IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-9, Sept 2019 E-ISSN: 2347-2693

A Real Time Gender Recognition System Using Facial Images and CNN

Taran Rishit Undru[1*], CVNS Anuradha[2]

[1,2]Dept. of Computer Science and Engineering, Gokaraju Rangaraju Institute of Engineering and Technology, Bachupally, Hyderabad, India

[*]*Corresponding Author: taranrishit1234@gmail.com, Tel.: 9182314137*

DOI: https://doi.org/10.26438/ijcse/v7i9.122126 | Available online at: www.ijcseonline.org

Accepted: 08/Sept/2019, Published: 30/Sept/2019

Abstract— With technological advancements many small to large, simple to complex activities are automated. Growth of Artificial Intelligent techniques has eased the way we would look to solve the real world problems. One such area which has recently gained lot of attention is the facial analytics. It involves extracting features such as face expressions, gender, age etc. Gender information plays a vital role in areas such as human computer interaction, crime detection, gender preferences, facial biometrics for digital payments etc.

This paper proposes **an improved** Convolutional Neural Network (CNN) framework for real time gender classification from facial images. A pretrained model Visual Geometry Group "VGGNet16" is used. It loads image datasets consisting of male and female images and trains consistently for 16 hours. Haar Cascade classifier is used to classify images based on facial traits. The proposed architecture exhibits a much reduced design complexity as compared to other CNN solutions applied in pattern recognition. A recognition accuracy of 90% was achieved with this method.

Keywords— *CNN, Face Images, Gender Recognition*

I. INTRODUCTION

As the saying goes "Face is the index of mind", human face can depict many characteristics such as ethnicity, gender, age, emotions etc. Facial analysis has recently gained lot of attention from research fraternity. Face is an important part of human body. Gender recognition from face images has become one of the mostly researched areas of computer vision. It can be applied to many areas such as security, human computer interaction, advertising and marketing. Smart phones with cameras are a common gadget nowadays and people often share their pictures on open platforms. Every day thousands of Terra bytes of such images are passed on Internet. We can use this big data to extract many important information.

In the last decade, studies on two-dimensional face recognition methods gained significant value in the field of computer vision. However, they are limited to changes in illumination conditions, occlusions, and facial expressions.

With the emergence of intelligent techniques, the challenges faced by facial recognition systems are met by applying machine learning algorithms. Machine learning is an area of Artificial Intelligent (AI) that provides machines an ability to learn and improve from experiences without having to be explicitly programmed. A CNN is a deep learning algorithm which takes an image as input, assigns weights to different objects in the image and differentiates one object from another. The pre-processing required in CNN is lower than other classification algorithms. Let us see briefly how CNN works. There are three basic components of a CNN. (i) the convolutional layer (ii) the pooling layer-optional and (iii) the output layer. "The convolution layer takes an image, define a weight matrix and the input is convolved to extract specific features from the image without losing the information about its spatial arrangement" [1]. Figure 1 depicts the image after convolution.

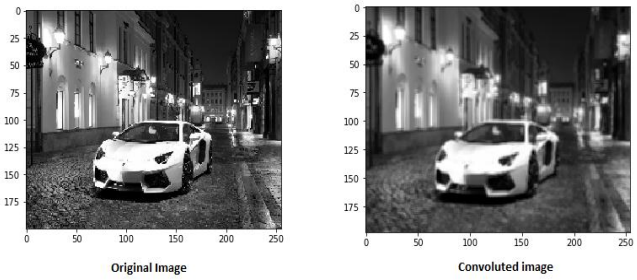

Figure1: Image after convolution [1]

Sometimes when the image is too large we need to reduce its size without losing information. A pooling layer could be used between convolution layers. The most common type of pooling layer used is the max pooling. Figure 2 depicts pooling.

Figure2: Pooling [1]

The output layer in a CNN is a fully connected layer, which takes input from the other layers and flattened and sent so as to transform the output into the number of classes as desired by the network.

This paper proposes a CNN based real time gender recognition system to classify images into man and woman. The effectiveness of this method is demonstrated by applying it to datasets. Section 2 gives an insight of the state-of-the-art in the problem area. Section 3 proposes a model for the implementation of the problem. Section 4 highlights the key implications of the proposed model along with future scope.

