IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.114126

Gender and Age Detection Using Artificial Intelligence In Python

T. Veena1, B. Lokesh2 , A. Sanjay3

[1]Assistant Professor, Department of Information Technology, S.A Engineering College, Chennai, India

[2,3]UG Scholar, Department of Information Technology, S.A Engineering College, Chennai, India

Abstract: The continuous progression of AI models for classification and Facial recognition has gained a lot of attention and importance these days and have immensely constituted in finding solutions for complex real-life issues. The Gender and Age Prediction is a Deep Learning application and falls under the area of Artificial Intelligence. Age and gender are considered as key attributes because they play a foundational role in social interactions. Estimating age and gender based on image(s) is considered as a crucial task in smart applications. The gender is expected to be classified into one of 'MALE' and 'FEMALE' however estimating age accurately using regression is considered as a monotonous job as even humans can't accurately predict the age by looking intently at a person. However, still it can be determined whether a person is in their 20s or in their 30s. through an approach namely Age Prediction as grouping and classification task using Audience Dataset, it consists of images labelled with subsequent age groups [(0 – 2), (4– 6), (8 – 12), (15– 20), (25 – 32), (35 – 43), (45 – 53), (60 –100)] and gender labels 'MALE' and 'FEMALE'. Convolution neural networks (CNN) are extensively being used for classification and facial recognition because of its exceptional efficiency in these tasks. "OpenCV" is an abbreviation for open-source computer vision. It's also a Machine Learning Library, capable of processing real-time image, video and supports the Deep Learning frameworks -Tensor Flow, Caffe, and PyTorch.

Keywords: Facial recognition, Convolution neural networks, Artificial Intelligence, Classification, Regression, OpenCV, Adience Dataset, Machine Learning Library, Tensor Flow, Caffe, and PyTorch.

INTRODUCTION:

The continuous progression of AI models for classification and Facial recognition has gained a lot of attention these days and have immensely constituted in finding solutions for complex real-life issues. The Gender and Age Prediction is a Deep Learning implementation to faces and falls under the area of Artificial Intelligence. This implementation is about determining the gender and age of a person from a given image. All the Facial Features play a crucial role in social interactions of a person by determining the person's race, age, gender and Identity. But age and gender are considered to be the key attributes among all, because of the increased usefulness of age and gender in smart applications.

The scope of this application in future is really bright as this can be used in several different areas for efficient results and helps mankind by significantly reducing the time and labour consumed to accomplish a tedious task. Few areas where age and gender prediction can be applied are human- computer interaction, law enforcement, visual surveillance, electronic customer relationship management, entertainment, cosmetology, electronic vending machines, forensic art and access control

A crucial part of gender and age prediction is **Facial Detection**. Face detection's sole responsibility is to locate the presence and highlight the face(s) in an image but it strictly does not identify the person's identity. In machine learning approach, Face Detection is performed using Haar cascades of OpenCV, where a cascade function is trained using a set of input data. The OpenCV contains many more pre-trained classifiers for face, eyes, smiles, etc.

With regards to the gender and age prediction trained model efficiency, it hugely depends upon the network and data used. Thus, we will be working with Convolution Neural Networks (CNN) because of its significant efficiency in classification tasks. Here we are basically considering an 8- layer CNN network consisting of an input layer, 6 Hidden layers (3 convolutional layers & 3 fully- connected layers) and an output layer. The reason for considering this 8-layer network is to avoid the risk of overfitting the model and also because it fetches the nature of our task. CNN network is basically used for extracting the features from the input image and follows a feed forward network to classify the gender and age.

One of the challenges with CNN network is the need for homogenous data to train the network besides large dataset appears with its own challenges such as cleaning the dataset and loading it to the model. The dataset must be properly

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.114126

labelled and low-quality images are to be removed from the dataset as it can affect the accuracy of the resultant trained model. OIU-Adience Dataset is a public dataset which consists 26,580 unfiltered images for the purpose of gender and age prediction. The unfiltered images are labelled with gender either 'MALE' or 'FEMALE' and age in classified groups [(0 – 2), (4– 6), (8 – 12), (15– 20), (25 – 32), (35 – 43), (45 – 53), (60 –100)].

LITERATURE REVIEW:

To choose an approach or to build an efficient model, it must be known/ aware of earlier works related to gender and age prediction. The past works may include information about different algorithms their efficiencies, and approaches used. The literature review can help to understand better of and consider it as a guiding principle and support the research to be taken up.

