ISSN: 2348 - 2117

International Journal of Engineering Technology and Computer Research (IJETCR)

Available Online at www.ijetcr.org

Volume 10; Issue 3; May-June; 2022; Page No. 25-31

Real Time Human Detection and Counting

Isha Verma

(ishavermaknp@gmail.com),

Neelesh Asthana

(astha.neelesh@gmail.com),

Yash Nigam

(nigamyash456@gmail.com)

Student, Pranveer Singh Institute of Technology, Kanpur, Uttar Pradesh, India

Conflicts of interest: Nil

Corresponding author: Isha Verma

Abstract

Detecting humans in images and videos is a challenging problem, the camera and the background and to variations in pose, appearance, clothing, illumination and background clutter. We have developed a detector for standing and moving people in videos, testing several different motions coding schemes and showing the best overall performance. Use of human modelling to recognize and monitor human activity in the scene such as human walking, running etc is tracked. In addition to videos, detection from a static image can also be carried out by providing image as an input instead of a live feed from a CCTV footage respectively. Human detection in videos (i.e., series of images) plays an important role in various real-life applications (e.g., visual surveillance and automated driver assistance). The task of human detection in a series of images is challenging due to various reasons. One of these reasons is the variation of human size in the video frame. This results from changing the altitude of the platform that the camera is attached to during the task. Accuracy and short training time are the two important factors that should be taken into consideration to get a robust human, nonhuman classification system. The current Covid-19 Pandemic has altogether increased the need of such sustainable system that is Real time human detection to avoid any mishap and limit the spread of the virus, with the help of detected persons required actions can be taken by the concerned authorities respectively.

1. Introduction

Real-time human detection and counting is a vast, challenging and important field of research. In this article, we are going to build the Human Detection and Counting System with python through images provided. This is an intermediate level **Deep Learning** project on Computer Vision, which will help you to master the concepts and make you an expert in the field of Data Science. The problem of object detection and tracking in video sequences becomes much harder when the camera is allowed to move uncontrollably. If the object of interest is a deformable object, like a human, the problem becomes even more challenging. Nevertheless, several interesting applications are waiting for solutions to this problem. A modern car prepared with a digital camera and a computer vision

software for human detection can automatically avoid running over pedestrians. A military unmanned vehicle equipped with similar technology can automatically detect and deal with an enemy before being attacked. In this paper, a real-time computer vision system for human detection and tracking in uncontrolled moving camera platforms is presented. The contribution presented in this paper is not in the algorithmic aspect of the system. Rather, our focus is on the system design and implementation aspects. Namely, our design was made to achieve two main goals: robustness and efficiency. Robustness was achieved through integration of algorithms, for human detection, tracking, and motion analysis, in one framework so that the final decision is based on the agreement of more than one algorithm. Efficiency was achieved through multithreaded design and usage of a high-performance library. A final merit of our system is its object-oriented design. Object orientation was adopted to abstract the algorithmic details away from the system design so that we can easily experiment with different algorithms. Therefore, our system can be regarded as a testbed in testing different algorithms for human detection, tracking, and motion analysis. One of the simplest ways to detect targets in images is to convolve an image with a filter or template that responds to the target. The output of the convolution should produce a large response where the target is present and a suppressed response over the background. Targets are then detected where the convolution output exceeds a threshold. The primary advantages of this approach is that it is extremely simple and very fast. Computer Vision is day by day becoming important and with that human detection for applications like video surveillance, autonomous driving vehicles, person recognition have also become important. Human Detection is challenging because everyone is different in appearances and there are wide range of poses. There should be a robust method for feature extraction even when the background is cluttered. The cameras used for these applications make use of RGB cameras during night when there is deficiency of light and the images are not clear. This makes the changes in lighting conditions an

$$h(\mathbf{x}) = \operatorname{sgn}\left(\sum_{j=1}^{M} \alpha_j h_j(\mathbf{x})\right)$$

important point as well. For this, many researchers have proposed different methods for detecting humans from any image. Human detection in video is important in a wide range of applications that intersect with many aspects of our lives: surveillance systems and airport security, automatic driving and driver assistance systems in high-end cars, human-robot interaction and immersive, interactive entertainments, smart homes and assistance for senior citizens that live alone, and people-finding for military applications. The wide range of applications and underlying intellectual challenges of human detection have attracted many researchers. This paper will discuss the application of Haar-Cascade classifier method to detect the presence of human beings in thermal images. Evaluation and analysis were conducted on the performance of the method of detecting human beings in thermal images taken with the variation of human poses in the image as well as environmental conditions during image acquisition. With these variations, the performance evaluation is expected to be done comprehensively.