II. RELATED WORK

This section describes the state of the art in the field of gender identification from facial images.

In [2] the authors propose a image-based gender recognition and age estimation model.A competition-winning deep neural networks with pretrained weights are used Transfer learning is explored using both VGG19 and VGGFace pretrained models by testing the effects of changes in various design schemes and training parameters in order to improve prediction accuracy. A hierarchy of deep CNNs is tested that first classifies subjects by gender, and then uses separate male and female age models to predict age. A gender recognition accuracy of 98.7% and an MAE of 4.1 years is achieved. The authors suggests that with proper training techniques, good results can be obtained by retasking existing convolutional filters towards a new purpose.

In [3] the authors **p**ropose a framework to classify the facial images based on their gender. The framework uses a modified form deep convolution neural network (CNN), to obtain higher performance and accuracy. This model can be used even for processing large quantity of data. They combined modified deep convolution neural network and KNN-classifier to create an application that can classify gender accurately. The rate of accuracy could be increased by increasing the number of layers and simultaneously training the images using back propagation.

Authors in [4] purpose a special convolutional module that allow a model to recognize emotions and gender with a considerable lower number of parameters, enabling real-time evaluation on a constrained platform. The reported accuracies were 96% in the IMDB gender dataset and 66% in the FER-2013 emotion dataset, while requiring a computation time of less than 0:008 seconds on a Core i7 CPU.

Authors in [5] present a real-time end-to-end gender detector based on deep neural networks. The proposed method detects and recognizes the gender of persons in images with a high inconsistency in pose, illumination and occlusions. To train and assess the results a new annotation set of Pascal VOC 2007 and CelebA were created and used. The experimental results indicate that combining both datasets during training can increase the mAp of the gender detector.

The paper [6] presents a robust gender classification model using deep convolutional neural networks. The proposed model segments the iris from a background image using the graph-cut segmentation technique. The model contains 16 subsequent layers; three are convolutional layers for feature extraction with different convolution window sizes, followed by three fully connected layers for classification. The proposed architecture achieved 98.88% accuracy although it increased the number of datasets by three folds.

The paper [7] focuses on the automatic extraction of persons and their attributes (gender, year of birth) from album of photos and videos. A two-stage approach is proposed in which, first, the convolutional neural network simultaneously predicts age/gender from all photos and additionally extracts facial representations suitable for face identification. The MobileNet is modified and is preliminarily trained to perform face recognition in order to additionally recognize age and gender. The age is estimated as the expected value of top predictions in the neural network. In the second stage of the proposed approach, extracted faces are grouped using hierarchical agglomerative clustering techniques. The birth year and gender of a person in each cluster are estimated using aggregation of predictions for individual photos. 94.1% gender recognition accuracy was obtained with the proposed method.

Paper [8] describes the details of a deep learning pipeline for unconstrained face identification and verification which achieves state-of-the-art performance on several benchmark datasets. A novel face detector is proposed , Deep Pyramid Single Shot Face Detector (DPSSD), which is fast and capable of detecting faces with large scale variations (especially tiny faces).

III. PROPOSED MODEL ARCHITECTURE AND IMPLEMENTATION

In this propsoed architecture , a dataset of images is fed as input to algorithm to identify male and female genders. A dataset of 6000 images is used here. Convolutional Nueral Networks is used as the classifying algorithm. In machine learning, CNN or ConvNet is aclass of deep feed forward, artificial neural networks , most commonly applied to analyzing visual imagery. CNNs requires less pr-processing as compared to other image classification algorithms.

Figure 3 depicts the train and classfiy modules of the proposed algotirthm. It indicates the steps involved in the entire process of gender classification. In the figure the train module has the CNN implementation, image is resized to (96,96,3) and converted into an array of (96*96*3,1) and image is augmented using ImageDataGenerator class which rotates the image in 3D to extract every feature and pass every image in dataset to this CNN. The output image os a 5 layer cosisting of Conv2D dimensions. The model is compiled using SGD and Binary crsoss-entropy and fit into the model. A dataset conatining iamges using –d option is passed. The dataset is splitted into woman and man directories. In the classify module in figure 3 test images are passed for classification, haar cascades detects faces.

Basically two steps are required to implement this architecture. Step 1 is to train the classifier with sufficinet number of images so that vareity of images are learnt by the classifier. Step2 is to test the claassifier for validation. Figure 4 depicts the steps in training the model.

