Age and Gender Detection using OpenCV

Mr.Suresh Samdade
Diploma in Information Technology
MAEER's MIT Polytechnic
Pune, India
sureshsamdade123@outlook.com

Prof. Pallavi Nehete
Information Technology Department.
Dr. Vishwanath Karad MIT World Peace University,
School of Polytechnic & Skill Development, India, Pune.

Mr.Pavan Patil
Diploma in Information Technology
MAEER's MIT Polytechnic
Pune, India pavanpatil5698@gmail.com

Mr.Roshan Maske
Diploma in Information Technology
MAEER's MIT Polytechnic
Pune, India
roshanmaske1323@gmail.com

Prof.Mrunal P. Fatangare
D V K MIT World Peace University
School of Polytechnic & Skill Development,Pune
mrunal.fatangare@mitwpu.edu.in

Mr.Vilas C. Rathod
DVK World Peace University
School of Polytechnic and Skill Development, Pune.

Abstract--- **In this fast-emerging world Artificial Intelligence plays a very vital role in every field of science . Everything is being automated from operating a remote to driving a car using Artificial Intelligence. We show a glimpse of such automated experience with this project. In this project we show how easy it is to detect faces and identify gender along with gender with the help of CNN(Convolutional Neural Networks) and OpenCV. Using these fields of Artificial Intelligence, we can reduce the use of hardware components and complexities in this project. Along with CNN and OpenCV we use Adience dataset so that the output is achieved with accurate values in training and validation. For the output to be determined even with multiple parameters we use pretrained model that is caffee model along with OpenCV. The proposed model can be used in surveillance purposes or in medical purposes.**

Keywords--- Face detection, OpenCV, Haar Cascade, Video surveillance, Security purpose, etc.…

I.INTRODUCTION

Artificial Intelligence (AI)is a computing technique which imitates human brain for the actions that are performed. These actions can be performed by the AI algorithms with the assistance of Machine Learning(ML) and Deep Learning(DL) algorithms. In order to be able to make decisions/predictions human-like, the model is required to be trained and then verified to decide the outputs. Testing is done to validate over what it has learnt at the training and verify the functionality. Based on input data, the neural network can use the algorithms of machine learning to improve accuracy. Machine learning algorithms like

Regression, Classification for Supervised Learning and Clustering for unsupervised learning etc. can be used which help to improve the model's efficiency and accuracy as a supporting algorithm for the output prediction to the main model being developed. The output prediction depends on the present inputs for those algorithms[1,2]. Deep Learning improves the overall performance and the efficiency of the model which has to detect characteristics of the person like age and gender by developing a neural network[3,4]. The model being developed can be used for surveillance purposes. Deep learning's neural networks forms the basis for the entire model and then entire decision making process is done by the neurons of the neural network. The main objective of the paper is to determine the parameters like the age, gender of the person by using the model being developed. It is makes it easier for the sake of the video analytics, for medical purposes for the surveillance purposes and it can be achieved by the use of the computer vision.

II.MOTIVATION

In this section we provide the age and gender classification literature and briefly describe about few early methods which are most related to our proposed method, focusing on age and gender detection. Many early methods in age and gender detection were handcrafted, focusing on manually engineering the facial features from the face. To mention a few, in 1999, Kwon and Lobo[5] developed the very first method for age estimation focusing on geometric features of the face that determine the ratios among different dimensions of facial features. These geometric features separate babies from adult successfully but are incapable of distinguishing between young adult and senior adult. Hence, in 2004, Lanitis et al. [6] proposed an Active Appearance Model (AAM) based method that included both the geometric and texture features, for the estimation task. This method is not suitable for the unconstrained imaging conditions attributed to real-world face images which have different degrees of variations in illumination,

expression, poses, and so forth. From 2007, most of the approaches also employed manually designed features for the estimation task: Gabor[7], Spatially Flexible Patches (SFP)[8], Local Binary Patterns (LBP)[9,10], and Biologically Inspired Features (BIF)[11]. In recent years, classification and regression methods are employed to classify the age and gender of facial images using those features. Classification methods in [12,13-15] used Support Vector Machine (SVM) based methods for age and gender classification. Linear regression[16,17], Support Vector Regression (SVR)[18], Canonical Correlation Analysis (CCA)[19], and Partial Least Squares (PLS)[20] are the common regression methods for age and gender predictions. Dileep and Danti [21] also proposed an approach that used feed-forward propagation neural networks and 3-sigma control limits approach to classify people's age into children, middle-aged adults, and old-aged adults. However they all were incompetent when given large datasets therefore, cannot be relied on to achieve respectable performance in practical application.
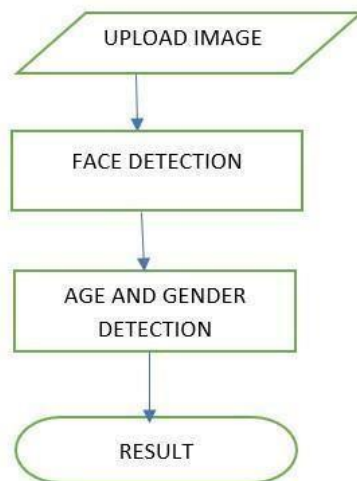
A.   PROPOSED MODEL