2. Related Works

Our work is inspired by the Viola-Jones Object Detection System, which analyses the picture in order to differentiate a face from a non-face. In this context, the object detection system uses the AdaBoost algorithm to pick the best features and train the classifiers to be used. The algorithm uses a linear combination of "weakly-approved classifiers" to create a classifier which is a good acceptable classifier.

Each weak classifier is a threshold function based on the feature fj.

$$h_j(\mathbf{x}) = \begin{cases} -s_j & \text{if } f_j < \theta_j \\ s_j & \text{otherwise} \end{cases}$$

Based on the algorithm (Viola-Jones) described above, additional references include Js Zhang and X. Xiao and their face Recognition algorithm based on multi-layer weighted LBP . In addition, references used in our work give priority to the Haar Cascade as opposed to the LBP method, which lacks the accuracy needed for the protection framework. The home safety feature of our project is influenced by N. A. Othman and I are. Aydin supports improved security work in smart homes and cities using IoT (Internet of Things). Here, multiple modules such as raspberry pi, etc. are used for image capture and analysis. Later, when a match is found, a notification is sent to the respective user via a telegram application. Another piece of work suggested by D. Hey, Pertsau and A. Uvarov is a more sophisticated variant of the Haar algorithm. Their framework modifies the GPU OpenCV algorithm. This (GPU programming) can be done using the CUDA (Compute Unified Device Architecture). The modified algorithm makes the device not only to process HD streams, In real time, quicker, but also lowers the total Processor load. Other similar works are N. It's Erdogmus and S. Marcel retrieves paper on how a face recognition device can be spoofed by using high-grade 3D face masks. The related work referred to above uses the Hair Cascade Classifier for identification and recognition of the face, but lacks any module that supports the ability to recognise the face. It is also mandatory for the protection system to function, both in terms of identification and recognition, which is available in our proposed system. Some essential literature reviews are as follows: V Latham et al. suggested the latest successful face search findings of Multiple CovNets or Deep CovNets. Recent findings, as shown by Yi Sun et.al, have shown that the methods that usually exist address the FR issue in two phases: extraction of the properties (for a better picture we must build or learn from each face of an image picture) and identification (calculates identical characteristics between the two face with a rep). Although CovNets shows positive FR performance, a well-designed architecture in CovNet remains unclear, since a specific classification project lacks theoretical guidance. Cox et al. reported that Brunelli and Poggio have determined geometrical characteristics such as mouth width and position, noise condition face size and chin form. They reported a recognition rate of 90% in a 47-person survey. However, we find that 100 per cent identification of the identical dataset is created by a rather simple method matching framework. The interrogatory database of 95 pictures of 30 manually extracted functions describing each face observed the mixing distance technique introduced by Mr Cox. Pentland et al. defines database (95 percent of 200 of the 3000 recognition) which incorporates best performance. Breakdown results are complicated since several images of the people appeared to be the same. The closest stored graph using elastic graph matching, applied a dynamic link architecture to define the distortive, invariable entity. With a sample of 87 individuals and 150 years old research images of different expressions and faces, we have seen strong findings. A 23-transuter parallel method uses a computer-cost methodology that takes around 25 years to compare 87 saved items. Eigen faces are thus an algorithm that is rapid, easy and functional. The pixel intensity for training and test images may therefore be decreased, because a high correlation is needed for the optimal performance. Another means of identifying the face is by visuals. Wikott et al. used better methods for the Facial Recognition Technology (FERET) database and contrasted 300 of the same men's sides against three hundred separate sides. You reported a

consciousness score of 97.3 percent. In restricted conditions for example local binary models and local step quantization, handmade findings have been respectable in the FR. Furthermore the efficiency is significantly reduced when used in pictures taken under uncontrolled conditions such as multiple face postures, voice and lighting. Turkaand and Pentland said that there is also a 2-dimention method for image processing. They suggested a way to classify the face portraits by projecting the principal elements of the initial trainers' photos. The resulting Eigen faces are ranked, as opposed to recognized individuals. High-level identification is usually based on multiple processing layers, for example the Marr processing framework for fitted objects from images to surfaces to 3D versions.

