.2 Artificial Intelligence

"Artificial Intelligence is neither a new technology nor a machine". Artificial intelligence is the recognition of outcome-direction which is the rapid analysis of live data to achieve the expected goal. Outcome-directed thinking splits from the confines of the rule-directed approach that is accomplished through artificial intelligence.

4. Description

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), and a class label for each bounding box.

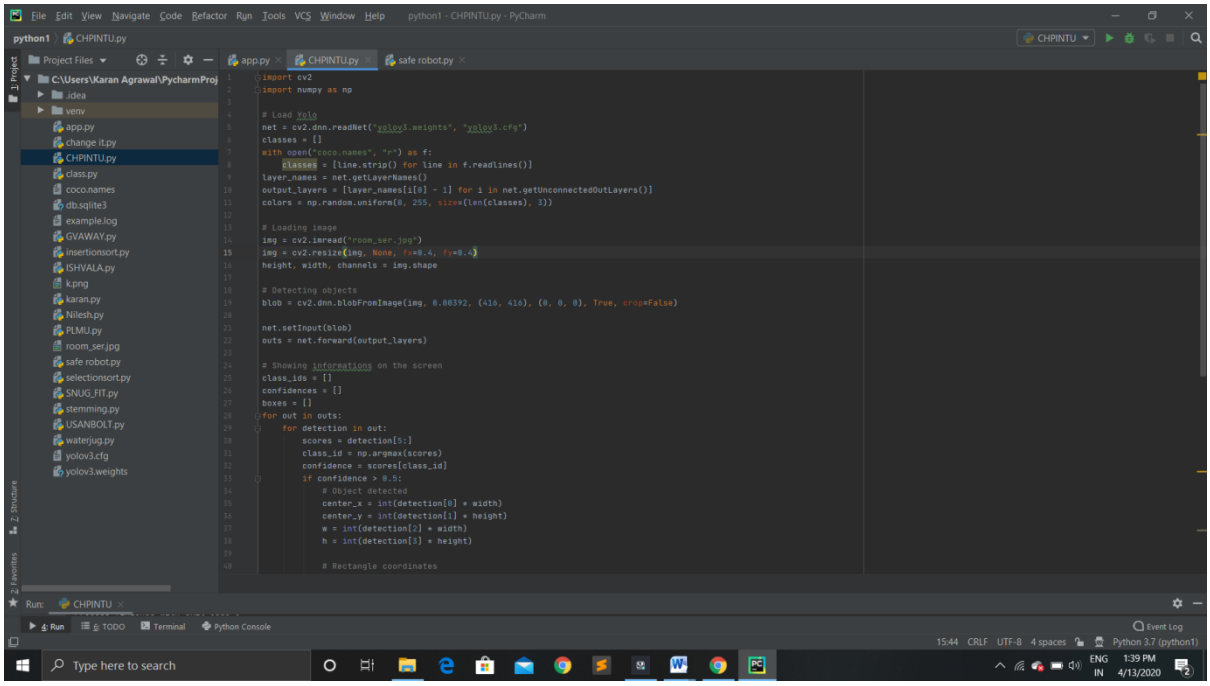
One of the further extension to this breakdown of computer vision tasks is object segmentation, also called "object instance segmentation" or "semantic segmentation," where instances of recognized objects are indicated by highlighting the specific pixels of the object instead of a coarse bounding box. From this breakdown, we can understand that object recognition refers to a suite of challenging computer vision tasks.

For example, image classification is simply straight forward, but the differences between object localization and object detection can be confusing, especially when all three tasks may be just as equally referred to as object recognition.

Humans can detect and identify objects present in an image. The human visual system is fast and accurate and can also perform complex tasks like identifying multiple objects and detect obstacles with little conscious thought. The availability of large sets of data, faster GPUs, and better algorithms, we can now easily train computers to detect and classify multiple objects within an image with high accuracy. We need to understand terms such as object detection, object localization, loss function for object detection and localization, and finally explore an object detection algorithm known as "You only look once" (YOLO).

