EasyChair Preprint

№ 4997

People Monitoring and Mask Detection Using Real-Time Video Analyzing

Yogesh Gowari, Ritik Gaikwad, Aniket Gurav and Vinit Raut

February 21, 2021

Volume 1, Issue 4 (2021)
Article No. X
PP XX-XX

VIVA Institute of Technology
9th National Conference on Role of Engineers in Nation Building – 2021 (NCRENB-2021)

VIVA-TECH IJRI
ANNUAL RESEARCH JOURNAL

VIVA-TECH INTERNATIONAL JOURNAL FOR RESEARCH AND INNOVATION
ANNUAL RESEARCH JOURNAL
ISSN(ONLINE): 2581-7280

People Monitoring and Mask Detection using Real-time video analyzing

Yogesh Gowari[1], Ritik Gaikwad[2], Aniket Gurav[3], Prof. Vinit Raut[4]
[1,2,3,4]*(Computer Engineering, VIVA Institute of Technology, India)*

Abstract: *People Counting and mask detection based on video is an important field in a Computer Vision. There is growing interest in video-based solutions for people monitoring and counting in business and security applications using Computer Vision technology. It has been effectively used in many Artificial Intelligence fields. Compareing to normal sensor based solutions the one with video based allows more flexible performance, improved functionalities with lower costs. The system with people counter program requires more processing because that deals with real-time video, so this particular proposed technique converts a color image into binary in order to minimize data of image. Reducing processing time is an important term in Software Engineering to build a good working system. People counting methods based on head detection and tracking to evaluate the total number of people who move under an overhead camera and check whether that people are wearing a mask or not. There basically four main features in this proposed system: People counting, Mask detection, Alarm alert and Scan ID. Based on tracking of head, this method uses the crossing-line judgment to determine whether the particular head object will get counted or not to be counted. The two main challenges overcome in this system are: tough estimation of the background scene and the number of persons in merge split scenarios. A technique for masked face detection using three different steps of estimating eye line detection, facial part detection and eye detection is used in this system. On exceeding the count of people or in case mask is not worn then alarm gets alerted*

Keywords - *Convolution Neural Network, MobileNet SSD, Dataset*

I. INTRODUCTION

Public safety has become a very major problem in areas like malls, railway stations and streets during festive seasons, concerts etc. during any pandemic situation. The massive disasters that happen worldwide include numerous instances of fatality where people gather in crowds. An efficient automated system to manage the crowd count is essential. People head tracking provides a way to detect the position, to obtain the motion trail and to maintain identities of persons in the scene. Managing a crowd of varying densities involves detection of the individual humans in the crowd. In a high density crowd, because of inter-object closure, detection and tracking of humans in the crowd will be a challenge in the computer vision field. This system focuses on training a model for human head detection by some positive samples and negative samples. The trained model is then used to process the video frames in which the human heads are detected and the count of humans in the scenario is provided. It also detects whether people are wearing a mask or not. If people are not wearing then the alarm gets alerted, the same alerting happens when the number of people gathering exceeds. This system can be used in malls or any other places where crowd should be minimum.

II. RELATED WORK

Mingjie Jiang [1], In the proposed paper and high accuracy and efficient mask detector i.e.an Retina Facemask is proposed which is a one-stage detector, which consists of a feature pyramid network and a module to focus on detecting face masks. Misbah Ahmad [2], in this paper, a deep neural network model SSD(Single Shot multibox Detector) is explored to solve problems like perspective distortion, variations in human pose, size or orientation; which gives a better accuracy.

M. Martínez-Zarzuela [3], An approach for AdaBoost face detection using haar features in GPU. The GPU speeds up the performance and a better video resolution. As CPU is ideal it can perform some other computer Vision tasks. Akshay Mangawati [4], This paper elaborates the exhaustive survey of various object tracking algorithms under different environmental conditions and identifies efficient algorithms in various types of tracking.

J. Grönman[5], This paper presents a real-life use case of collecting statistics about bus passengers on a free-to-ride bus route. The use case utilized cost-effective and off-the-shelf components. Prof. P Y Kumbhar[6], In this paper the author presents that face detection is an computer technology that determines location and sizes of human faces in digital image, which detects face and ignores other things like buildings, trees or bodies. Locating and tracking human faces is a prerequisite face recognition analysis.

