
REAL - TIME HUMAN DETECTION & COUNTING

PROJECT REPORT

BY

Akash Ramanand Rajak (CSE/435/19008)

Amaan Khan (CSE/438/19011)

Kumar Saurabh (CSE/470/19043)

Pallav Dubey (CSE/481/19054)



submitted to

Indian Institute of Information Technology, Kalyani

for 3rd Year Project

Bachelor of Technology

In

Computer Science and Engineering

Nov, 2021

Acknowledgments

First of all we would like to take this opportunity to thank our supervisor Dr. Anirban Lakhsman without whose efforts this project would not have been possible. We are grateful to him for guiding us towards the project wherever possible. We are most grateful to Department of Computer Science and Engineering, IIIT Kalyani, India, for providing us this wonderful opportunity to complete our 3rd year project.

And last but the biggest of all, We want to thank to each of the group members, for always helping keeping a continuous check that project never wandered off the track from our goal.

Akash Ramanand Rajak, Amaan Khan, Kumar Saurabh, Pallav Dubey

Department of Computer Science and Engineering

Indian Institute of Information Technology Kalyani

Kalyani, W.B.-741235, India.

Abstracts

This project investigates and reports benchmarks for detecting and enumerating humans through real time images, videos and camera.

This is very useful in various image processing and performing computer vision tasks. This schemes have been implemented in Python programming language, and using various tech-stacks like OpenCv, Tensorflow, etc.

Keywords : Computer Vision, Human Detection, Enumeration

Contents

List of Acronyms	5
1.) Introduction	6
1.1) Computer Vision	6
1.2) Application of Computer Vision	7
1.3) Detection & Enumeration in Computer Vision	7
1.4) Roadmap of the report	8
2.) Human Detection & Counting	9
2.1) Human Detection	9
2.2) Detection & Counting through Image	10
2.3) Detection & Counting through Video	10
2.4) Detection & Counting through Camera	11
3.) Accuracy	12
3.1) Introduction	12
3.2) Maximum Accuracy	12
3.3) Maximum Average Accuracy	12
4.) Plots	13
4.1) Plots	13
4.2) Enumeration Plot	13
4.3) Average Accuracy Plot	13
5.) Conclusion and Further Scope	14

List of Acronyms

GUI – Graphical User Interface

CNN – Convolutional Neural network

HOG – Histogram of Oriented Gradients

Chapter 1

Introduction

This chapter resembles the brief introduction about the most widely used field of study “Computer Vision”. Here talked about the various aspects and uses of computer vision, basic meaning and keywords like detection, enumeration, and discussed the roadmap to the report.

1.1) Computer Vision

Computer vision is an interdisciplinary scientific field that deals with how computers can gain high-level understanding from digital images or videos. It is most widely used field of artificial intelligence (AI) that enables computers and systems to derive meaningful information from digital images, videos and other visual inputs.



Different types of computer vision include image segmentation, object detection, facial recognition, edge detection, pattern detection, image classification, and feature matching.

Computer Vision itself is a big domain and is divided into various subdomains like scene reconstruction, object detection, event detection, video tracking, object recognition, 3D pose

estimation, learning, indexing, motion estimation, visual servoing, 3D scene modeling, and image restoration.

1.2) Application of Computer Vision

It has various different application that too in various fields. Some of them are listed below:

- Object Detection
- Screen Reader
- Intruder Detection
- Code and Character Reader
- Robotics
- Motion Analysis
- Image Restoration

There are many left to list as it is very wide topic and here in this project we have used one of the application i.e. Object Detection.

1.3) Detection and Enumeration in Computer Vision

Detection is a computer technology related to computer vision and image processing that deals with detecting instances of semantic objects of a certain class (such as humans, buildings, or cars) in digital images and videos.

For Detection process in computer vision, there are various methods and each one have different level of accuracy according to their advancement level, like is some methods that is invented in very early stage, they give more cases of false detection as compared to the advanced methods that had been discovered after that.

