**Age, Gender and Ethnicity Detection System**

Tejaswini Parlapalli

Fall 2022 - INFO - H518

Faculty - Sunandan Chakraborty, Ph.D. SOIC, IUPUI - Indianapolis

# Introduction

# With the increase in the number of images available worldwide combined with an increase in the number of social media applications with ethical identity uses of consumers, there has been an exceptional demand for age, ethnicity and gender-detecting frameworks. Such frameworks can be utilized mainly in air terminals, police stations, customized product recommendation systems, to restrict age-based access, transportation stations, high-threat areas, and so on to tackle any potential threats, and to provide effective personalized recommendations.

# Despite their many uses, there exist few working models that effectively detect the age, gender and ethnicity of real-world images. The proposed system is developed to detect representations through the use of deep convolutional neural networks (CNN) and recurrent neural networks (RNN).

**Problem description**

Although biometric face recognition systems have been actualized, they are still met with a lot of skepticism and have not yet been universal. Also, Facial Recognition technology doesn’t always work as well as it should as it can be impacted by multiple external factors. With this project, we intend to simplify the detection of age, gender, and ethnicity of images from the dataset. Facial features are extracted by merging and comparing multiple models, and then a deep neural network is constructed to train and construct the combined features. In this way, the advantages of multiple models can be combined to mention the recognition accuracy.

# Data Description

# Our dataset comprises data in the form of CSV format. This includes a CSV of facial images that are labeled on the basis of age, gender and ethnicity. Our dataset consists of about 199.7 MB with 23315 unique values. It consists of 5 columns with age, ethnicity, gender, image name, and pixel value data needed to implement the detection system and has been extracted from Kaggle.

# The total number of pixels in each image is found to be 2304 which means it would be a 48\*48 matrix data set.

**Methodology**

Our project aims to detect a person's age, gender and ethnicity based on the features of their facial images. The dataset we choose consists of pixel values of images with labels of their respective age, gender and ethnicity. Keras models are proven to be better performing compared to the traditional machine learning models, so we planned on building different keras models and looking which model is having better accuracy in prediction.

**Environment Setup Google Drive & Colab**

Google colab was set up using a personal google account and data files were imported into the drive of the same account.

# Exploratory Data Analysis and Data Pre-Processing

# Before building the model, we performed necessary data cleaning steps of removing the rows with null values. We chose Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) models over fully connected network since CNN model performs better on image datasets and RNN performance on sequential data proven to be more accurate. First CNN model is built on the cleaned data and desired accuracy is not obtained. Here after looking at the results, we felt the data size we have used is not sufficient to train the model well and data augmentation is done. Since we have pixel data of the images to perform data augmentation, flipping the images will produce twice the data that already exists. Pixel data is first converted into pixel array of size 2304. This array is then reshaped into a matrix of size 48x48. Matrix is then flipped using fliplr function of numpy library and then reshaped back into array. This reshaped pixel array is appended to the dataset with original labels.

**Data Visualization**

Data visualizations are performed to understand the distribution of the data. There are almost equal numbers of rows for male and female, all groups of ethnicities are not of same size Group A is having very high data compared to any other group and Age is distributed in range 1 to 116 with age 20 to 30 having peak values.

**Model Building and Results**

A model is built and trained separately for age, gender and ethnicity. To do this age, gender and ethnicity columns are stored separately in different variables. These separated data is then split into test and train sets of 20% and 80% respectively. Initial model we built was CNN, to train gender and ethnicity data two CNNs are built 2 convolutional layers and for age 3 convolutional layers followed by max-pooling and activation function ‘relu.’ After fitting these models on train dataset, we got value accuracies of 89.59% for gender, 1.3% for age and 41.15% for ethnicity. Very low value accuracy for age model and not desired accuracy for ethnicity model were obtained.

Next step that followed by question why our accuracy for age is very low. After examination of data distribution, we got to know that age is a continuous variable distributed between 1 and 116. For approximately 23000 rows, age range of 116 will make it very less data to train a model to obtain desired results. Here we decided to make age from continuous variable to categorical variable by grouping age into 5 categories 0-15, 15-30, 30-45, 45-80 and 80+. Hence, we decided not to predict the exact age of the person but to predict the age group of a person.

After making age as a categorical variable, it is then split into test and train sets with same ratio of 20:80. The second CNN model is built with three convolutional layers each followed by batch normalization, max pooling and random dropout of 20% nodes. Due to the constraint of availability of GPU and huge runtime for models to train we used two functions of early stopping and learning rate reduction. Early stopping of training will be stopped when there is no desired improvement of accuracy (0.001) in our model is seen for 5 consecutive epochs. Learning rate reduction is done by the factor of 0.5 when there is no change in accuracy for 2 consecutive epochs and minimum learning rate is set to 0.00001.

This second built CNN model is called to fit age, gender and ethnicity. This model resulted in accuracies of 89.8% for gender, 71.2% for age and 78.8% for ethnicity. Here we can see a very good improvement in the model accuracies for age and ethnicity compared to the initial model with not much change in accuracy of gender. Random image is taken from the dataset, and we predicted the images age, gender and ethnicity using this model and compared with the actual labels. Predicted values are the same as the actual labels.

The third model that is built is RNN model using Long short-term memory (LSTM) and we used learning rate decay in optimizers to train the model on high learning rate and then reducing the learning rate. This model fits on all three variables age, gender and ethnicity separately resulting in accuracies of 86.9% for gender, 75% for age and 72.8% for ethnicity. While accuracy of gender and ethnicity are slightly decreased compared to CNN model, age accuracy has been increased. The reason for this improvement is RNN stores the data of previously passed data and age is sequential in our data, it better predicts the age. Random images are passed to RNN model and predicted values are same as the actual labels.

**Predicting User Input Image**

We tried to predict the age, gender and ethnicity of the user input image. We used one of our friend’s images where his face is filled with colors and making the image noisy. This image is initially reshaped into 48x48 size and pixel array data is derived from image and this array is passed to CNN model (since CNN has better accuracies) CNN predicted the age, gender and ethnicity of this user input data correctly even though image is noisy.

**Results**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Initial CNN Model** | **CNN Model** | **RNN Model** |
| **Age** | <1% | 72.1% | 75% |
| **Gender** | 89% | 89.8% | 86.9% |
| **Ethnicity** | 41.5% | 78.8% | 72.6% |

Accuracies of the models for three variables age, gender and ethnicity are as shown in the table above

# Limitations

# References

1. <https://www.kaggle.com/datasets/nipunarora8/age-gender-and-ethnicity-face-data-csv?resource=download>
2. <https://keras.io/api/optimizers/adam/#adam-class>
3. K. Zhang, C. Gao, L. Guo et al., “Age group and gender estimation in the wild with deep RoR architecture,” *IEEE Access*, vol. 5, pp. 22492–22503, 2017.
4. <https://kitchell.github.io/DeepLearningTutorial/7lossfunctionsoptimizers.html>
5. <https://keras.io/api/callbacks/reduce_lr_on_plateau/#reducelronplateau-class>
6. Zheng, Shuai, et al. "Conditional random fields as recurrent neural networks." Proceedings of theIEEEInternational Conference on Computer Vision. 2015.
7. <https://towardsdatascience.com/early-stopping-a-cool-strategy-to-regularize-neural-networks-bfdeca6d722e>
8. <https://www.tensorflow.org/tutorials/images/cnn>
9. K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” 2014, <https://arxiv.org/abs/1409.1556>.