# Age, Gender, and Ethnicity Prediction and Classification – CNN

George Washington University
DATS 6203: Machine Learning II
Final Group Project – Group 1
Individual Final Report
Instructor: Dr. Amir Jafari
Weirui (Rayna) Liu
04/27/2021

## I. Introduction

Facial recognition is now widely used in financial transactions, intelligent devices, and criminal investigations. We are interested in using a neuron network for automatic face recognition and inferring the person's characteristics in the image. This project used 2 Convoluted Neuron Network models on face recognition tasks, one for age estimation and others for ethnicity & gender classification.

I worked with Zheyue Wang on this final project. We were in charge of a model separately, from building, training, forecasting, to analysis. We optimized the model as much as possible and then combined the two models to predict the people's age, gender, and ethnicity by image. I worked on the age model, and my partner worked on the ethnicity & gender model. We completed the final group report together, and then each of us was responsible for half of the PPT and presentation according to the report's content.

## II. Individual Work

Since our dataset has three features, we have discussed whether to make a multiple-class model or three models. At that time, I put forward the idea that it would not be easy for a neural network to make age predictions because people's aging frequency may vary by ethnicity differences, gender differences, age differences, and even individual differences. So, using one model to produce three outputs will make the model inaccurate in predicting individual features. On the other hand, gender and ethnicity both are classification problems, we could train them together by using multiple-label classification, and we want our forecasts to be gender accurate and ethnicity accurate—that why I proposed to use two models in our project.

I am mainly responsible for the age model. In the beginning, when I did the visual analysis of age data, I found that the distribution of age data was very uneven. The number of 26-yea-old is the largest, and the number of 1-year old is also more. Missing age also exists in the 23,705 images, all of which are concentrated after 95 years old. There are few older people in these images, most of whom are between the ages of 20 and 60. So in my work in splitting training and testing, try to stratify by age. However, in some age only have one image, it could not do

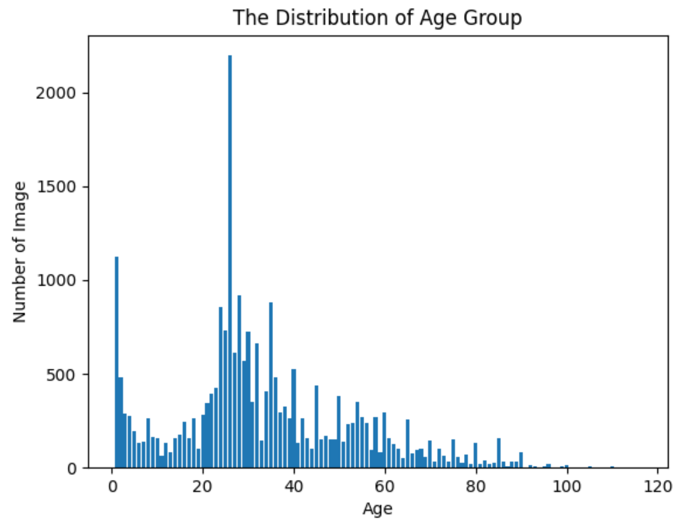
Out[1]:

| | age | Number of Image |
|---|---|---|
| 0 | 1 | 1123 |
| 1 | 2 | 482 |
| 2 | 3 | 289 |
| 3 | 4 | 273 |
| 4 | 5 | 196 |
| .. | ... | ... |
| 99 | 105 | 5 |
| 100 | 110 | 5 |
| 101 | 111 | 1 |
| 102 | 115 | 3 |
| 103 | 116 | 4 |

[104 rows x 2 columns]

| | age | Number of Image |
|---|---|---|
| 25 | 26 | 2197 |
| 0 | 1 | 1123 |
| 27 | 28 | 918 |
| 34 | 35 | 880 |
| 23 | 24 | 859 |
| .. | ... | ... |
| 102 | 115 | 3 |
| 90 | 91 | 2 |
| 97 | 101 | 2 |
| 98 | 103 | 1 |
| 101 | 111 | 1 |

[104 rows x 2 columns]