And here we have used the human as an entity which we are detecting our project and along with that, we are also counting the humans through image, video and camera.

1.4) Roadmap to the report

The structure of the report is as follows:

Chapter 1: It discusses about brief introduction of what computer vision is, what are there wide applications, some important keywords like detection and counting in computer vision, and finally roadmap of the report.

Chapter 2: It is based on the discussion of one of the domain of computer vision i.e. Human Detection and further is emphasize the detection of humans in real time image, video and through camera and discussed the meaning of Maximum Human Count.

Chapter 3: This chapter is based on key point of various computer vision project or basically any project i.e. Accuracy. Inside that we have discussed about Max. Accuracy we got in detection process, Max. Average Accuracy.

Chapter 4: It talks about the graphical representation results of detection process, on the basis of the data we got. In that we focussed on two basic plot that are Enumeration Plot and Average Accuracy Plot.

Chapter 5: At last, this chapter deals with a brief conclusion and further scope of this project.

Chapter 2

2.1) Human Detection

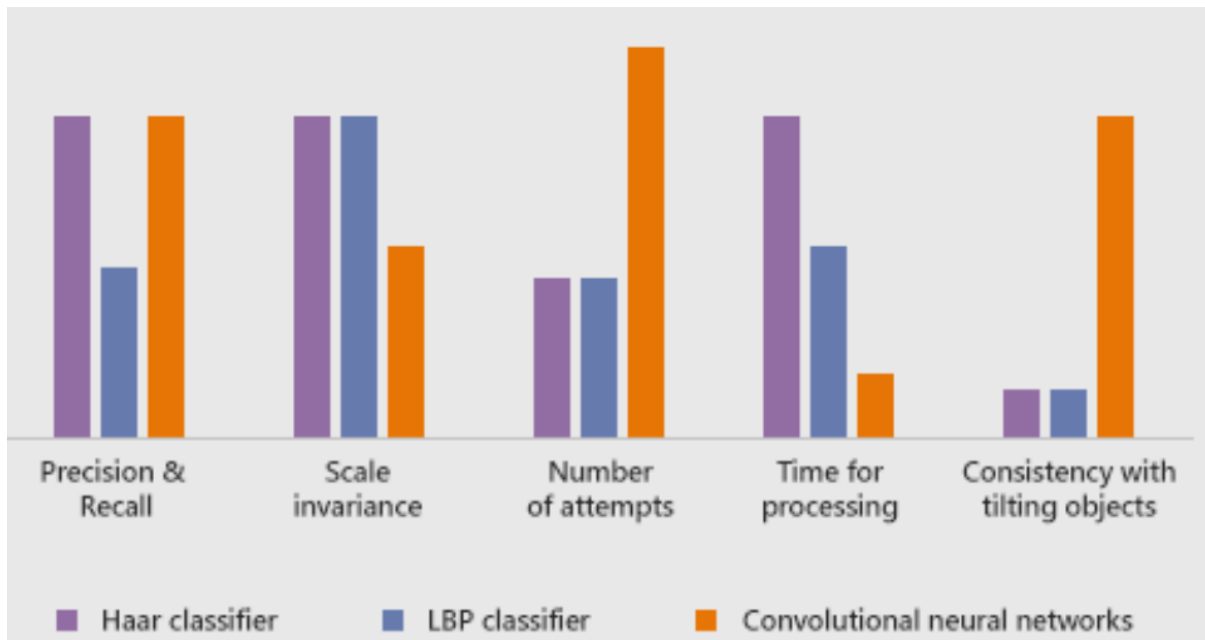
Human detection is the task of locating all instances of human beings present in an image, and it has been most widely accomplished by searching all locations in the image, at all possible scales, and comparing a small area at each location with known templates or patterns of people.

In this we can use various predefined methods and can detect the human in any image, video and can even get various factors like accuracy, each detections counting, etc.

