

Real Time Gender Classification using Convolutional Neural Network

Pooja Mishra^{1,*}, Neha Singh^{2,**}, and Pallavi Chavan^{3,***}

¹Information Technology, Ramrao Adik Institute of Technology, Nerul, Navi Mumbai, MH, India.

²Information Technology, Ramrao Adik Institute of Technology, Nerul, Navi Mumbai, MH, India.

³Information Technology, Ramrao Adik Institute of Technology, Nerul, Navi Mumbai, MH, India.

Abstract. This paper presents real time gender classification using Convolutional Neural Network. Automatic classification of gender has become important to an growing array of applications particularly with the emergence of web networks and social media. Slavery was a significant moral problem in the nineteenth century. It was a struggle toward fascism in the modern period. The fight for gender equality across the world, as well as the need to divide gender for meaningful purposes, would, we conclude, be the most critical moral issue of this century. Differences are needed at different places, such as restrooms for men and restrooms for women; attire for men and attire for women; and so on, in order to plan and advance further in the technological sector. To decrease crime rates, to place the advertisements in malls precisely attracting more people based on gender, to keep track of genders in respective toilets or in trains, for personal services, etc. The authors propose the gender classification dilemma for real-time applications, in which a tool decides if the faces within the exposure belong to a female or a male. The primary fundamental region of experimentation in this venture is adjusting a few already distributed, successful designs utilized for gender orientation classification. Generally, facial structure variations have an effect on gender classification accuracy considerably, as a result of facial form and skin texture modification as they become old. This requires re-examination on the sexual orientation classification framework. By learning representations through the utilization of deep-convolutional neural networks (CNN), a major increase in performance is obtained on these tasks.

1 Introduction

Gender plays a central role in social experiences. The number of picture uploads to the Internet has risen at an almost unprecedented pace over the last decade. This new discovered abundance of data has allowed computer scientists to solve previously either irrelevant or unsolvable machine vision issues. Recognition of gender is an inevitably challenging problem, far more than most other computer-vision activities. The main explanation for this complexity difference lies in the design of the data used to train these types of systems. According to a record, out of 7.8 billion people (current stats of year 2020) in total population of world, there are 4.48 billion internet users (statistics of October 2019) which has increased in this COVID pandemic situation. So there is a growth in the usage of automatic gender recognition because of online website shopping, online transactions, online social networking websites and social media. However, recognition of gender is an inevitably challenging problem, far more than most other computer-vision activities. The main explanation for this complexity difference lies in the design of the data used to train these types of systems. And for many such digital progress we should be ready with a better gender classification system which can process every

person with a high accuracy even of a low quality light surroundings. Creating efficient, precise frameworks for tasks of gender distinction, and seeking to expand their solutions to boost performance and have a very good accuracy. Nevertheless, there is still a significant lack of performance of current methods on real-world photographs, particularly as compared to the enormous performance strides recently recorded for the related face recognition mission. In this modern world where everything is getting automated and online, starting from shopping for phones to jewellery. It is very important for websites to be efficient in giving right choice for customers as per their gender, taste and interest.

2 Gender Detection Data Analysis

Thanks to the increasing usage of face recognition technologies in the field of safety, a slew of profile databases have sprung up in recent years. The authors have developed the data set of faces to get the right idea of how our data enhances the accuracy for Indian faces' gender recognition in real-time. Gender classification schemes are in a peculiar situation since they only describe 2 groups (male and female). As a result, classification and authentication methods are used in this situation without incurring any penalties.

*e-mail: mishra.pooja.17it1082@gmail.com

**e-mail: singh.neha.17it2012@gmail.com

***e-mail: pallavi.chavan@rait.ac.in

2.1 Data collection

The authors collected data from the public domain, Social media like Instagram and Facebook (with the user’s permission). The authors uses a web scraping tool to scrap the photos from their profile pictures and crop them around the frame. These pictures are improved in terms of quality by reducing the noise. The authors collected photos in one folder which was divided into 2 folders namely: Male and Female. These photos were then subdivided into 2 folders in each folder: Modified and Raw. Raw folder consist of photos which are collected from the Web scraping tool and through friends and families using WhatsApp and google drive link and it contains folders of Instagram, Group, Whats app. Modified folder consisted of photos which are improved versions of the raw photos by cropping and data improving tools which are manually segregated in male and female photos. As the difficulties increased in improving accuracies authors added more photos to the Raw and Modified photos folder.