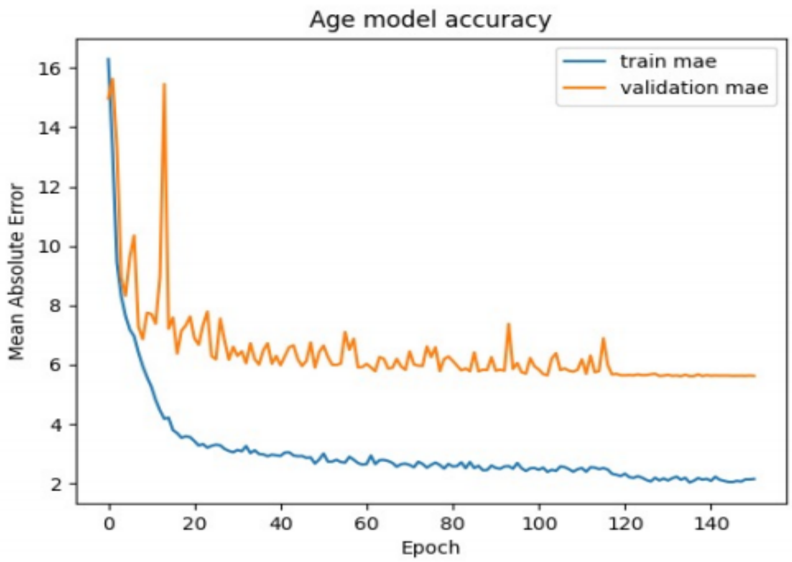



the stratify. If I remove that part of age, although I can do the stratify, my age model cannot catch the features of the images on these ages well. Therefore, I tried the data augment plan to make up for the uneven data volume of different ages by generating pictures through rotation, scaling, etc. I spent several days trying data augmented, but I found that after data augmented, the MAE value of the age model converged around 6 and did not appear to decline again. Before that, I was convinced that data augmentation would improve the model, so I kept changing the parameter in *ImageDataGenerator*. However, it still did not achieve the desired effect.

Before constructing the age model, I knew that age estimation by using neural networks is a scorching topic, and there are many articles for reference. I was very interested in CNN at that time, so I referred to the article *Deep Convolutional Neural Network for Estimation based on the VGG-face model* 1 to build the architecture of my age model since the architecture of the CNN model for age estimation in the article has been already optimized based on VGG-Face model.

## III. Result

After much training, I adjusted various parameters on this framework of age model. For example, add a more convoluted layer. Finally, my age model MAE dropped from 6.1 to 5.5. Since my age model takes more than half an hour to train, I adjust the parameter is not arbitrary but based on some literature. For example, I refer to the article *How to Choose the Right Mini-Batch Size in Deep Learning when choosing the batch size*. 2 To handle the overfitting problem, I used Early Stopping and Dropped out. I add a dropout layer after the last two flatten layers. Fore early stopping, when the validation loss stops improving for 25 consecutive epochs, it will stop. I also used a performance schedule to adjust the learning rate dramatically increases the utility of the model. The learning rate 1-E3 is decreased by a factor of 0.105 as soon as the validation loss has not improved for 15 consecutive steps.3




From the previous experiment, the learning rate and epochs number should be changed simultaneously, a lower learning rate needs more epochs, and a higher learning rate needs fewer epochs. A learning rate greater than 1e-2 is significant in the age 11 model, causing the model to converge too quickly to a suboptimal solution. Learning rate less than 1e-5 is too low to cause the process to get stuck. After 100 epochs, Age's MAE remains below 6.0 and gradually decreases as the epoch increases. However, the training time of the age model is as long as half an hour. We finally set the learning rate to 0.001 with 200 epochs.

After repeated training of the model, I find that the MAE of age is stable at about 5.7.

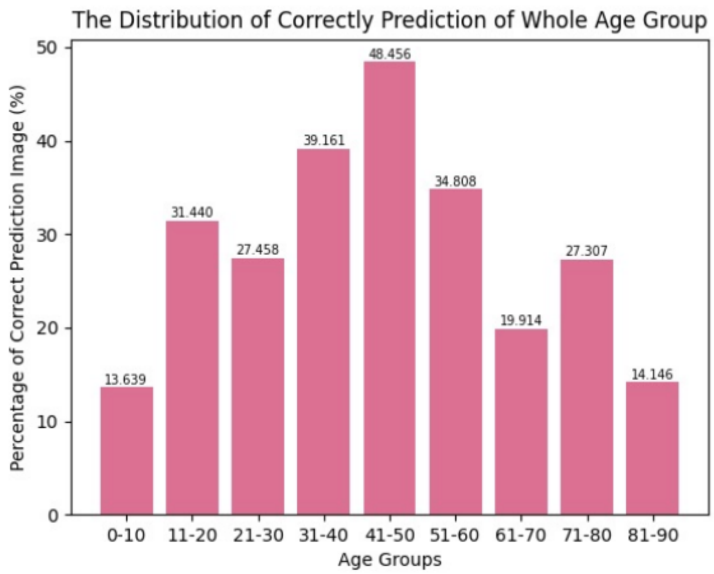



Figure 8: The Distribution of correct prediction

Above is the distribution of correctly prediction of whole Age Group, since no people over 90 were be predict correctly, this graph did not show the 91-100, and 100+ group. Actual, this distribution will have a large vary depend on the test set. Cause as I remind before; the age data is much imbalanced.