3. Motivation

Object detection and tracking is of utmost importance for different kinds of applications such as safety, surveillance, man-machine interaction, driving assistance system, traffic monitoring, crowd control. Due to the ongoing worldwide Covid-19 Pandemic, it is the upmost need of an hour to implement sustainable systems to provide smooth functioning.

4. Problem Statement

Automatic human detection and tracking is an important feature of video surveillance systems, web proctoring, crowd controlling systems etc. It can improve a system’s performance in fields such as security, safety, human activity monitoring etc. Human detection systems can have different goals such as detecting the presence of humans, recognition of abnormal behaviour (falls, climbing, running, etc.), identification of specific individuals, etc. The concept of human detection tracking and counting is to give a computer system the ability of finding humans precisely in images or videos. Numerous algorithms and techniques have been developed for improving the performance of human detection. Recently Deep learning has been highly explored for computer vision applications. Human brain can automatically and instantly detect and recognize humans. But when it comes to computer, it is very difficult to do all the challenging tasks on the level of human brain. The main focus in this automated detection is that human activities are recognized correctly and based on human body part motion, the human activity analysis can happen. If abnormal behaviour is detected, an alarm can be triggered. To identify individuals, face detection and recognition can also be applied.

5. Scope

Detecting human beings accurately in a visual surveillance system is crucial for diverse application areas including abnormal event detection, crowd control, web proctor, human gait characterization, congestion analysis, person identification, gender classification and fall detection for elderly people. The biggest use of such system is during the ongoing Worldwide Covid-19 Pandemic where intelligent tracking of mass gathering is of utmost importance to avoid the community spread of the disease. As the system comes with Real-time human counting, based on the statistics the task of the governing body gets reduce significantly to identify crowdy places or streets with the help of the proposed system respectively.

ALGORITHMS AND TECHNIQUES

6. Current System

The most successful and popular vector-form feature: Harr cascade It is shown that the Harr cascade features are based on the contrast of silhouette contours against the background. Despite all the difficulties on human detection, a lot of work has been done recent years. First, we

may use different features such as edge, Haar features and gradient orientation features; second, we may use different classifiers such as Nearest Neighbour, Neutral Network, SVM and Adaboost the second step of human detection is designing classifier. Large generalization ability and less classifying complexity are two important criteria for selecting classifiers. Linear support vector machine (SVM) and AdaBoost are two widely-used classifiers satisfying the criteria. So the traditional approach of AdaBoost for face detection and has demonstrated both high recognition accuracy and fast run-time performance. However, in most cases the classification accuracy is lower than that of the first proposed algorithm based on Harr Cascade.

7. Proposed System

To overcome the problems in the existing system, we shall develop an Automated detection and tracking system over manual system. Initially system consists of reading an image and detecting the existing people in this image, using Harr Cascade. Secondly the program must read a video streams recorded in the hard drive or collected by the webcam then tracking and detection is done. Initially the program carried out consists in reading an image and detecting the existing people in this image, using Harr Cascade Classifier, Harr Cascade Classifier optimized descriptor or Haar features. Secondly the program must read a video streams recorded in the hard drive or collected by the webcam or by an external camera; then make the same treatment for each video frame. We must test if the passed arguments are a static image or a recorded video stream. If there are no arguments the function directly will seek the existence of a webcam or a camera installed on the computer. Harr Cascade classifier for person detection in video: In the case of a video stream, we treat the flow components as independent images, we take their local gradients separately and we compute Harr Cascade's Classifier as in static image In the case of a video stream, we treat the flow components as independent images, we take their local gradients separately and we compute Harr Cascade's Classifier as in static image.

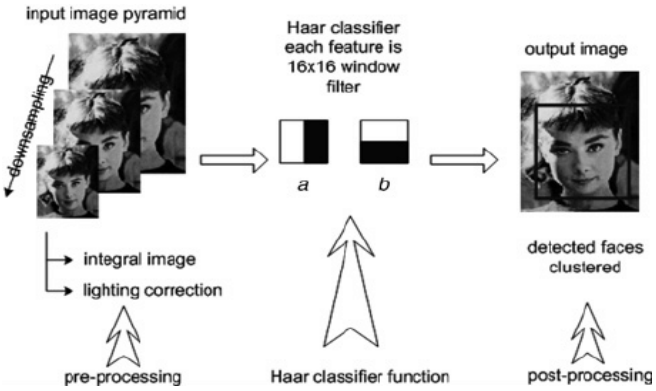

Fig 1: Algorithm Implementation

8. Algorithm

The currently available Human Recognizer Algorithms in OpenCV are:

- SIFT: Scale Invariant Feature Transform
- SURF: Speeded-Up Robust Feature
- Harr Cascade Classifier Algorithm

For our purpose, we would be using the last algorithm Harr Cascade Algorithm.