2.2 Facial Image Databases

For improved accuracy, authors have used different types of faces and the number of times we have trained the model with each type of face to generate an accurate report of the database which was most useful to predict better results. While the images in real-time tend to recalculate the gender continuously so it requires every point to distinguish between the genders. Authors have created a mixture of each database type and trained the model with one database one at a time.

2.2.1 ORL database

The first widely used database created is of type YALE database which was first originated in the Olivetti Research Laboratory database. These datasets contain pictures shot from a frontal or nearly frontal face posture shown in Fig. 1.



Figure 1. Examples from collected ORL database

2.2.2 YALE database

The YALE database, on the other side, displays people in all sorts of combinations, including front/sides of face, high brightness/low brightness, with/without spectacles, and joyful/surprised/tired/sluggish/sleepy/laughing/crying/unhappy or wink emotions. As shown in Fig. 2 .



Figure 2. Examples from collected YALE database

2.2.3 YALEb database

Large security measures placed at airport terminals or public spaces are used in more strong and complicated scenarios. These systems must contend with uncontrollable lightning, – anti individuals, and huge stores of identity. In order to assess systems against these kind of scenarios, the dataset contains almost all the features.As shown in Fig. 3 .



Figure 3. Examples from collected YALEb database

2.2.4 FERET database

The FERET dataset includes pictures taken in a moderately controlled setting from 200 people in various face postures. As shown in Fig. 4. This dataset has the fascinating virtue of providing substantial statistics describing positions of face features, ethnicity, sexual orientation, and face traits such as moustache, beards, and spectacles details.



Figure 4. Examples from collected FERET database

2.2.5 FRGC database

Another comprehensive resource is the FRGC dataset. The FRGC dataset, like FERET, is made up of high-resolution pictures from a group of over 200 individuals, as well as comprehensive statistical information. This dataset, on the other hand, contains pictures of the entire body captured in various settings, implying significant

variations in backdrop and lighting conditions. As shown in Fig. 5. In summary, while evaluating a program, it is critical to consider the type of dataset being utilised. Using different databases creates diverse circumstances, allowing us to verify it against variety of challenges.



Figure 5. Examples from collected FRGC database

3 Proposed Architecture

In this work authors are using Convolutional Neural Network to train the model for Gender Detection. It is a Deep Learning Algorithm which specialises in taking input as image and learning various aspects from it. It processes images as per the data given to train and the output expected out of it. It is able to perform this classification because the data given to train the model are assigned among two class. Female and Male are the only two classes it has to classify between. As per the label assigned and using back propagation it assigns various learnable weights and biases to the objects present in the image. The pre-processing required are less in ConvNet when compared with other classification algorithms. Convolutional Network has ability to teach itself various aspects, characteristics and filters. The architecture of ConvNet is similar to how neurons are connected in Human Brains. Individual neurons respond to specific region in front of the visual field known as receptive field. ConvNet's primary aim is to reduce pictures in a way that makes it easier to analyse while preserving essential classification information.

3.1 Methodology

Following is the methodology:

1. $C = \{Instagram, Whatsapp, GroupPhotos, Family\}$
2. $Data = D - A(Data) - Noise$
 where $Data \rightarrow Enhanced\ Data,$
 $D - A \rightarrow Data\ Augmentation,$
 $Noise \rightarrow Noise\ removal,$
3. $CrImg = mtcnn.faceDetect(Data)$
 Mtcnn uses face detection model to crop face, which is stored in Male and Female Folders
4. $TrainSet, TestSet = sklearn.split(Folders),$
 where $Folder \rightarrow Female\ Folder, Male\ Folder,$
5. $Model = CNNModel.add(TrainSet),$
6. $RealTime\ Video = Opencv.videoCapture(Camera),$

7. Gender = Model.Predict(RealTime Video)

Gender is predicted and displayed on Realtime Video

As mentioned above the flow of work started with data collection by sources mentioned in first step. Collected Raw data is then preprocessed to enhance for better classification and faces are cropped from the fed data which is followed by segregating the data into male and female folders respectively. Two different image folders are then divided into train and test set which trains the model and tests model accuracy on test set. The generated model is used to classify the gender on realtime video.

4 System Architecture

The Figure 6. shows the flow of System that is divided into different stages. It starts with data collection and ending Gender Detection by the features that got extracted using the CNN Model. The middle stages contain stages on data preprocessing, manual segregation to provide for classification, training the model and testing its accuracy.