So far, all the earlier methods of Gender and Age Prediction were handcrafted where the facial features were manually engineered from images taken under controlled conditions. In 1999, Kwon and Lobo were the first to develop Gender and Age Prediction model using geometric features of face. This method was successful to distinguish between the face of babies and adults, but unfortunately it couldn't distinguish between the face of an adult and a senior adult. Lanitis et al., 2004 proposed a model that included both geometric features of face and texture features using Active Appearance Model for the prediction task. However, these methods were a failure as they were not suitable for uncontrolled or real-life images.

To overcome the above drawback many researches have started using CNN networks for better results on real-life images. CNN can effectively classify unfiltered images using its good feature extraction methods. The availability of large datasets made it possible train CNN models which can learn compact and discriminative facial features making it even more efficient for classification task. The CNN network should consist sufficient number of layers (not more or less) according to the task so as to avoid the risk of overfitting and underfitting the model. The CNN solution consists of a strong layer for face alignments in images for which we need to properly pre- process the images and then feed it to the model.

An important challenge of Gender and Age Prediction is to detect faces in images, but there might be several visual differences in each image such as lighting, posing and angle of the camera. A way to address this challenge is by training a model that can differentiate between the face and the background in an image, Haoxing li et al. has proposed that using CNN with multiple Prediction layers which can adjust the image's alignment and detect potential face(s) in images.
Also, Z. Tan, J. Wan, Z. Lei, R. Zhi, G. Guo, and S. Z. Li, have studied an age encoding CNN Model, where they explored the relationship between the real age of a person and adjacent ages to identify Age Groups for classification.

PROPOSED METHODOLOGY:

Our Proposed System determines the gender and age of person(s) in images using Facial Recognition and CNN Network (trained using the OIU-Adience Dataset). This model is a classification model and has improved level of identification between similar faces in images. And, as we have discussed earlier this implementation will be beneficial for human race in coming times.

Let's us now look closer into the implementation and functions of our model. The **Fig-1** represents the flow of implementation of our model. There are basically few prime functions involved here – they are, Image Processing, Face Detection, Gender and Age Prediction.
Image processing stage is responsible for aligning the image and detecting the face(s) in the input image. In a dataset not all images are aligned the same as there might be variations in the lighting, background and image angle. So, the images are to be aligned properly in compatibility to our window alignment and then this updated image is converted from RBG to Grey Scale, the Grey scale conversion will discard useless data in RGB for easy processing. Also, RGB to Grey scale conversion signifies conversion of 3D (R, G, B) image to 2D (Black, White) which helps in faster processing of images. This Converted Grayscale image is then passed on for Face detection.

The Face detection's sole responsible is to locate the presence and highlighting the faces displayed in an image but it strictly doesn't identify the identity of the person. In machine learning approach, we implement OpenCV library. OpenCV consists predefined modules for identification of text, objects, and faces in an image. Face detection is performed using Haar cascades of OpenCV, where a cascade function is trained using a set of input data. Next the highlighted faces in image are fed into Adience Dataset Trained CNN Network for Feature extraction, Gender and Age

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.114126

classification.

There are few set of pixels called as interest points which are unique in nature and can be robustly identified in an image. These interest points play an important role in feature extraction as they are mathematically obtained to form an affirmative shape or structure around the edges or corners of the features for identification. A human face consists of 80 nodal points representing distance between two eyes, shape of your cheekbones, nose etc., which are analysed for mapping of interest points accordingly. Once Feature extraction is finished, the next step is to accomplish gender and age Classification using the extracted information.

In this stage, the process is to estimate the gender of the person in the image accurately and this process is called as **Gender Prediction**. This is considered as gender prediction approach as a classification problem as discrete output values can be ascertained for gender i.e., "**MALE**" and "**FEMALE**". We train our prediction model using these two classes for gender prediction and the output is either "**MALE**" or "**FEMALE**".

Age Prediction is a very challenging task to accomplish. At this level, instead of predicting an accurate integer it was perceived and inclined to cluster and classify age into age groups. In Clustering, similar adjacent ages are grouped together to form small clusters and a range is set for each cluster which is later used as a class label to train the prediction model for age classification Therefore, the **Adience dataset** has 8 classes split into the following clustered age groups [(0 – 2), (4
– 6), (8 – 12), (15 – 20), (25 – 32), (35 – 43), (45 – 53), (60 – 100)]. Also, the prediction network has 8 nodes in the final SoftMax layer consisting the mentioned age ranges.

It is also to be noted that age prediction from a single image is not an easy problem to solve as the **perceived age** depends on a lot of influential factors that affect a photo for example: lighting, angle, posture and also people of the same age may look different in discrete parts of the world.