Figure 1: Flowchart of our proposed model

In our proposed model we use CNN and Opencv for facial recognition. This proposed model can detect faces, divide into Male/Female based facial features, divide an image with face of a person into one of 8 age ranges. Convolutional neural networks (CNN): There are various neural networks available which can be used as per the requirement or inputs being given. They have 3 main layers are input layer, hidden layer(s) and the output layer. Each layer has large number of neurons where each is associated with a certain value of weights. The values of the weights are updated at the time of forward and backward propagations, along with the help of an activation function at every layer/neuron in order to activate them. Updating the weights, governs the overall accuracy of the neural network model, as the cost/loss function is reduced to a minimum value, at a certain point in the gradient descent[3,4].
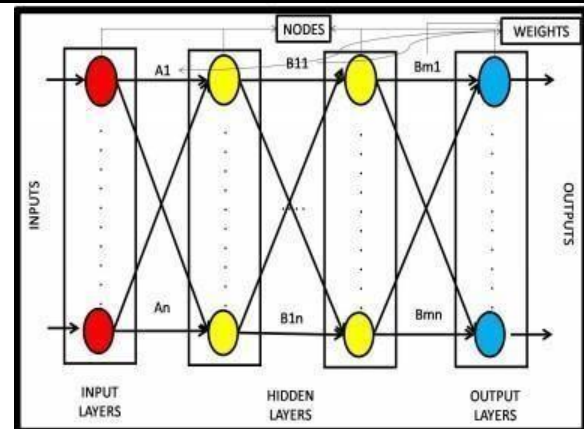

Figure 2: A Basic Neural Network

Artificial Neural Network (ANN) which is used to process the images are known as the Convolution Neural Networks (CNN). The 3 convolutional layers in convolutional neural network are:

● Convolutional layer: 96 nodes, kernel size 7
● Convolutional layer: 256 nodes, kernel size 5
● Convolutional layer: 384 nodes, kernel size 3[22]

It has 2 fully connected layers, each with 512 nodes, and a final output layer of softmax type. It is used for the features to be extracted every time when the convolutions are done. From the input image, a particular region is selected and then convolutions are done upon the intensity values of the pixels when the image is segmented. The convolutions are done in a matrix, wherein matrices of same dimensions are used for the convolutions across rows and columns on same input dataset with some dimensions. As the convolutions are completed in the convolutional layer with some kernel size, the data is given to the max pool layers to reduce the dimensions of the matrix so as to be able to do the computations on the large set of values. The data is sub sampled initially and after the max pooling by the help of strides, optimizing the neurons connections or by zero padding[23,24].
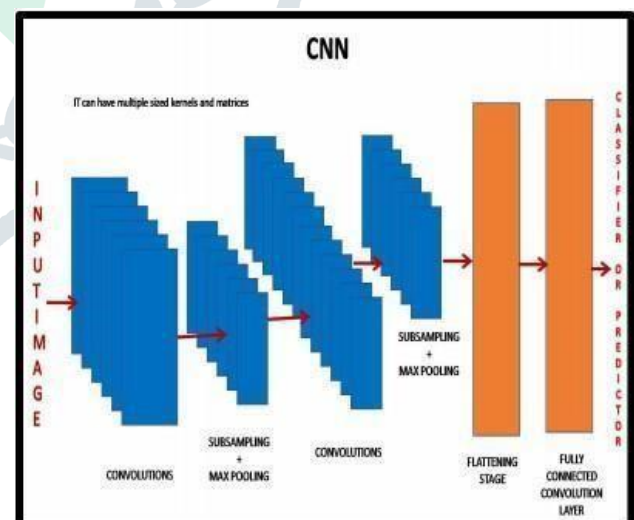

Figure 3: Convolutional Neural Network (CNN)

B.   EXPERIMENTAL SETUP

In this section, we describe all the elements used in the experiment to explore our proposed model approach in age and gender detection. This includes the dataset description and implementation of the proposed model.