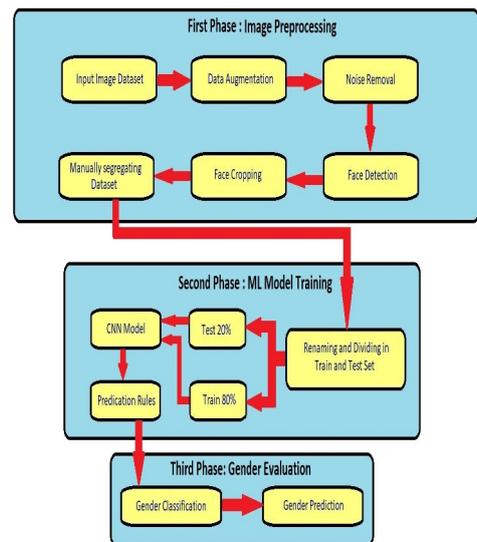


Figure 6. Real Time Gender Classification Architecture

4.1 CNN Algorithm

In this work, Keras Tensorflow library is used. Different layers are added to the sequential Model starting from conv2D. This layer marks the start of taking 2d images as input. Various layers that are included in the model are explained below

The Layers used are as follow:

1. **Conv2D:** The first layer of the Convolutional Neural Network remains the layer of the Convolutional Layer. Convolution layers apply the convolution effect to the input, transferring the result to the following layer. The answer converts all pixels in

their reception field to a single value. The foremost common kind of convolution used could be a 2D convolution layer and is commonly abbreviated as conv2D. The filter or kernel within the conv2D layer "slides" over 2D input file, enabling element-wise duplication. As a result, it summarizes the results into one single pixel output. The kernel will perform the identical function everywhere it goes, converting a 2D feature matrix into a separate 2D feature matrix.

2. **BatchNormalization:** To avoid overfitting, familiarity with a pre-data processing tool used to bring numerical data to a standard scale without distorting its structure. Generally, when authors enter data into a machine or in-depth reading algorithm authors tend to change values on an average scale. The reason authors are accustomed in part ensures that the model can adapt properly. Now back to the Batch normalization, it is a process of making neural networks faster and more stable by adding additional layers to a deeper neural network. The new layer performs the measurement and normalization functions of the layer input from the previous layer.
3. **Max Pooling:** Pooling layers are worn to reduce the scale of the feature maps. Therefore, reduce the quantity of learning parameters and also the number of computers created within the network. The integration layer summarizes the features present within the feature map region formed by the convolution layer. Therefore, some activities are performed on abbreviated elements instead of precisely set elements produced by the convolution layer. This makes the model more powerful in changing the form of the elements within the input image. Integration function includes sliding a two-sided filter at each feature map channel and summarizing features within a filter-covered region. Max Pooling maybe a merging function that selects the highest element from the feature map region covered by the filter. Therefore, the output after the max-pooling layer are a feature map that contains the foremost prominent features of the previous feature map.
4. **Flatten:**Classifying the 'Gender.' We created a Classification model, which means that the data used should be a good fit for the model. It needs to be in the form of a 1-sided vector. Rectangular or cubic shapes cannot be a straight input. And that's why we need flatten layers and fully connected. Flattening converts data into a 1-dimensional array to insert it into the next layer. We consider the removal of convolutional layers to create a single vector of long feature. It is also connected to a final separation model, called a fully connected layer. In other words, we put all the pixel data in a single line and make connections to the last layer. And again.

'Gender' segregation.

5. **Dense:**A dense layer is a layer of deeply connected neural network, which means that each neuron in a dense layer receives input from all the neurons of its previous layer. The dense layer is found as the most widely used layer in models. In the background, a dense layer creates a matrix-vector multiplication. The values used in the matrix are actually in parameters that can be trained and updated with the help of back propagation. The output produced by the dense layer is in the 'm' dimensional vector. Therefore, a dense layer is used to change the vector size. The dense layers work and function as rotation, scale, vector translation.
6. **Sigmoid:**Sigmoid activation function, also called activation function, is traditionally the most active opening function of neural networks. Input input is converted to a value of between 0.0 and 1.0. Inputs greater than 1.0 are converted to a value of 1.0, similarly, values greater than 0.0 are reduced to 0.0.

5 Experimental Results and Performance Analysis

5.1 Accuracy and Loss

Figure 7 shows the accuracy and loss comparison between train and test set when only celeb-A dataset is used.

