

Real Time Human Detection and Counting

Ms. Mohini Mohurle¹ Ms. Tejaswita Likhar² Ms. Shubhangi Hatwar³ Ms. Jayashri Sharnagate⁴
Prof. Abhimanyu Dutonde⁵

^{1,2,3,4}BE Student ⁵Head of the Department

^{1,2,3,4,5}Department of Computer Science Engineering

^{1,2,3,4,5}Abha Gaikwad-Patil College of Engineering Nagpur, India

Abstract— The goal of this project is to use OpenCv and Tensorflow to build and implement a human detector and counting system. The quantity of people in a given location at a given moment is measured using human detecting technologies. The accuracy of this measurement is totally determined by the sophistication of the equipment employed. 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. To build and implement such a system. It was attempted to assess its viability, improve its correctness, and highlight the problems encountered in its execution. The experiment was declared successful since it was able to detect and count a certain number of persons in a given location using photographs, Videos and camera. There are various vision-based algorithms for counting people. For various indoor and outdoor circumstances, each algorithm performs differently in terms of efficiency, adaptability, and accuracy. For this reason, the most widely used Haar Cascade Classifier and Histogram of Oriented Gradient based approaches are reviewed in our study with regard to many aspects such as camera orientation, lighting, occlusion, and so on. The accuracy and speed with which these algorithms performed in various settings indicates the need for more accurate and faster people counting methods.

Keywords: Computer Vision, Human Detection, Enumeration

I. INTRODUCTION

In this project discusses about brief introduction of what computer vision is, what wide applications are there, some important keywords like detection and counting in computer vision. 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.

Also, discuss the 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. Also 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.

A human counter is a gadget that counts the quantity and direction of individuals passing through a certain path or entry in a given amount of time. Many technologies are in development.

Each technology has advantages and disadvantages, and no system can guarantee 100% accuracy.

The computer vision system is used in this work, out of all the counter technologies described above.

Computer vision is a field that encompasses methods for acquiring, processing, analysing, and comprehending images and videos, as well as high-dimensional data in general, in order to generate numerical or symbolic information, such as judgments.

The goal of this field's growth has been to replicate human vision by electronically perceiving and comprehending images and videos. This knowledge of images and videos can be viewed as the disentangling of symbolic information from data utilising models built using geometry, physics, statistics, and learning theory.

II. LITERATURE SURVEY

A broad example-based framework for recognising objects in static photos by components was proposed by Anuj Mohan, Constantine Papageorgiou, and Tomaso Poggio. The method is demonstrated by creating a system that can locate people in congested environments. The method uses four different example-based detectors to recognise the components of a person's body in a photograph, such as the head, left and right arms, and hence the legs, rather than the entire body. After that, the system verifies that the discovered components are in the correct geometric configuration.

Failure of Facial Recognition in China In November 2018, Chinese police admitted to wrongfully humiliating a millionaire entrepreneur after a face recognition system used to identify jaywalkers mistakenly identified her on a billboard on a passing bus. In major Chinese cities, traffic cops utilize smart cameras with face recognition software to detect jaywalkers, whose names and faces appear on a public monitor. After the video went popular on Chinese social media, a Cloud Walk researcher speculated that the algorithm's lack of live detection was to blame.

III. RELATED WORK

A. Computer Vision

PC vision is an interdisciplinary consistent field that courses of action with how PCs can obtain certain level appreciation from mechanized pictures or accounts. 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.



Fig. 1:

Various sorts of PC vision incorporate picture division, object identification, facial acknowledgment, edge location, design recognition, picture characterization, and element coordinating.

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, movement assessment, visual servoing, 3D scene displaying, and picture rebuilding.

B. 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.

C. Detection and Enumeration in Computer Vision

Recognition is a PC innovation connected with PC vision and picture handling that arrangements with recognizing examples of semantic objects of a specific class (like people, structures, or vehicles) 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 humans through image, video and camera.

D. Human Detection

Human recognition is the undertaking of finding all cases of individuals present in a picture, and it has been most generally achieved via looking through all areas in the picture, at every single imaginable scale, and contrasting a little region at every area with known formats or examples of individuals.

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:

1) 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

2) Using HOG (Histogram of Oriented Gradients):

- Here we make used of predefined functions and with that we detect, and this case gives somehow better accuracy as compared to Harr Cascade Classifier.
- The Histogram of Oriented Gradients (HOG) approach: It's one of the most well-known human detection techniques. It's a type of feature descriptor made up of small gradient-oriented sections. This strategy works effectively when there are a lot of photos to train with, therefore it requires careful selection of different training images.

3) 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.

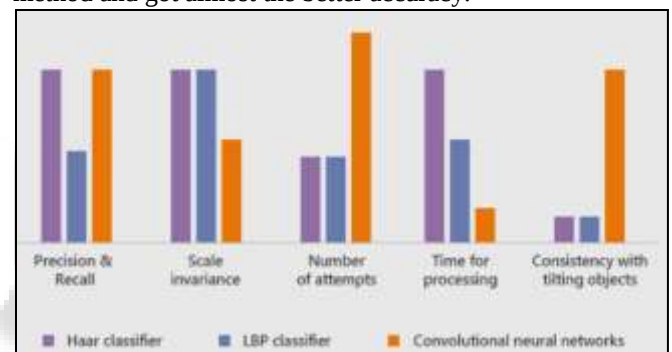


Fig. 2:

We have implemented the project for three different cases:

- Image
- Video
- Camera

E. 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.

F. 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.

G. 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.

H. 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 det 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.

I. 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 factors basically tells us about the preciseness of our implemented application.