Dataset

OIU-Adience is a collection of face images from real-life and unconstrained environments. It gives all the features that are anticipated from an image that is collected from various real -world scenarios etc are facial images that were uploaded to Flickr website from smart phone without any filtering. Adience images, therefore, display a high-level of variations in noise, pose, saturation, brightness and appearance, among others. , entire collection of OIU-Adience dataset is about 26,580 face images of 2,284 subjects and with an age group label of eight comprising 0–2, 4–6, 8–13, 15–20, 25–32, 38–43, 48–53, and 60+ [25].
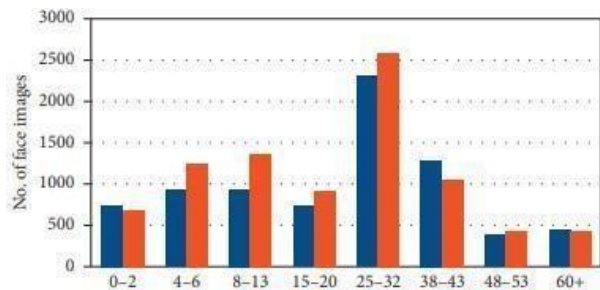

Figure 4: Age-groups for male and female (in years)

OpenCV

OpenCV (Open Source Computer Vision Library) is a library of programming functions used for image processing. It is available for free of cost at Berke;y Software Distribution License. This library has 2500 algorithms which can be used to identify objects,recognize human faces, etc. OpenCV was started at Intel in the year

1999 by Gary Bradsky. It has interfaces for Python,Java and C++. OpenCV-Python is the python API for OpenCV. OpenCV-Python is not only fast but is also easy to code and deploy [26]. This makes it a great choice to perform computationally intensive programs. Packages for standard desktop environments (Windows, macOS, almost any GNU/Linux distribution)

- run(pip install opencv-python)if you need only main modules.
- run(pip install opencv-contrib-python)this gives other modules including main modules.

Face detection

For facial recognition a protocol buffer file can be used which has all the trained weights of the model. The protobuf files with .pb extension hold data in binary format whereas the files with .pbtxt hold data in text format. These can be used to run the trained model. These protobuf files also involve in age and gender detection for our model. These are tensor flow files.

Gender and Age detection

CAFFE Model: CAFFE (Convolutional Architecture for Fast Feature Embedding) is a deep learning framework, originally developed at University of California, Berkeley. It is open source, under a BSD license. It is written in C++, with a Python interface. Caffe supports types of deep learning concepts related in the fields of image classification and image segmentation. It supports CNN and fully connected neural network designs.Caffe supports kernel libraries such as NVIDIA, CNN and Intel MKL.In this project caffe model helps us define the internal states of the parameters of the layers[27].

Protocol Buffer Files: Protocol Buffers (Protobuf) is a free and open source cross-platform library. They are used for data serialization. These are tensorflow files which are used to describe the network configuration. The protobuf files are written in xml which has .pbtxt extension. Whereas the files with .pb extension contain data in binary format which is hard to read. Google developed Protocol Buffers for internal use and provided a code generator for multiple languages under an open source license.These Protocol Buffers were designed with an aim for simplicity and better performance. Also were aimed to be faster than XML. However these are used at Google to store and interchange various kinds of data. Also used for many intermachine communication.

C. EXPERIMENTAL RESULTS

This section specifies the results obtained by conducting the experiment. We depict the testing results on various conditions.

IV. LITERATURE SURVEY

A new architecture for face image classification named unsupervised CNN was introduced by S. U. Rehman et al. [2]. A CNN that handles multitask (i.e. Facial detection and emotional classification) is made by merging CNN with other modules and algorithms. A hybrid deep CNN and RNN (Recurrent Neural Network) model was introduced by N. Jain et al. [4]. This model aims to improve the overall result of face detection. MI Facial Expression and JAFFE dataset were used to evaluate the model. A convolutional network architecture was proposed by G. Levi et al. [5] that classified the age with small amounts of data. The Audience Benchmark was used to train the model. A system in which a real time automatic facial expression system was designed was proposed by S. Turabzadeh et al. [6]. It was implemented and tested on an embedded device which could be the first step for a specific facial expression recognition chip for a social robot. MATLAB was first used to build and simulate the system and then it was built on an embedded system. The hardship of performing automatic prediction of age, gender and ethnicity on the East Asian Population using a Convolutional Neural Network (CNN) was explored by N. Srinivas et al. [3]. A fine-grained ethnicity has predictions based on a refined categorization of the human population (Chinese, Japanese, Korean, etc.). Previous results suggest that the most critical job is to predict the fine-grained ethnicity of a person, followed by age and lastly gender. An automated recognition system for age, gender and emotion was presented by A. Dehghan et al. [7] that was trained using deep neural network. At the ImageNet LSVRC-2010 contest, A. Krizhevskyetal. [8] presented a paper which suggested segregation of 1.2 million images into 1000 different categories with the help of a deep Convolutional neural network. The results which were obtained suggested that supervised learning can deliver exceptional accuracies. Some datasets have annotations on the face images which are not considered to be of any use for face recognition. Some papers have also used RNN but it is not applicable for our project as the RNN takes text or speech as an input