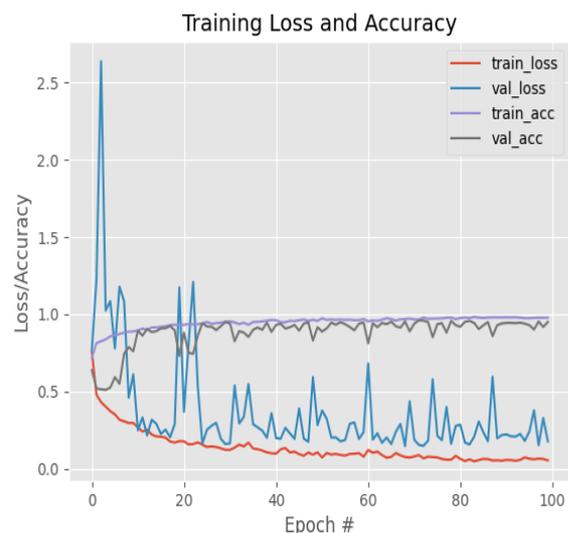


Figure 7. Loss Accuracy Graph on Celeb-A

Figure 8 shows the accuracy and loss comparison between train and test set when combined data of celeb-A and our collected data was fed to the Model

It is easily clear seeing both the graphs that the accuracy of Model when trained with only Celeb-A dataset which is readily available on Internet has a poor performance on

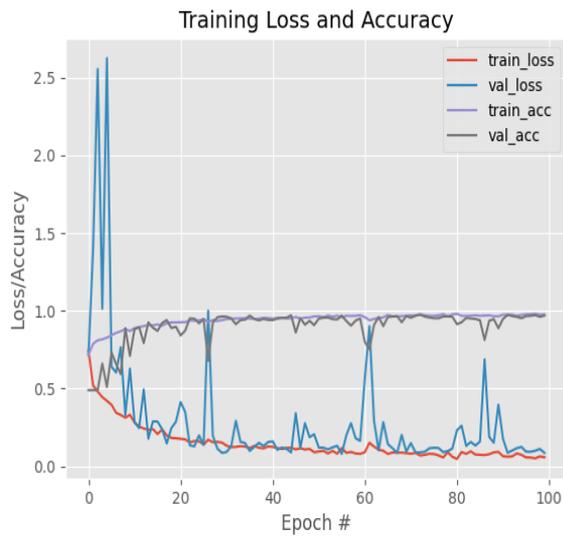


Figure 8. Loss Accuracy Overall

Unseen data whereas when it is trained with combined data yields a better performance.

5.2 Result Images

Test on realtime images are shown in the figures below. Different angles and look have been tried to check whether it was able to distinguish clearly or not.

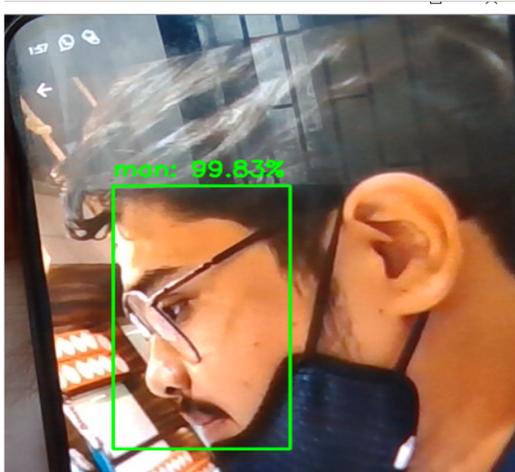


Figure 9. Male Face with side face and one-third face covered with mask

5.3 Performance Analysis

1. Earlier the model were trained with less raw images and that caused decrease in accuracy when classifying the human in realtime with different surrounding conditions. This problem was overcome by creating our own dataset which resembled realtime images to train the Model.