Harr Cascades classifiers are an effective way for object detection. This method was proposed by Paul Viola and Michael Jones in their paper Rapid Object Detection using a Boosted Cascade of Simple Features. Haar Cascade is a machine learning algorithm / approach where a lot of positive and negative images are used to train the classifier.

Positive Images – These images contain the images which we want our classifier to identify.

Negative Images – Images of everything else , which do not contain the object we want to detect

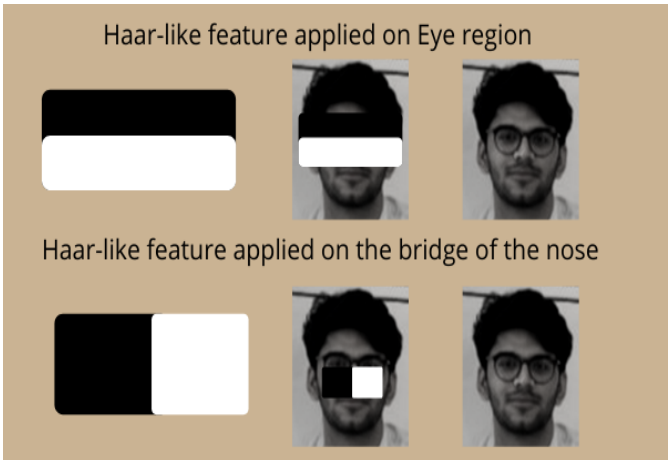

Fig 2 : Harr Classifier applied to Eyes

Haar-like features are digital image features used in object recognition. They owe their name to their intuitive similarity with Haar wavelets and were used in the first real-time face detector.

9. Conclusion

Haar Cascade Detection is one of the oldest yet powerful face and human detection algorithms invented. It has been there since long, long before Deep Learning became famous. Haar Features were not only used to detect faces, but also for eyes, lips, license number plates etc. The models are stored on GitHub, and we can access them with OpenCV methods.

Fig 3: Applying Harr Cascade to obtain results

10. Limitations

Deformable Part is a discriminatively trained, multiscale model for image training that aim at making possible the effective use of more latent information such as hierarchical (grammar) models and models involving latent three-dimensional pose.

11. Applications

In security and surveillance and crowd controlling, human presence detection are often used to detect intruders, whether it's in your personal home, an office or retail setting, or even an industrial environment. With computers, human presence detection is also used for security purposes. In this pandemic situation where number of peoples are limited to certain count by government to avoid spreading of covid counting the detected human can help to regulate the laws. The scenes obtained from a surveillance video are usually with low resolution. Most of the scenes captured by a static camera are with minimal change of background. Objects in the outdoor surveillance are often detected in far field. Most existing digital video surveillance systems rely on human observers for detecting specific activities in a real-time video scene. However, there are limitations in the human capability to monitor simultaneous events in surveillance displays.

12. References

[1] Paresh M Tank, Hitesh A Patel. Survey on Human Detection Techniquesin Real Time Video. IJCSI International Journal of Computer Science Issues Vol. 7, Issue 5, May 2018.

[2] N. Dalal and B. Triggs, "Human detection using oriented histograms of flow and appearance," in European Conference on Computer Vision (ECCV), 2006, pp. 428–441 . [

3] Mohamed E Hussein, W. Abd-Almageed, Yang Ran, Larry S. Davis. Real-Time Human Detection,

Tracking, and Verification in Uncontrolled Camera Motion Environments, 10.1109/ICVS.2006.52 Computer Vision Systems, 2006 ICVS '06. IEEE International Conference

[4] Manoranjan Paul, Shah M E Haque, and Subrata Chakraborty 2013 Human detection in surveillance videos and its applications, EURASIP Journal on Advances in Signal Processing 2013(1):25 .DOI: 10.1186/1687-6180-2013-176