whereas we required an image to be as the input. Hence, CNN is chosen over RNN for the sake of our project. Some papers also suggest the use of unsupervised CNN, but, for this project supervised learning is more appropriate. The UTKFace dataset is used as dataset for the project.

V.METHODOLOGY/ WORKING

We are going to purpose the safe working by using the automatic face detection machine in which we can tract the data not only of person but also with the background of the human. persons expression. clothes, personality and background combinations can also detect the victims data . For measuring the system work there are four real time demos.

- Real time web demo system
- Comparison demo system
- Training a classifier system
- Real time Sphere Visualization system

In order with creating new directory of images and then processing it into raw form of code is main task to open cv where the main work is to classify the model and depict the structure of images in videos Video and clips should be only in mp.4 format other .filetype can cause error in loading the data and it will show the output as system destroy .
We have to first build the record of people and then create directory path for the video and process our data in format to create directory to find image in video.

The image should be bright and colourful which will be similar to identify in video.
Blur and Xerox copy pictures are a source of corrupt data format when the machine will process the data the image will not be visible and then the data entry in output remain blank and it will show no result . Lots of people have pose variation and pose confliction in video so here we have kernals and open cv process demo models which will give independent data .
We have setup various outstanding designs using open cv model which will show various animated data on time.
We can implement the correct time and place of victim.
Just instead of watching whole video clip of cctv just sort the image of criminal you have doubt on and the process will find that where the victim is criminal or not . It is employed in video survelliance, human laptop interface and image management.
Video process has become a significant demand in current world. this method is majorly wont to sight, acknowledge and track numerous objects. Face sighion and trailing is that the part wherever we tend
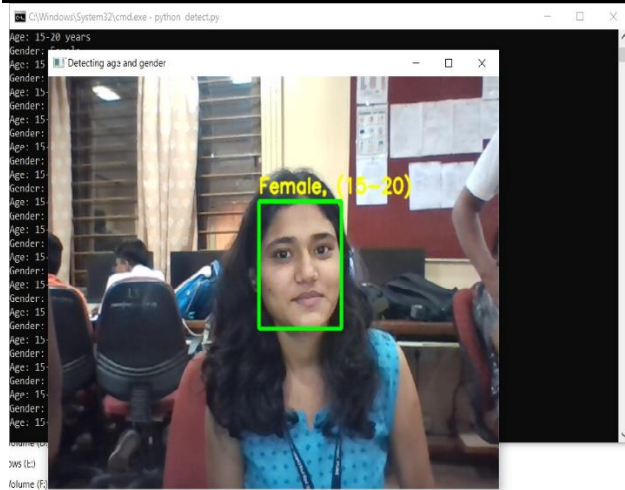
to detect someone's face from a video sequence and track him/her throughout the video. It plays important role in video corrections, police investigation, military trailing therefore on.
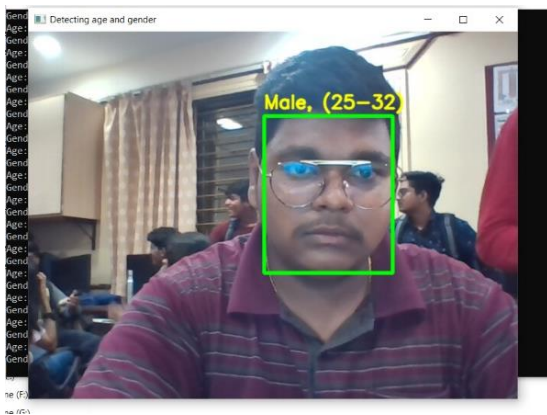
VI. SYSTEM ARCHITECHTURE
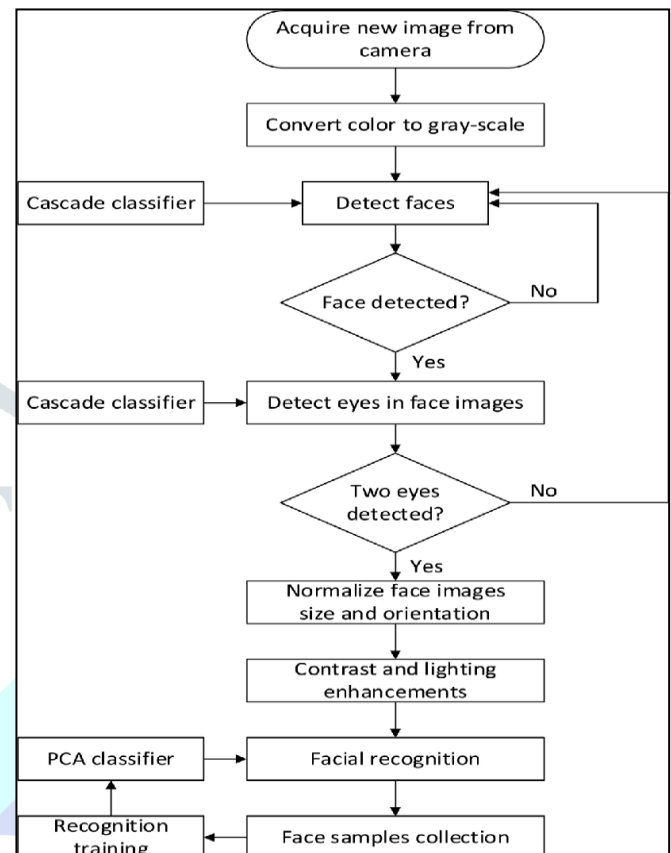
(Fig. 6 system architecture)
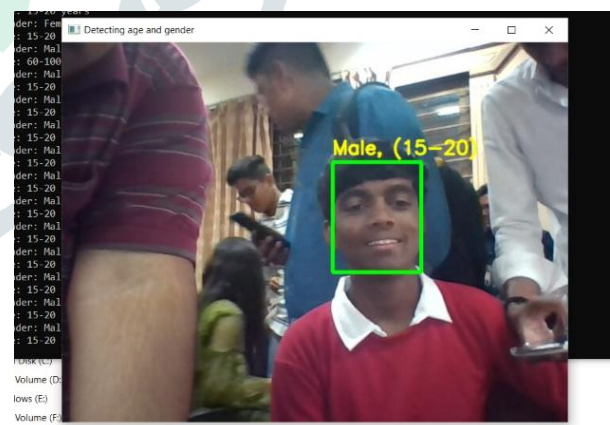
(Fig 7. Gender and Age detected using image)

