Gender Facial Recognition by Convolutional Neural Network

1st Ahmed Naji Musleh Nusari
Department of Electrical and Electronic Engineering
Faculty of Engineering
Ataturk University
Erzurum, Turkey

2nd *Dr. Öğr. Üyesi Ferhat BOZKURT*
Department of Electronic Engineering
Faculty of Engineering
Ataturk University
Erzurum, Turkey
fbozkurt@atauni.edu.tr

***Abstract*—Various human attributes could be facially fetched, such as gender, age, and ethnicity. In many computer vision applications, gender classification can play a significant role, for the reason that could be appliable to many domains such as commercial marketing, surveillance, analyzing of facial expression, face recognition, and human-computer interactions. In this study, we used Convolutional Neural Networks (CNN) to create a classifier that can identify the gender of a human being through a face image. for the CNN training parameter we used a dataset with contains 35557 samples, the best performance rate was 89.42%.**

Keywords— Gender recognition, gender classification, deep learning, CNN.

I. Introduction

When it comes to classification major, CNN has achieved considerable results. For example face recognition, CNN has moved the ability of face recognition to advanced levels. The importance of face detection is gender recognition cannot be achievable unless the face has been detected in the picture. To explore the impact of face alignment methods, efforts have been made based on larger datasets on the performance of face recognition.

For machines, gender recognition function poses various challenges. More importantly, when it comes to dealing with image misalignment, poor quality, and obstruction. For such constraints, the distributions of pixel intensity may differ among the two genders' faces and such differences have been deployed to train binary classifiers. Zhou et al [4] presented a new method that processes gray level information using Principal Component Analysis (PCA) and Genetic Algorithm (GA) with previous preprocessing steps and forms a network of neurons on the resulting features.

In this study, we will create a classifier that can identify the gender of the facial image. We will use one of the deep learning methodologies which are CNN. By inserting a facial image for an unknown person the classifier will attempt to predict the person's gender.

II. Previous Studies

Most existing solutions to pattern recognition problems apply trainable or untrained classifiers preceded by heuristic feature extractors. This section briefly discusses previous work from the perspective of applied classification methods.

In 2017, Yaman AKBULUT and others used LRFELM and CNN as deep learning methods to create facial gender classification [1]. They used 13,239 images as their dataset. The face data set was generated for gender and age identification. The achieved performance rate for LRF-ELM and CNN was 80% and 87.13%, respectively.

Moreover, the human being has various daily activities such as shopping, running in nature, working, or picnicking. To automatize recognition of daily human demographical attributes, Sergiu Cosmin Nistor et al used ~70000 facial images gathered, merged, and annotated from the internet, intending to classify the gender and human attributes [2]. By using the CNN Inception-v4 network, the best-obtained accuracy was 98.2% on their dataset, and 84% on the Adience dataset.

In 2019, *Shiva Mittal et al* implemented the public LFW-Gender dataset, to investigate the reusability of Visual Geometry Group-16 (VGG16), which is an off-the-shelf CNN model that is pre-trained upon a large dataset of natural images [3]. Tuning a CNN segment on a medium-sized dataset provides better performance than peak approaches when implemented on the public LFW Gender dataset. The best performance rate obtained was 87.5%.

III. Material and Method

The main purpose of this study is to create a system based on artificial intelligence to identify the gender of a person. Naturally, a person can know various attributes of another human being from his image such as age, gender, and also race, however, to implement this with a computer we need a complex system to know what a given face represents [5].

Creating a gender classification system goes through stages: data processing, network training, and network performance testing. The mechanism of training the system is that we will batch and insert images into the created neural network, the network will extract the face features and classify them based on the gender given in the dataset - male or female - as shown in Figure 1

A. Data processing

On the Internet, there is a lot of data set, especially on face recognition topics. In this study, we used a data set containing 5 columns and 27,305 rows. Where the columns represent age, gender, race, image name, and image.

We chose this data because it contains various samples, ages, and ethnicity. As a result, we will have a lot of pictures of different ages and ethnicity which means the dataset contains people of different ages and ethnicity, which will make the system more practical and reliable.

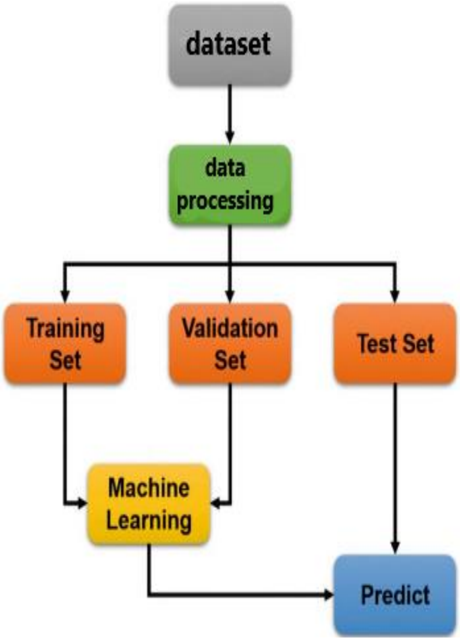

Figure 1: Gender Classification Stages.