RafaelMuñoz-Salinas[7], This work presents a system able to visually detect and track multiple people using a stereo camera placed at an under-head position. This camera position is especially appropriated for human–machine applications that require interacting with people or to analyze human facial gestures. Tracking based exclusively on position information is unreliable when people establish close interactions Thus, we also include colour information about the people clothes in order to increase the tracking robustness. Zebin Cai [8], In this paper, we propose a people counting method in crowded scenes by detection the head information from the video taken from a camera installed straight down on the ceiling.is proposed for people detection. Combining the head detection and tracking together, a people counting strategy is presented to count the number of the people in the video frames.

Heemoon Yoon[9], Within this paper the aim is to develop an user friendly Graphical Framework for object detection API on TensorFlow which is called TensorFlow Graphical Framework (TF-GraF). The TF-GraF provides independent virtual environments according to user accounts in server-side, additionally, execution of data preprocessing, training, and evaluation without CLI in client-side.Since TF-GraF takes care of setting and configuration, it allows anyone to use deep learning technology for their project without spending time to install complex software and environment. Gretchel Karen L. Alcantara[10], In this paper, the researchers familiarizes and exposes themselves with OpenCV. OpenCV is an open source computer vision library that is written in C and C++. First the group aims to have a deeper knowledge and understanding about head detection and tracking using OpenCV.

S. Syed Ameer Abbas[11], In this paper, we propose a method to manage the crowd by keeping in track the count of the people in the scene. In our study, we develop a system using Raspberry Pi 3 board that consists of ARMv8 CPU that detects the human heads and provide a count of humans in the region using OpenCV-Python. Fabio Dittrich[12], The author presents two novel approaches for people counting in crowded and open environments that combines the information gathered by multiple views.

III. METHODOLOGY

The Crowd Monitoring and Mask Detection is a simple system used for people counting and detection of mask in crowded places. This system uses Convolution Neural Network (CNN), which is an image classification algorithm as well as MobileNet SSD which is used for the same. CNN is made up of neurons, each having an independent weight assigned to it. CNN is a class of deep neural networks specially used for image recognition and image processing. MobileNet is a simple but efficient and not very intensive convolutional neural network for mobile vision applications. MobileNet is widely used in many real-world applications which include fine-grained classifications, object detection, face attributes, and localization. CNN takes the input as an image, identifies and assigns priority to various features of the image and it differentiates the features from one another. Mobilenet is a neural network that is used for classification and recognition whereas the SSD is a framework that is used to realize the multi detector. Only the combination of both can do object detection. SSD can be interchanged with RCNN. The preprocessing required for CNN is less and has the ability to learn image characteristics. CNN consists of several sets of convolution layers, pooling layers, flatten and dense. The sets of convolution and pooling layers are used for feature extraction and the number of such sets may vary. Convolution layer is the basic building block of the CNN and is used for extracting features from an input image. The proposed system uses Convolution model which consists of multiple layers for the purpose of feature extraction from the image. Training data is provided to the model for better prediction of people wearing

a mask or not. The classification of people wearing mask, the input video is converted into frames and then into RGB format and then is flattened in matrix to extract the information by convolution layer. Multiple convolutional layers used to provide better predictions with higher accuracy. Figure 1 represents the mask detection system flow using CNN with MobileNet algorithm is used in this system as it consumes less data processing time. The testing of the module is done using real time images of people with masks and no mask to reflect the accuracy of the model. Hence, the model classifies the real time people counting and masks detection in an efficient way.