| | age | Number of Image | Sum of Correct Image | Accuracy | Age Groups |
|---|---|---|---|---|---|
| 0 | 1 | 1123 | 4 | 0.356189 | 0-10 |
| 1 | 2 | 482 | 2 | 0.414938 | 0-10 |
| 2 | 3 | 289 | 3 | 1.038062 | 0-10 |
| 3 | 4 | 273 | 13 | 4.761905 | 0-10 |
| 4 | 5 | 196 | 18 | 9.183673 | 0-10 |
| .. | ... | ... | ... | ... | ... |
| 86 | 87 | 10 | 1 | 10.000000 | 81-90 |
| 87 | 88 | 34 | 3 | 8.823529 | 81-90 |
| 88 | 89 | 33 | 1 | 3.030303 | 81-90 |
| 89 | 90 | 82 | 1 | 1.219512 | 81-90 |
| 90 | 91 | 2 | 1 | 50.000000 | 91-100 |

| | age | Number of Image | Sum of Correct Image | Accuracy | Age Groups |
|---|---|---|---|---|---|
| 25 | 26 | 2197 | 766 | 34.865726 | 21-30 |
| 27 | 28 | 918 | 312 | 33.986928 | 21-30 |
| 23 | 24 | 859 | 298 | 34.691502 | 21-30 |
| 29 | 30 | 724 | 265 | 36.602210 | 21-30 |
| 34 | 35 | 880 | 261 | 29.659091 | 31-40 |
| .. | ... | ... | ... | ... | ... |
| 85 | 86 | 35 | 1 | 2.857143 | 81-90 |
| 86 | 87 | 10 | 1 | 10.000000 | 81-90 |
| 88 | 89 | 33 | 1 | 3.030303 | 81-90 |
| 89 | 90 | 82 | 1 | 1.219512 | 81-90 |
| 90 | 91 | 2 | 1 | 50.000000 | 91-100 |

From above table we can see that for age model the best predictor was 26, which also had the most pictures. It has large number of images to let the CNN capture the characterize of feature.