Image classification also involves assigning a class label to an image, whereas object localization involves drawing a bounding box around one or more objects in an image. Object detection is always more challenging and combines these two tasks and draws a bounding box around each object of interest in the image and assigns them a class label. Together, all these problems are referred to as object recognition.

Object recognition refers to a collection of related tasks for identifying objects in digital photographs. Region-based Convolutional Neural Networks, or R-CNNs, is a family of techniques for addressing object localization and recognition tasks, designed for model performance. You Only Look Once, or YOLO is known as the second family of techniques for object recognition designed for speed and real-time use.


File Edit View Navigate Code Refactor Run Tools VCS Window Help
python1 - CHPINTU.py - PyCharm
python1 CHPINTU.py
CHPINTU
Project Files
app.py
CHPINTU.py
safe robot.py
C:\Users\Karan Agrawal\PycharmProj
.idea
venv
app.py
change it.py
CHPINTU.py
class.py
coco.names
db.sqlite3
example.log
GVAWAY.py
insertionsort.py
ISHVALA.py
k.png
karan.py
Nilesh.py
PLMU.py
room_ser.jpg
safe robot.py
selectionsort.py
SNUG_FIT.py
stemming.py
USANBOLT.py
waterjug.py
yolov3.cfg
yolov3.weights
import cv2
import numpy as np
# Load Yolo
net = cv2.dnn.readNet("yolov3.weights", "yolov3.cfg")
classes = []
with open("coco.names", "r") as f:
classes = [line.strip() for line in f.readlines()]
layer_names = net.getLayerNames()
output_layers = [layer_names[i[0] - 1] for i in net.getUnconnectedOutLayers()]
colors = np.random.uniform(0, 255, size=(len(classes), 3))
# Loading image
img = cv2.imread("room_ser.jpg")
img = cv2.resize(img, None, fx=0.4, fy=0.4)
height, width, channels = img.shape
# Detecting objects
blob = cv2.dnn.blobFromImage(img, 0.00392, (416, 416), (0, 0, 0), True, crop=False)
net.setInput(blob)
outs = net.forward(output_layers)
# Showing informations on the screen
class_ids = []
confidences = []
boxes = []
for out in outs:
for detection in out:
scores = detection[5:]
class_id = np.argmax(scores)
confidence = scores[class_id]
if confidence > 0.5:
# Object detected
center_x = int(detection[0] * width)
center_y = int(detection[1] * height)
w = int(detection[2] * width)
h = int(detection[3] * height)
# Rectangle coordinates
Run: CHPINTU
Run TODO Terminal Python Console
Event Log
15:44 CRLF UTF-8 4 spaces Python 3.7 (python1)
Type here to search
ENG IN 1:39 PM 4/13/2020