To train the network we will not use all the existing data. We will only use the images and gender column; the rest of the columns will not be considered. In this data set, the male gender is represented by the number 1, and the female gender is represented by the number 0. 75% of the set will be randomly selected and used as network training data, and the remaining 25% will be used to check network performance. For validation, 50% of the data set will be selected randomly Figure 3 shows randomly selected face images from the dataset.

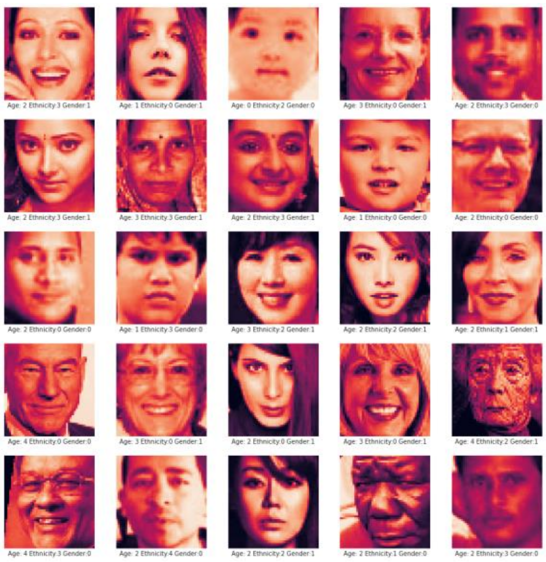

Figure 2: Dataset Face Images.

B. Network architecture:

CNN is a feedforward network composed of numerous layers of convolutional filters interleaved with subsampling filters, then fully linked layers.

In general, there are two types of layers in CNN: First, convolutional layers This layer contains many filters and equalizers, to extract the facial attributes of the input image [7]. Second, a fully connected layer is a set of linear equations that connect all perceptrons in one layer to all perceptrons in all layers, as shown in Figure 3.

Figure 4: an Example of CNN

The process of convolution layers can be described by:

- Training dataset in the size of 48x48x1 was applied to the network input layer.
- Applied images to the input layer, are processed by convolution process. The number of kernel filters and kernel size used in this step were selected as 3x3 and 75, respectively. The Padding and strides hyperparameters were 'same' and 1x1 respectively. This layer is named "conv2d_1" and its output is 48x48x75. Then, the "MaxPool2D" function was applied with pool size 2x2. After pooling, 25% of the pooling output was dropped out and layer output became 24x24x75.
- Similarly, “conv2d_2” was applied with hyperparameters (kernel size:3x3, number of filters:125, Stride 1, padding same) to the output conv2d_1. After that, with size 2x2 of "MaxPool2D" and 25% dropout functions were applied to the conv2d_1 output. The final "conv2d_2" output is 12x12x125.
- The fully connected layer inputs are the flattened output of "conv2d_2" and have two hidden layers: the first layer contains 500 perceptrons with 40% dropped out. The second layer has 250 perceptrons and is connected to the previous 500 perceptrons output layer. The number of network output will be equal to the number of classification, which is 2 in our system, the entire network parameters can be summarized as shown in Figure 4.

IV. Experimental Results

A. Training the Network

During the training network, the optimizer was 'adam', loss function was 'categorical_crossentropy', epochs was 60, batch size was 30 and the train set samples was 17778 samples. The training results were 0.970 for training accuracy and 0.091 for training loss. The validation set samples were 11852 samples. The validation results were 0.11 for validation loss and Val accuracy was 0.97 for validation accuracy.

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 48, 48, 50)	500
conv2d_1 (Conv2D)	(None, 48, 48, 75)	33825
max_pooling2d (MaxPooling2D)	(None, 24, 24, 75)	0
dropout (Dropout)	(None, 24, 24, 75)	0
conv2d_2 (Conv2D)	(None, 24, 24, 125)	84500
max_pooling2d_1 (MaxPooling 2D)	(None, 12, 12, 125)	0
dropout_1 (Dropout)	(None, 12, 12, 125)	0
flatten (Flatten)	(None, 18000)	0
dense (Dense)	(None, 500)	9000500
dropout_2 (Dropout)	(None, 500)	0
dense_1 (Dense)	(None, 250)	125250
dropout_3 (Dropout)	(None, 250)	0
dense_2 (Dense)	(None, 2)	502

Figure 4: CNN Network Parameters

The train and validation accuracy is shown in Figure 5, and network train and validation loss are shown in Figure 6.