It consists of a stack of five different layers. These layers are (i) convolutional layer (ii) activation layer (iii) batch normailization layer (iv) max pooling layer and (v)dropout layer.

We train the model on a image by image basis, the dataset consists of 6000 images both male and female. Pre Porcessing and resizing is done using CNN.

Step 2 is the testing phase of the model. In this phase images are fed as input and loop over the faces in the image using haar cascade. Haar cascade is a machine learning based approach where a cascade function is trained from a lot of positive and negative images. It is then used to detect objects in other images. In the proposed method , the model needs a lot of positive images i.e., the images conatining human faces and negative images i.e., the images without human faces to train the classifier. Then we need to extract features from it. For extracting the feautures haar features shown in figure 5 are used. Each feature is a single value obtained by subtracting sum of pixels under white rectangle from sum of pixels under black rectangle.

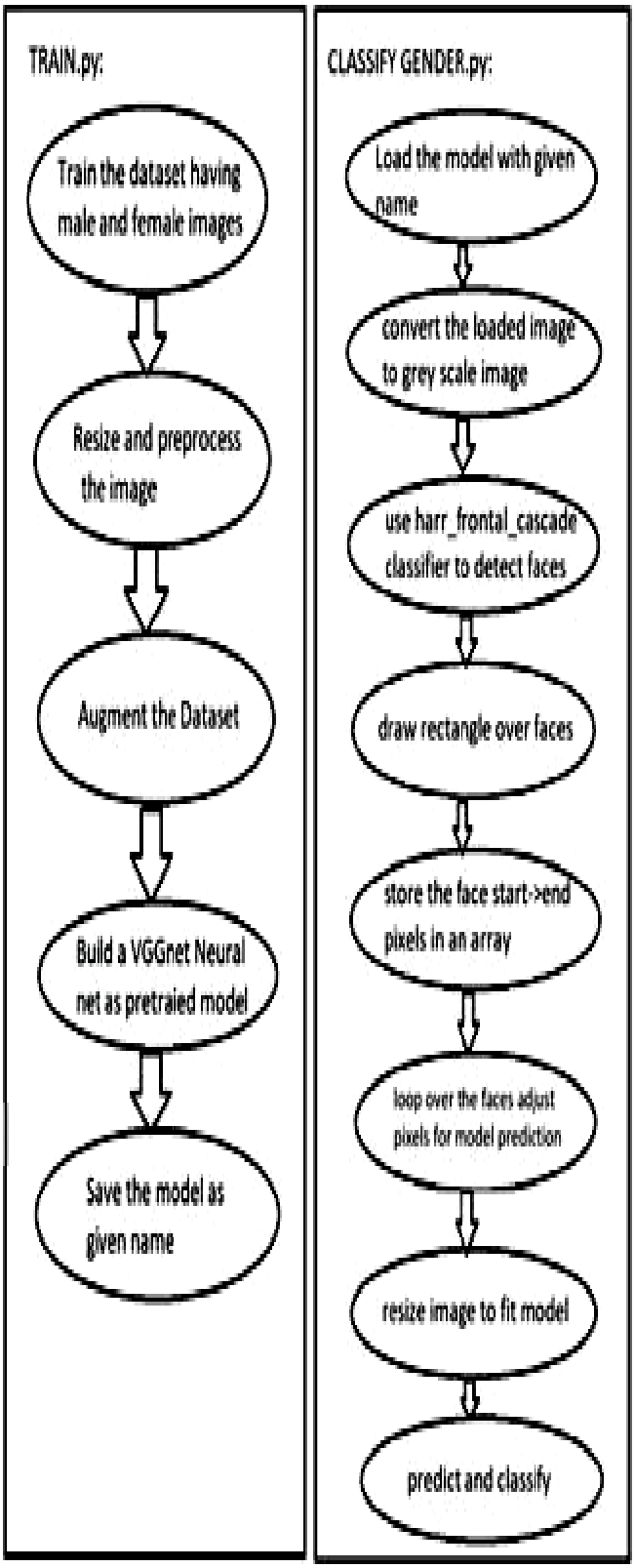

Figure 3: Process Model

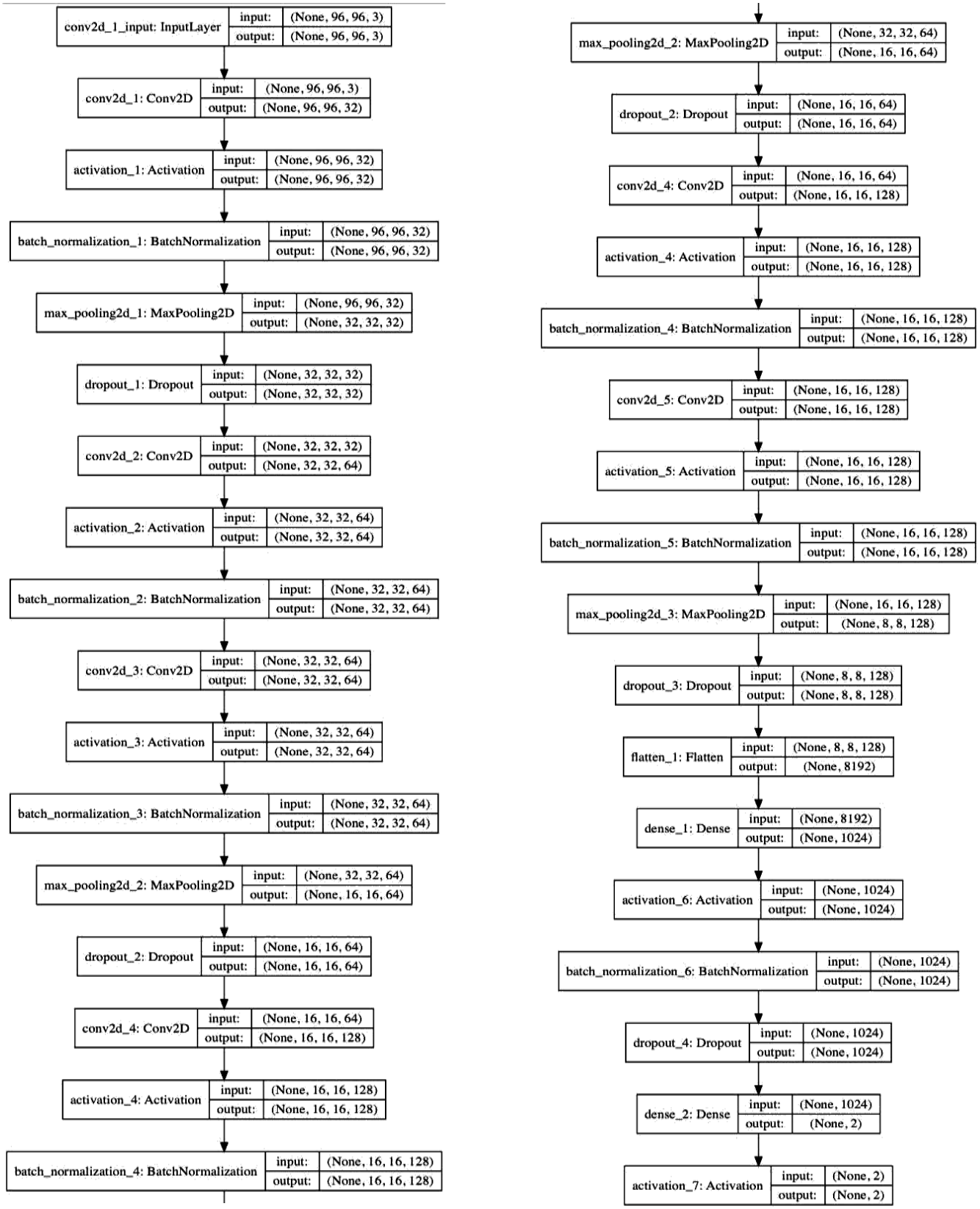

Figure 4: Training Model for Dataset

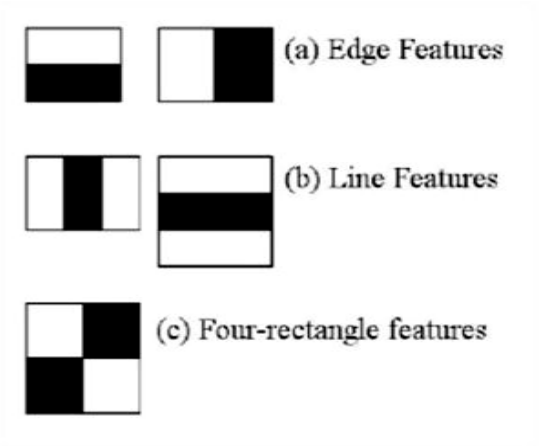

Figure 5: Haar Features

Figure 6 shows the output of the proposed gender classification model.