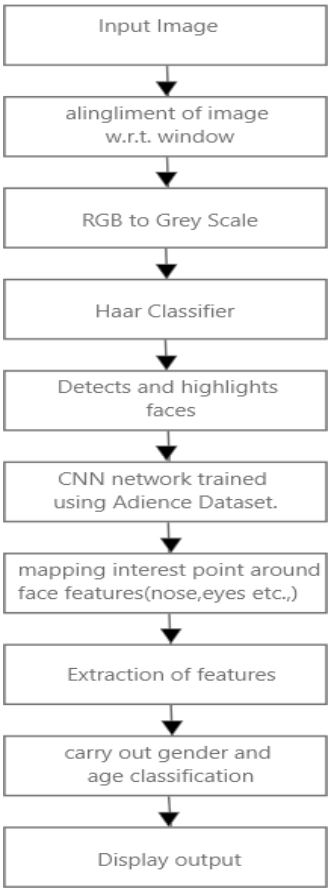

Fig 1– Flow of Implementation

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.114126

DATASET:

OIU-Adience Dataset

Adience dataset is an open-source dataset provided by the Open University of Israel (OIU). The dataset has 26,580 unfiltered images of 2284 subjects gathered together from Flickr for the sole purpose of Gender and age prediction applications. The images in the dataset are taken under uncontrolled situations and have high-levels of variances in Noise, pose and lighting among all other datasets. And, the images are labelled with gender one of MALE and FEMALE and Age is split into 8
subsequent groups or classes - [(0 – 2), (4– 6), (8 – 12), (15– 20), (25 – 32), (35 – 43), (45 – 53), (60 – 100)].

OpenCV - Haar cascade:

OpenCV is an abbreviation for open-source computer vision library. OpenCV was first developed by intel and is a cross platform framework written using C++. It consists basic infrastructure and pretrained modules required for real-time computer vision applications. The pretrained models like Haar Cascade are located in the data folder in the OpenCV installation.

Haar Cascade is a feature Based classifier used for object detection in given input visual source. This cascade function is trained using set of data consisting of images with faces and without faces. The basic approach of Haar Cascade is to divide the features required for face detection into different layers and applying them one by one on the image making the task easier and efficient. The model discards an image if it fails in a layer and considers it as a failed case and if an image passes all the layers then it is understood that the image has face(s) in it for facial detection.

CNN Architecture:

CNN Architecture is end-to-end connected and is sequential in nature. As per our paper, CNN network is responsible for feature extraction, gender and age classification task. Also, we are following Feed-Forward model which is the simplest form of CNN network, where data flows only in forward direction and there is no going back to previous step by using loop or cycle. Our CNN network consists of 8 input layer, 6 hidden layers, an output layer. The hidden layers consist of 3 convolutional layers and 3 fully-connected layers. Further let's look into the functionality of these hidden layers and how the information flows through the CNN network.

After aligning and detecting the face(s) in the image, this data is then passed on to the CNN input layer. The input layer then passes this to the hidden layers.

- The first hidden layer is a convolutional layer, this layer consists of 96 filters of size 3 x 7 x 7 applied on the Input, followed by rectified linear operator layer (ReLU) (activation layer), Max pooling layer, and normalizing layer.

- The output of first hidden convolutional layer acts as the input to second hidden convolutional layer which has 256 filters of size 96 x 5 x 5 applied on the Input, followed by rectified linear operator layer (ReLU), Max pooling layer, and normalizing layer.

Similarly, the output of second hidden convolutional layer acts as the input to third hidden convolutional layer which has 384 filters of size 256 x 5 x 5 applied on the Input, followed by rectified linear operator layer (ReLU), Max pooling layer, and normalizing layer.

- Next, there are 3 fully connected layers. The output of third hidden convolutional layer is the input for first fully connected layer. This layer has 512 neurons, a ReLU and a dropout layer (fixed 25% dropout). This layer produces 512-dimensional output.
- The 512-dimensional output is the input for second fully connected layer. This layer also has 512 neurons, a ReLU and a dropout layer. This layer too produces 512-dimensional output.
- The output of second fully connected layer is then passed as input to the third fully connected layer, where this layer performs classification by mapping the input to plausible gender and age classes.

Then the output of third fully connected layer is sent to SoftMax layer placed just before the output layer, where it

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.114126

assigns the mapped plausible classes to defined gender and age classes.