J. 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.

K. Plots

This section basically deals with the graphical representation of the data[4][5] 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

1) 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.

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.

IV. PROPOSED ARCHITECTURE

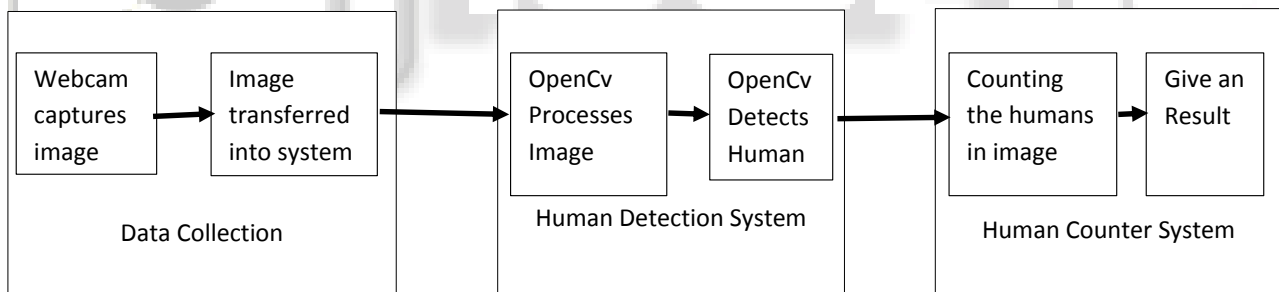
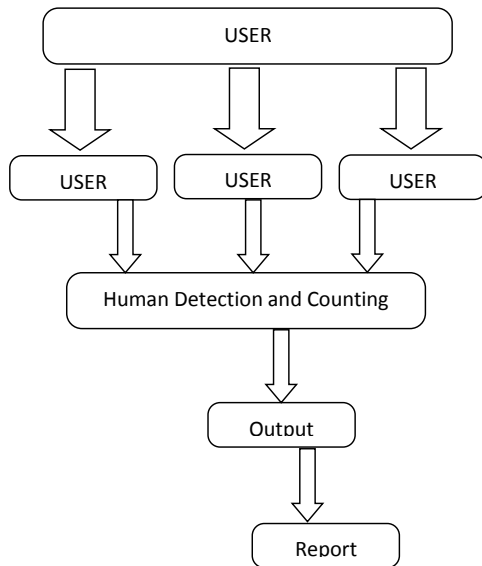


Fig. 3: Architecture of the system

In this proposed design, we talk about the framework engineering of our work. The above framework engineering shows the Human Detection and counter system. When the picture or video is caught in the webcam, the frozen casing of the video is taken care of into OpenCv for processing. After the picture is inputted, the PC vision works out the appropriated weight of the picture and distinguishes any occasions of the human. When the people in the casing are recognized, then in the third stage it will include the quantity of people in the frame. After that it will give an outcome.

V. METHODOLOGY

A. Flow Diagram



- On the starting window of the operation, user will be suitable to see START and EXIT option, using which user can start the operation or exit from the operation.
- When user starts the operation using Launch button, a new window will open, which allows user with options like, discover FROM IMAGE, discover FROM Videotape or discover FROM CAMERA.
- When user selects any of the first two option, he/ she needs to elect the separate lines using SELECT button.
- User can exercise the named train using Exercise button, and descry and count the humans using DETECT button.
- And when user selects, the last option of detecting through camera, user need to open the Camera, using OPEN CAMERA button, As soon as camera opens, discovery process will start.
- After discovery process gets completed or user manually completes it, two graph get colluded,
 - 1) Recitation Plot (Human CountVs. time) and
 - 2) Avg. Delicacy Plot (Avg. AccuracyVs. time).

Along with these two plots, an option to induce crowd report also appears, on clicking on it, a crowd report in form of PDF is generated and saved automatically at the design train position.

In the group report generated, there will be data like, what's Max Human Count, Max Accuracy, Max Avg. Fineness, and likewise a two- line status about mass.

B. Computer vision technology have several advantages.

- Exceptionally great perfection (95%)
- Customizable options.
- Fluently scalable
- Integration with other systems is simple.

C. Uses

- Calculating the chance of store callers who make purchases is needed.
- Staff shift optimization
- The number of workers needed is commensurable to the number of callers.

- When business is light, services similar as drawing and conservation are frequently performed.
- Client shadowing and line operation
- Allows for line control in workplaces, banks, and seminars.
- Statistics on callers can be tracked to assess marketing performance.

D. Security

- An meddler can be detected using a people counting/ discovery system, which can spark an alert.

VI. CONCLUSION

We induce Crowd Report, this can be used in colorful promenades and other areas, to dissect the maximum people count and also furnishing some restrictions on number of people to have at a time at that place. This can replace colorful internal jobs, and this can be done more efficiently with machines.

Which will give some communication on the base of the results we got from the discovery process. For this we took some threshold mortal count and we gave different communication for different results of mortal count we got form discovery process.

REFERENCE

- [1] Jan Eric Solem, 2012, O' Roily, first Edition, Programming Computer Vision with Python
- [2] Adrian Kaehler and Gary Rost Bradski, 2008, O' Reily, Learning OpenCv
- [3] Giancarlo Zaccane, Md. Rezaul Karim, Ahmed Menshawy, 2017, Deep Learning with TensorFlow
- [4] Alan D. Moore, 2018, Python GUI Programming with Tkinter
- [5] Fredrik Lundh, 2001, O', Python Standard Library