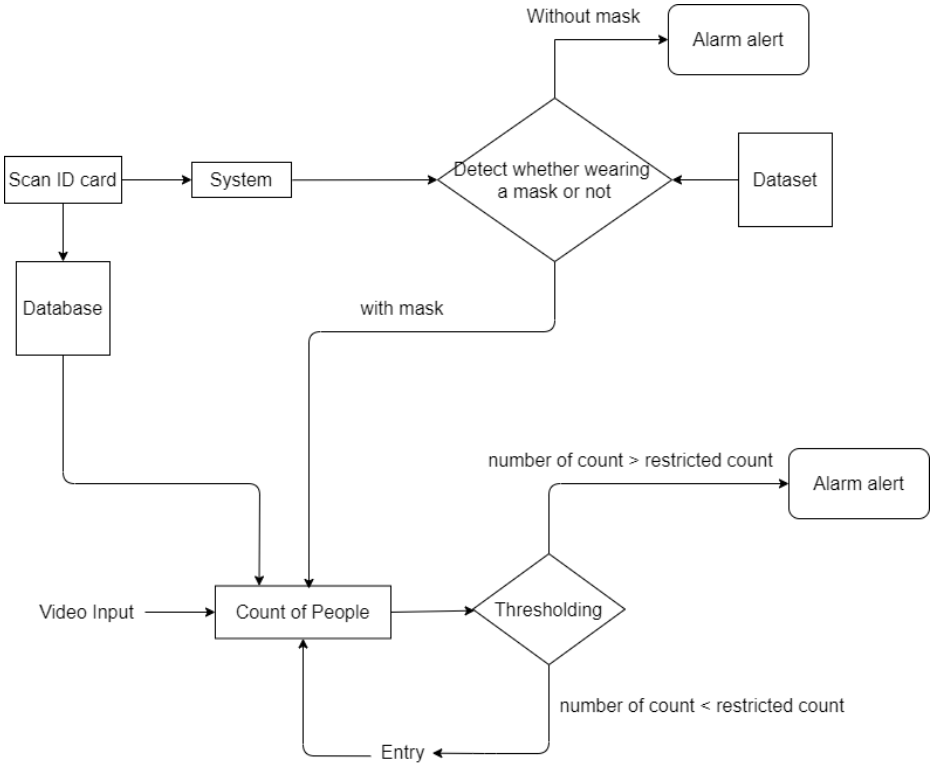

Figure 1: System Flow Diagram

IV. CONCLUSION

The mask detection using CNN with MobileNet algorithm is used in this system as it consumes less data processing time. This System presents a people counting system as a way to manage crowds by keeping the count of people. Keeping in mind the Pandemic situation Mask-Detection feature is added if the count exceeds the prohibited count or if the model recognizes whether people are not wearing masks then the alarm gets alerted. This system will reduce the time taken for humans for counting or checking purposes and ensure them, this work is done by the system itself in no time. By this model human errors will be reduced to great extents as the system itself gets trained through large datasets. This process requires comparatively less time and provides great accuracy. As the system trains itself by doing the same tasks of mask detection so that there is less loss and provides a better accuracy. As this system is still under progress so we can't predict accurate accuracy but it offers better accuracy.

REFERENCES

[1] Mingjie Jiang, Xinqi Fan ―RetinaMask: A Face Mask detector,2020, 7th International conference on Artificial Intelligience, IEEE, 2020, pg.9.

[2] M. Ahmad, I. Ahmed, K. Ullah, I. Khan, A. Khattak and A. Adnan, ―Energy efficient camera solution for video surveillance‖, International Journal of Advanced Computer Science and Applications, vol. 10, no. 3, IEEE, 2019, pg.2.

[3] Mario Martínez Zarzuela, Francisco Javier Díaz-Pernas, Miriam Antón-Rodríguez, ―AdaBoost Face Detection on the GPU Using Haar-Like Features‖, Proceedings of the 4th international conference on Interplay between natural and artificial computation: new challenges on bioinspired applications-Volume III, IEEE, 2018, pg.9

[4] Akshay Mangawati; Mohana; Mohammed Leesan; H.V. Ravish Aradhya ―Object Tracking Algorithms for Video Surveillance Applications‖, 2018 International Conference on Object detection, and motion sensor, IEEE, 2018, pg.6.
[5] J. Grönman; P. Sillberg; P. Rantanen; M. Saari;" People Counting in a Public Event—Use Case: Free-to-Ride Bus",IEEE,2019.
[6] Prof. P Y Kumbhar1 , Mohammad Attaullah2 , Shubham Dhere3 , Shivkumar Hipparagi:" REAL TIME FACE DETECTION AND TRACKING USING OPENCV",2019.
[7] RafaelMuñoz-SalinasaEugenioAguirrebMiguelGarcía-Silventeb "People detection and tracking using stereo vision and color",2007.
[8] Zebin Cai; Zhu Liang Yu; Hao Liu; Ke Zhang "Counting People in Crowded Scenes by Video Analyzing",IEEE,2014.
[9] Heemoon Yoon, Sang-Hee Lee, Mira Park," TensorFlow with user friendly Graphical Framework for object detection API",2020.
[10] Gretchel Karen L. Alcantara; Ivan Darren J. Evangelista; Jerome Vincent B. Malinao; Ofelia B. Ong; Reginald Steven DM. Rivera "Head Detection and Tracking Using OpenCV",IEEE,2018.
[11] S. Syed Ameer Abbas; P. Oliver Jayaprakash; M. Anitha; X. Vinitha Jaini "Crowd Detection and Management using Cascade classifier on ARMv8 and OpenCV-Python",IEEE,2017.
[12] Fabio Dittrich, Luiz E. S. de Oliveira, Alceu S. Britto Jr. and Alessandro L. Koerich "People Counting in Crowded and Outdoor Scenes using a Hybrid Multi-Camera Approach",IEEE,2019.