(Fig 8. Gender and Age detected using image(False Detection))

VIII. FUTURE SCOPE

This project can be enhanced in few ways such that this project can be used to its fullest:

1) Application- The project can be developed into a web application or a mobile application such that it is easily accessible.

2)In public places- using sensors this can be used in public places like restaurants, ATM places, shops such then when a theft happens the scope of finding the person could be much more easy.

3) Enhancing this project to detect multiple individuals-this project can be enhanced such it can estimate age and gender even for a group of individuals in the image. This model does detect the face of individuals in a group but cannot give the accurate age and gender estimation.

IX. CONCLUSION

In this project we've used Python and OpenCV because it is simple to do code and simple to understanding also. Face recognition is an useful technology nowadays, that can provide many benefits to every industry. Face recognition can save resources and time and even develops new income streams, for companies that implement it right. The main goal of this project is to detect face in video or image and track person which we want as victim in it.

XI. REFERENCES

[1]   Van der Aalst, W. (2016). Data Science in Action. Process Mining, 3– 23. doi:10.1007/978-3-662-49851-4_1 .

[2]   Sanchez-Pinto LN, Luo Y, Churpek MM, Big Data and Data Science in Critical Care, CHEST (2018), doi: 10.1016/j.chest.2018.04.037 .

[3]   Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda. International Journal of Information Management, 48, 63– 71. doi:10.1016/j.ijinfomgt.2019.01.021 [4] Salehi, H., & Burgueño, R. (2018). Emerging artificial intelligence methods in structural engineering. Engineering Structures, 171, 170–189. doi:10.1016/j.engstruct.2018.05.084 .

[5]   Y. H. Kwon and N. Da Vitoria Lobo, "Age classification from facial images," Computer Vision and Image Understanding, vol. 74, no. 1, pp. 1–21, 1999.

[6]   A. Lanitis, C. Draganova, and C. Christodoulou, "Comparing different classifiers for automatic age estimation," IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics), vol. 34, no. 1, pp. 621–628, 2004.