Figure 6: Output of classification

IV. CONCLUSION AND FUTURE SCOPE

This paper proposes a supervised CNN model to detect gender from the facial features from a given image. The proposed CNN architecture exhibits a much reduced design complexity when compared with other CNN solutions applied in pattern recognition. We first convert every image in the dataset into an array and pass them as input to CNN. Then the CNN strides over every image and condense into smaller images. We use max pooling to extract features. Pretrained model named "VGGNet16" which stands for visual geometry group is used .It basically loads images datasets consisting of male and female and trains them consistently increasing accuracy. In this case the trained time was 16hrs.We used classifier "haar cascades" to detect faces in a image. Other known classifiers are LBP but since LBP isn't as robust as haar cascades for classification, we used "haar_frontal_face_cascades. The model gives an average accuracy of 90%. Although it is a robust model to detect gender but sometimes it may not detect faces due to environmental conditions and dynamic hyper parameter tuning and might misclassify them too. A perfect image is required for this task to happen. This research model can be further improved to detect faces from images with varying intensities, illuminations and face expressions.

REFERENCES

[1] D. Gupta, "*Architecture of Convolutional Neural Networks (CNNs) demystified*", Analytics Vidhya, June 29, 2017.

[2] P. Smith, C. Chen, "*Transfer Learning with Deep CNNs for Gender Recognition and Age Estimation*", In the Proceedings of the 2018 IEEE International Conference on Big Data, Seatle USA,pp. 2564-2571,2018

[3] S. Choudhary, M.Agarwal, M. Jailia, *"Design Framework for Facial Gender Recognition Using MCNN*", International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Vol.8 Issue-3, pp**. 209-213,2019**

[4] O. Arriaga ,M.Valdenegro-Toro, P.G. Ploger, *"Real-time Con tional Neural Networks for emotion and gender classification",* In the Proceedings of European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, Bruges, Belgium, pp. 221-226, 2019

[5] L. F. d. Araujo Zeni and C. Rosito Jung, "*Real-Time Gender Detection in the Wild Using Deep Neural Networks*," In the Proceedings of 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Parana, pp. 118-125, 2018.

[6] N. M. Khalifa , M.H. N. Taha , A.E. Hassanien, H. N. E. T. Mohamed *"Deep Iris: Deep Learning for Gender Classification Through Iris Patterns",* ACTA INFORM MED. Vol.27, Issue2, pp**. 96-102,2019**. doi: 10.5455/aim.2019.27.96-102

[7] A. V. Savchenko "Efficient facial representations for age, gender and identity recognition in organizing photo albums using multi-output ConvNet ",PeerJ Comput. Sci, June, pp.**1-26, 2019.** DOI 10.7717/peerj-cs.197

[8] R.Ranjan, A. Bansal, J. Zheng, H. Xu, J.Gleason, B. Lu, A. anduri, J. Chen, C. D. Castillo, R. Chellappa " *A Fast and Accurate System for Face Detection, Identification, and Verification"*, Journal of Latex Class Files, Vol. 14, no. 8, pp.1-16, 2015

Authors Profile

Mr. Taran Rishit Undru is a student of Computer Science and Engineering in GRIET Hyderabad. He research interests include coding, image proceesing and Artificial Inetlligence. He is interested in sports, debates, and music. He has won several awards in sports at school level.

CVNS Anuradha is an Assistant Professor of Computer Science and Engineering Department, completed her M.TECH fromJNTUH in 2014. She has over 4years of academic experience in various institutions in India. She had earned Bachelors of Engineering in Computer Science and Engineering from JNTUH,in 2008. CVNS Anuradha is class coordinator for IV-Year F sec and exam coordinator for IV-Year and active member of NBA. Her areas of interests include Distributed Systems, Network Security, C programming. She is also a CISCO certfied Associate.