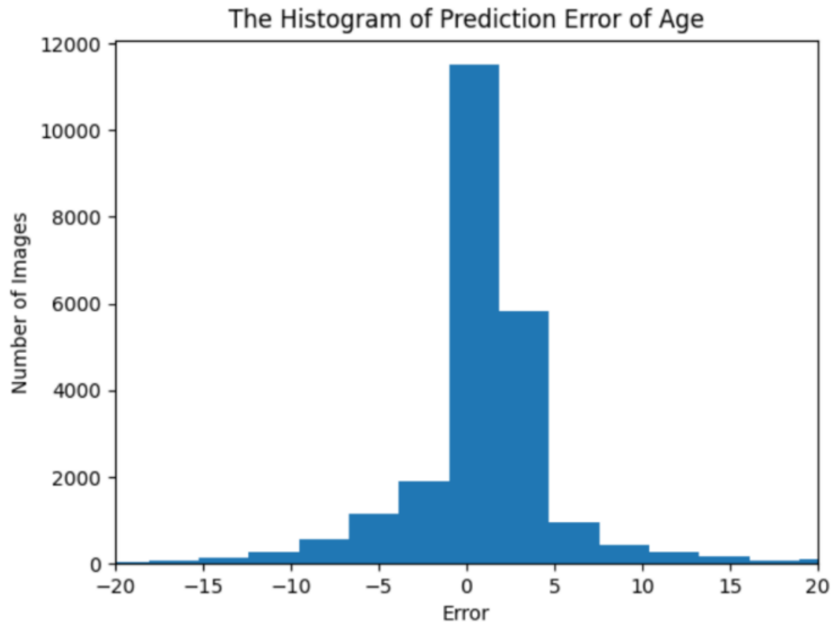
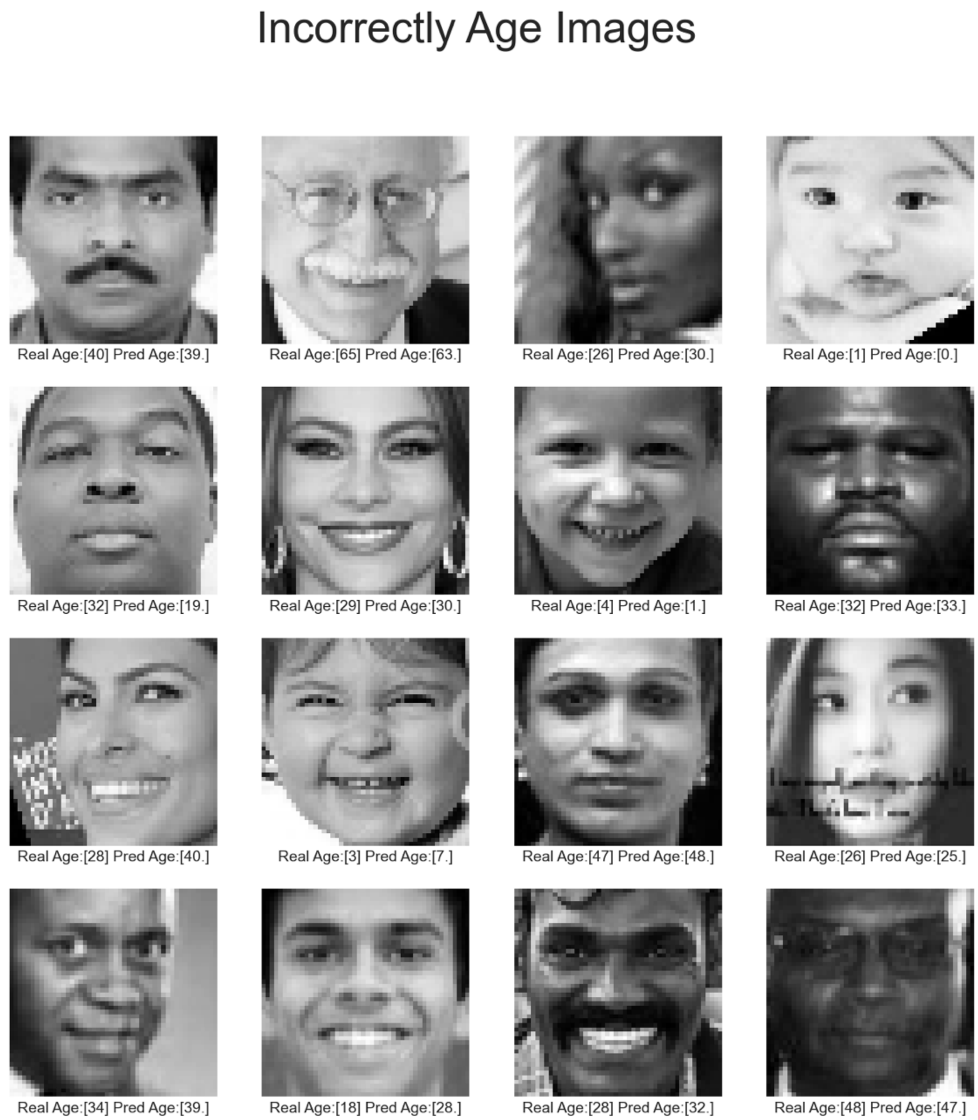
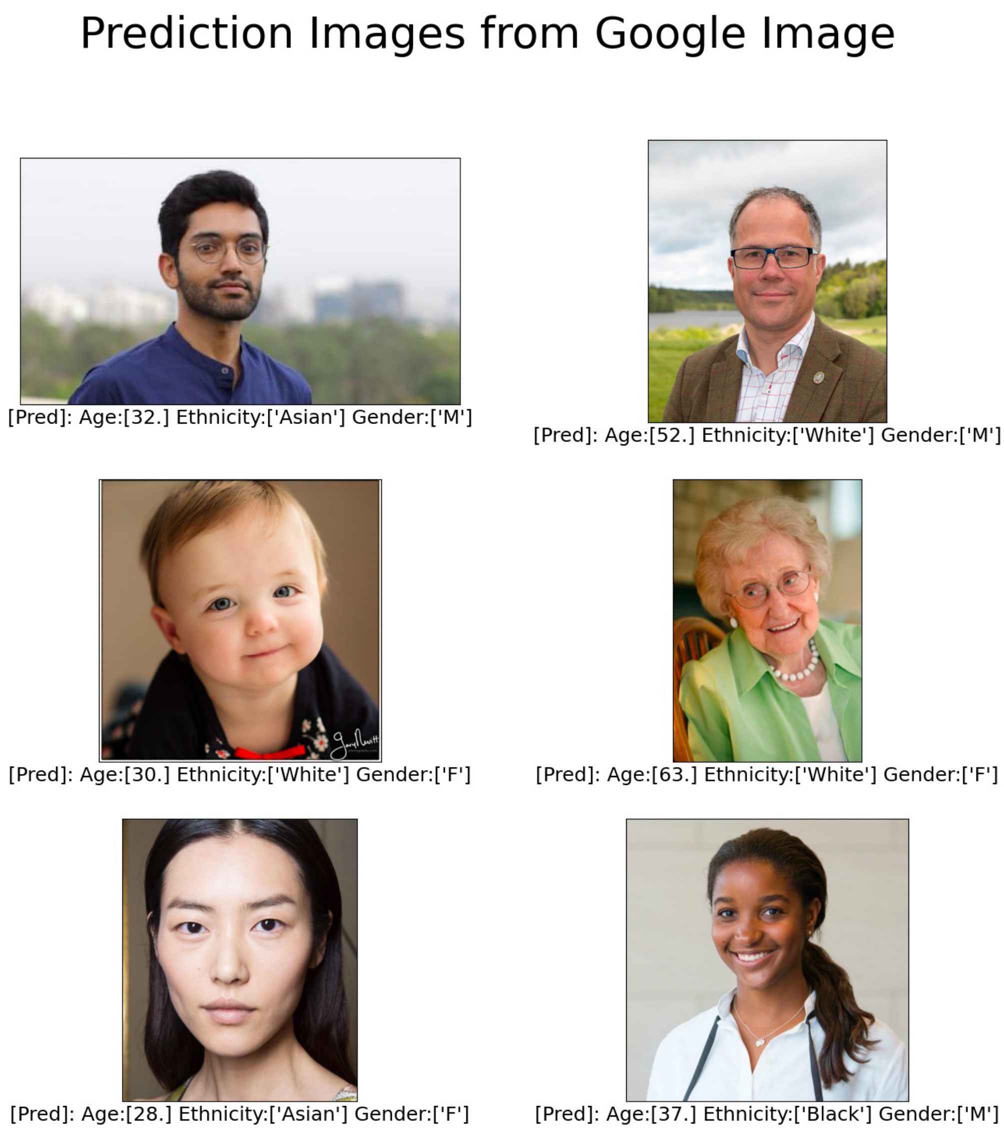