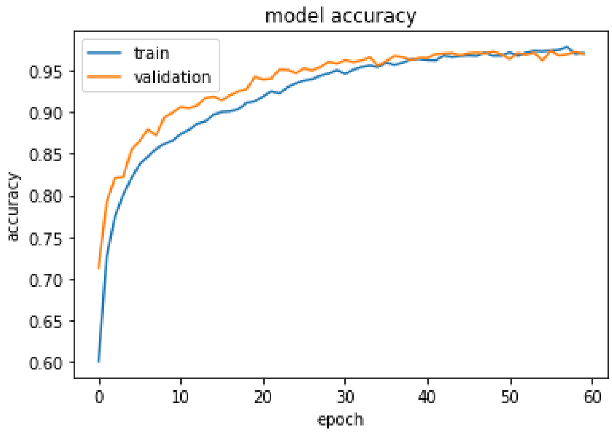

Figure 5: Train and Validation Accuracy

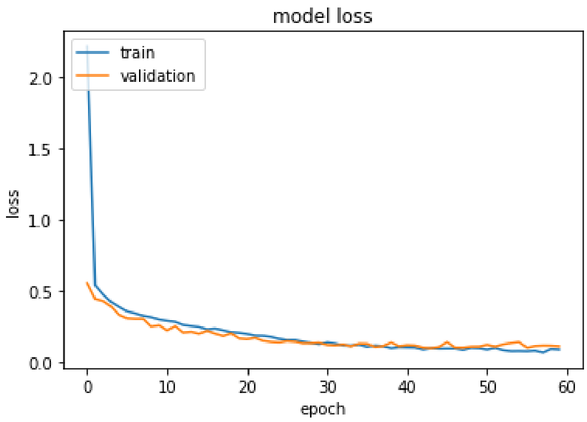

Figure 6: Train and Validation loss

B. Training Result

For testing network performance, we took 5927 samples randomly from the dataset and predict them through the network.

The model prediction showed True positive (TP), False negative (FN), False positive (FP) and True negative (TN) values 2751, 377, 250, and 2549, respectively, model confusion matrix is shown in Figure 8. The accuracy of the network was 89.42%.

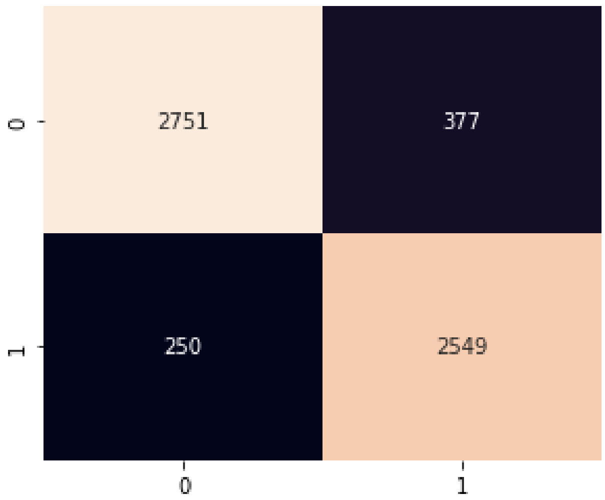

Figure 8: Model confusion matrix.

V. Conclusion and Future Work:

In this study, we created a classifier that can classify human gender through a facial image. The deep learning method that we used is CNN. Generally, CNN is a powerful network that can extract the features from images and represent them like weights. The performance rate of the classifier obtained was 89.42%. Compared with previous studies, our classifier is more effective and reliable, because our result is higher than theirs, also, we used more than samples to train the network. In the future, this classifier can be extended. By creating a classifier that can classify more than human features, such as a model that can classify human gender, age, status, and race.

Magnetization (A/m)" or "Magnetization {A[m(1)]}", not just "A/m". Do not label axes with a ratio of quantities and units. For example, write "Temperature (K)", not "Temperature/K".

References

[1] Akbulut, Yaman, Abdulkadir Şengür, and Sami Ekici. "Gender recognition from face images with deep learning." 2017 International artificial intelligence and data processing symposium (IDAP). IEEE, 2017.

[2] Nistor, Sergiu Cosmin, et al. "Automatic gender recognition for "in the wild" facial images using convolutional neural networks." 2017 13th IEEE International Conference on Intelligent Computer Communication and Processing (ICCP). IEEE, 2017.

[3] Mittal, Shubham, and Shiva Mittal. "Gender recognition from facial images using convolutional neural network." 2019 Fifth International Conference on Image Information Processing (ICIIP). IEEE, 2019.

[4] Zhou, Yuxiang, et al. "Face and gender recognition system based on convolutional neural networks." 2019 IEEE International Conference on Mechatronics and Automation (ICMA). IEEE, 2019.

[5] Dhomne, Amit, Ranjit Kumar, and Vijay Bhan. "Gender recognition through face using deep learning." Procedia computer science 132 (2018): 2-10.

[6] Lemley, Joseph, et al. "Comparison of Recent Machine Learning Techniques for Gender Recognition from Facial Images." MAICS 10 (2016): 97-102.

[7] KARAKAŞ, Eren, and İbrahim Yücel ÖZBEK. "Age Group and Gender Classification from Facial Images Based on Deep Neural Network Fusion." Erzincan Üniversitesi Fen Bilimleri Enstitüsü Dergisi 14.1: 150-158.

[8] Abbas, Farhat, et al. "Gender Classification Using Proposed CNN-Based Model and Ant Colony Optimization." Mathematics 9.19 (2021): 2499.

[9] Liew, Shan Sung, et al. "Gender classification: a convolutional neural network approach." Turkish Journal of Electrical Engineering & Computer Sciences 24.3 (2016): 1248-1264.