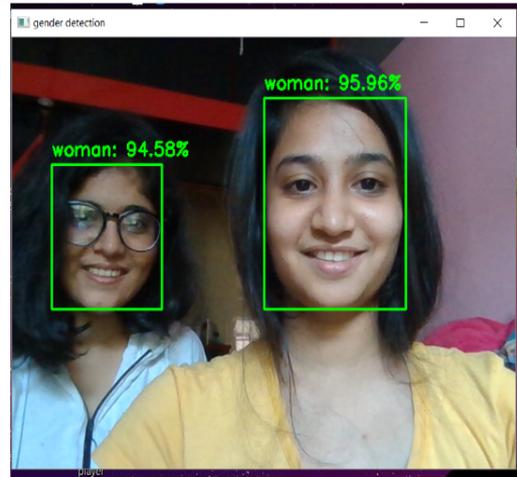


Figure 10. Two Females

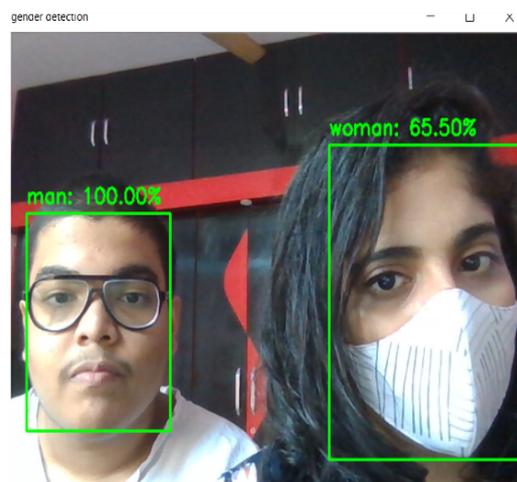


Figure 11. Unmasked and Masked Male-Female

2. Second major problem was that many of the literature surveys tried gender classification but did not perform on realtime, while had slow performance in classifying the real-time feed instantly. This was solved by giving the model the required data and adding the right layers on the right places such as dropout and batch normalization in order to prevent overfitting on the training data.

6 Conclusion and Future Scope

The live gender classification and prediction dilemma, from which a machine automatically decides whether a source face relates to a female or a male, was developed in this chapter. While several previous approaches have tackled the problem of gender classification of images, this paper lays out a benchmark for the challenge based on state-of-the-art network architectures and shows that chaining age prediction with gender prediction will increase overall precision. Through this work, authors can draw two significant conclusions. First, CNN should be used to produce greater outcomes in gender labeling, even with the much

smaller scale of contemporary unconstrained picture sets that are classified for gender. A gender classified properly can help cut down half of the choices for clothing, jewellery etc, which helps to give better experience to users surfing on the website. Not only for shopping but gender classification can also help for marketing. With the upcoming new social changes in the future there is one more gender which needs to be added Gender classification models that is for transgenders. To expand their methods so that outcomes can be strengthened. An end-to-end multi-task learning scheme for categorizing gender.

References

- [1] Amit D. Ranjit K. Vijay B. Gender Recognition Through Face Using Deep Learning. *Procedia Computer Science* 132 (2018) 2–10
- [2] Sajid Ali Khan, Muhammad Nazir, Naveed Riaz(2013) "Gender Classification using Multi-Level Wavelets on Real World Face Images", *Springer Acta Polytechnica Hungarica* Vol. 10, No. 4.
- [3] S. Baluja and H. A. Rowley. Boosting sex identification performance. *Int. J. Comput. Vision*, 71(1):111–119, 2007.
- [4] Jridi, M. Napoléon, T. Alfalou, A. One lens optical correlation: Application to face recognition. *Appl. Opt.* 2018.
- [5] Tivive FHC, A. Bouzerdoum (Sep 2016) "A gender recognition system using shunting inhibitory convolutional neural networks" In: *International Joint Conference on Neural Networks*; Vancouver, Canada. New York, NY, USA: IEEE.
- [6] Vaishnavi Y. Mali. Human Gender Classification using Machine Learning. *y (IJERT)* Vol. 8 Issue 12, December-2019
- [7] Ma, Z. Ding, Y. Li, B. Yuan, X. Deep CNNs with Robust LBP Guiding Pooling for Face Recognition. *Sensors* 2018.
- [8] Van de Wolfshaar, J. Karaaba, M.F. Wiering, M.A. Deep convolutional neural networks and support vector machines for gender recognition. In *Proceedings of the 2015 IEEE Symposium Series on Computational Intelligence*, Cape Town, South Africa, 7–10 December 2015.
- [9] Duan, M. Li, K. Yang, C. Li, K. A hybrid deep learning CNN-ELM for age and gender classification *Neurocomputing* 2018.
- [10] Levi G., Hassner T. Age and gender classification using convolutional neural networks. Boston, MA, USA, 7–12 June 2015.
- [11] Rajeev, R. Vishal, M.P. Rama, C. HyperFace. A deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition. *IEEE Trans. Pattern. Anal. Mach. Intell.* 2019
- [12] Yassin Kortli, Maher Jridi, Ayman Al Falou and Mohamed Atri. *Face Recognition Systems: A Survey*. *Sensors* January 2020
- [13] Khalil Khan, Muhammad Attique, Ikram Syed and Asma Gul. Automatic Gender Classification through Face Segmentation. *Symmetry* 2019
- [14] Alessandra Lumini, Loris Nanni, and Sheryl Brahnam. 2016. Ensemble of texture descriptors and classifiers for face recognition. *Appl. Comput. Info.* (2016).
- [15] Shubham Patil Bhagyashree Patil, Ganesh Tatkar. Gender Recognition and Age Approximation using Deep Learning Techniques. *y (IJERT)* Vol. 9 Issue 04, April-2020