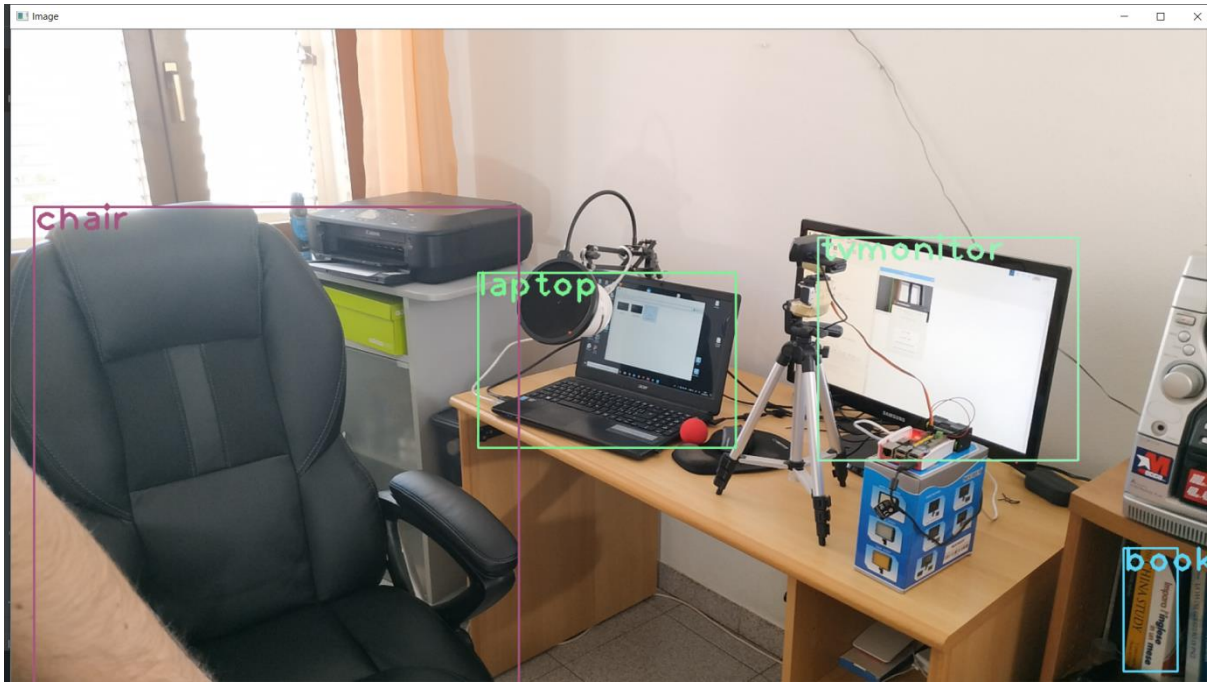


Image
chair
laptop
tvmonitor
book

4.2 Individual Student Role for chatbot

Different student studied different modules. Every one has equal contribution in coding. Karan Agrawal and Sarvesh Kumar made the report. Farman Alam and Sarthak Panwar did the testing and uploading in Git-Hub.

4.3 Technologies and Framework to be used

1. OpenCV:

OpenCV is an library of programming functions mainly aimed on real time computer vision. originally developed by Intel, it is later supported by Willow Garage then Itseez. The library is a cross-platform and free to use under the open-source BSD license.

2. Numpy:

NumPy is library of Python programming language, adding support for large, multi-dimensional array and matrice, along with large collection of high-level mathematical function to operate over these arrays. The ancestor of NumPy, Numeric, was originally created by Jim Hugunin with contributions from several developers. In 2005 Travis Olphant created NumPy by incorporating features of computing Numarray into Numeric, with extension modifications. NumPy is open-source software and has many contributors.

pip install numpy -command

3. Pillow:

Python Imaging Library is a free Python programming language library that provides support to open, edit and save several different formats of image files. Windows, Mac OS X and Linux are available for this.

pip install pillow -command

4.4 SWOT Analysis achieved in project

By using this thesis and based on experimental results we are able to detect obeject more precisely and identify the objects individually with exact location of an obeject in the picture in x,y axis.This paper also provide experimental results on different methods for object detection and identification and compares each method for their efficiencies.

The object recognition system can be applied in the area of surveillance system, face recognition, fault detection, character recognition etc. The objective of this thesis is to develop an object recognition system to recognize the 2D and 3D objects in the image. The performance of the object recognition system depends on the features used and the classifier employed for recognition. This research work attempts to propose a novel feature extraction method for extracting global features and and obtaining local features from the region of interest. Also the research work attempts to hybrid the traditional classifiers to recognize the object.

BONAFIDE CERTIFICATE

Certified that this project report " OBJECT DETECTION " is the bonafide work of " Names: SARVESH KUMAR, FARMAN ALAM, KARAN AGRAWAL, SARTHAK PANWAR" who carried out the project work under my supervision.

<<Signature of the Supervisor>>

<<Name of supervisor>>

<<Academic Designation>>

<<ID of Supervisor>>

<<Department of Supervisor>>