Figure 10: The histogram of prediction error of age

The Histogram of predicted error of age is followed the normal distribution. The graph intuitively shows that the average error of the age-predicted by our age model compared with the actual age is less than 5 years. That is to say, if our model gives an image label a person as 28, the predicted age probability will be between 26 and 31.




For most adults, the margin of error for age prediction was about a year. After several random checks of incorrectly age images, we found that our age model could hardly predict the age of children (less than 5) accurately and had significant errors because the model could not well identify the age characteristics of children. Once there were some odd expressions, like grimaces, there would show a significant margin of error.




## IV. Summary & Conclusion

One of the biggest challenges in age estimation is that humans age grows at different rates, with different races and genders showing different characteristics at the same age.

I looked up some literature and worked hard to develop a deep neural network to let it catch up with the age characteristics as much as possible. Since our input image size is (48x48x1), which only has one channel. It is hard for us to resize to 3 channels to fit the pre-trained model VGG16 or ResNet. In the future, I still want to try data augmentation and try to figure out to change the channel from 1 to 3 and use the pre-trained model to improve the accuracy.

## V. Percentage of Code

(19-14)/(19+17) *100 = 13.88%

### References:


1. Deep Convoluted Neural Network for Estimation
   https://arxiv.org/pdf/1709.01664.pdf.

2. How to choose the right mini-batch size in deep learning
   https://www.mikulskibartosz.name/ how-to-choose-the-right-mini-batch-size-in-deep-learning

3. Amir_Jafari: Deep Learning
   https://github.com/amir-jafari/Deep-Learning/blob/master/Exam_MiniProjects/9-Keras_Exam1_Sample_Codes_S21/load_data.py

4. Kaggle: Age, Gender & Ethnicity Prediction
   https://www.kaggle.com/rithikb24/age-gender-ethnicity-prediction.