RESULTS:

Below are few screenshots of the results obtained by implementing our proposed model. Webcam was used as the source of input to the model. In our test cases we have used pictures consisting faces of different ages and genders to see whether our system was efficient in predicting age and gender accurately. Additionally, from **Fig-4: Test case 3,** it is significant that our system was successful in detecting multiple faces in an image

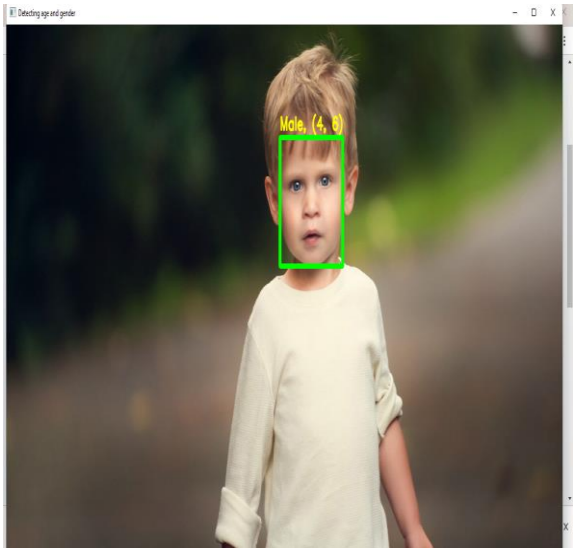

Fig-2: Test Case 1

Fig-3: Test Case 2

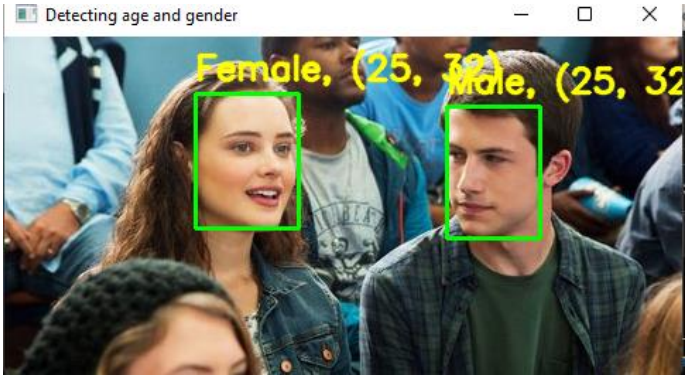

Fig-4: Test Case 3

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.114126

CONCLUSION:

This methodology has been a challenging one while implementing it in real-life because of the usage of large data set and ever existing scope of overfitting. But it can be concluded that the results of the above considered model when implemented in real-life are well in terms of performance and accuracy when compared with performance of other approaches. The Adience Dataset has also allowed us to fine tune our training model with its large varied collection of images.

FUTURE ENHANCEMENTS:

Coming to future enhancements, presently most of the image detections algorithms are efficient detecting objects in frontal images, further it can be improvised to identify objects in side angles images and inverted images. While, we can also extend our present methodology by incorporating emotion prediction in faces. If this can be achieved, we can open doors to many solution-less avenues.

REFERENCES:

1. "Age and Gender Prediction", Harcharan Kaur Universal group of institutes, lalru Punjab – India.
2. A. Kumar and F. Shaik," Importance of Image Processing", 2016.
3. H. Han and A. K. Jain, "Age, gender and race estimation from unconstrained face images," Tech. Rep., Michigan State University, East Lansing, MI, USA, 2014, MSU Technical Report, MSU-CSE-14-5.
4. G. Levi and T. Hassncer, "Age and gender classification using convolutional neural networks," in Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 34–42, Boston, MA, USA, June 2015.
5. "Review Paper on Real Time Age Rank Estimation with Gender Prediction with Image Processing ", Sharayu S. Ganorkar, Prof. S. B. Rathod, Student (M.E.), Professor, Computer Science and Engineering, Sipna College of Engineering and Technology, Amravati, India.
6. Sarah N. Kohail," Using Artificial Neural Network for Human Age Estimation Based on Facial Images", 2012 International Conference on Innovations in Information Technology.
7. R. Begg1 and J. Kamruzzaman," Neural networks for Prediction and classification of walking pattern changes due to ageing", 2006 ACPSEM/EA.
8. Eran Eidinger, Roee Enbar, and Tal Hassner," Age and Gender Estimation of Unfiltered Faces", 2014 IEEE.
9. Aditya K. Saxena, Shweta Sharma and Vijay K. Chaurasiya," Neural Network based Human Age-group Estimation in Curvelet Domain", 2015 The Authors Published by Elsevier B.V.
10. Hang Qi and Liqing Zhang," Age Classification System with ICA Based Local Facial Features", 2009 Springer-Verlag Berlin Heidelberg.
11. "Deeply Learned Classifiers for Age and Gender Predictions of Unfiltered Faces", Olatunbosun Agbo-Ajala and Serestina Viriri.