Some common methods are :

- **Using Haar Cascade Classifier:**
 - Here we make use of .xml file for human detection, and using that we detect the humans in real time videos and images
- **Using HOG(Histogram of Oriented Gradients) :**
 - Here we make used of predefined functions and with that we detect, and this case gives some how better accuracy as compared to Harr Cascade Classifier.
- **Using Tensorflow:**
 - TensorFlow is an open-source API from Google, which is widely used for solving machine learning tasks that involve Deep Neural Networks. And again this method gives even better accuracy than above two methods.

Here we have implemented the application using the third method and got almost the better accuracy.



We have implemented the project for three different cases:

- Image
- Video
- Camera

2.2) Detection & Counting through Image

This section works with real time images. Here will allow user to select any real time image from the local system and then user can detect the humans in it. And along with that it also gives the count of humans detected.

2.3) Detection & Counting through Video

This section works with real time videos. Here will allow user to select any real time video from the local system and then user can detect the humans in it.

Now in case of video, since it is running, while the detection process is going on user will be able to see the detected peoples and their count for each frames per second of the video.

2.4) Detection & Counting through Camera

This section works somehow similar to case of video. Here user will be asked to first open the webcam, and it will detect humans that will comes in that webcam during the detection process.

Chapter 3

3.1) Accuracy

Now here we have discussed about the main keypoint of all computer vision project i.e. Accuracy. During the detection process of human, we along with process also kept track of the accuracy with each human is getting detected in image, video and camera.

In our method, we have set the threshold accuracy for the detection process as 70%, so the object detected with accuracy more than the threshold accuracy, we declared it as the well detected human, and display detection indicator around that human during process. We have set this threshold in order to prevent false detection to get displayed while detection process.

Now whenever term accuracy comes, there is always a general question, "What is the maximum accuracy of the detection?" and that we have discussed in the next topic.

3.2) Maximum Accuracy

Since we got the factor of accuracy with which each human is getting detected, so we also kept the track of maximum accuracy which we are getting throughout the detection process. This factor basically tells us about the preciseness of our implemented application.

3.3) Maximum Average Accuracy

This factor comes in the case of detecting through video and camera. Because in case of image, since the image is static, there is no meaning of maximum avg. accuracy.

In case of video and camera, it basically reflects the maximum of all the average of accuracy that we get for each frames in running

video and webcam. This factors we have used in plotting purpose, which is covered in next chapter.

Chapter 4

4.1) Plots

This section basically deals with the graphical representation of the data we got from the detection process. Using this graphical representation, one can do the analysis of the human count and accuracy very well.

In our application, we have basically talked about two basic plots.

- Enumeration Plot
- Avg. Accuracy Plot

4.2) Enumeration Plot

This plot basically represent the plot between humans count against each time interval. For this plot, the parameter we took on X-axis is time(in seconds) and on Y-axis, we took, human count at that particular time.

And the highest peak in this enumeration plot, indicates the maximum no. of people detected in whole detection process.

4.2) Average Accuracy Plot

This plot basically represent the plot between Average Accuracy against each time interval. For this plot, the parameter we took on X-axis is time(in seconds) and on Y-axis, we took, average accuracy with which humans got detected at that particular time.

And the highest peak in this plot, indicates the maximum avg. accuracy with which people detected in whole detection process.

Chapter 5

Conclusion and Future Scopes

In the last section of the project, we generate Crowd Report, which will give some message on the basis of the results we got from the detection process. For this we took some threshold human count and we gave different message for different results of human count we got from detection process.

Now coming to the future scope of this project or application, since in this we are taking any image, video or with camera we are detecting humans and getting count of it, along with accuracy. So some of the future scope can be :

- This can be used in various malls and other areas, to analyse the maximum people count, and then providing some restrictions on number of people to have at a time at that place.
- This can replace various manual jobs, and this can be done more efficiently with machines.
- This will ultimately leads to some kind of crowd-ness control in some places or areas when implemented in that area.