Age and Gender Prediction

Jakkaphob Kongthanarith
Assoc.Profr. Ekachai Phaisangittisagul
Department of Engineering, Faculty of engineering, Kasetsart University
Call 0-2942-8555, E-mail: fengecp@ku.ac.th

Abstract

From the past to present, security has been improved time by time to prevent the crime from happening and also bring the criminal to justice. There are several methods for detect and identify the identity of people, one of them is detection by gender and age.

The project goal is to find the best deep learning model for predicting age and gender using face image of the target. The model build can also be used for further recognition procedure for those who are interested.

Keywords: Deep learning, Age estimation, Gender dentification

1. Introduction

To train the model, one must specify and choose the dataset that will fit their task. Data cleaning and augmentation are recommended since we need to avoid the bias of the dataset and incompatibility of the model. I want to create a model hoping for them to be able to predict the human age and gender with the performance as high as the situation allows.

2. Related Theories

2.1. Deep Learning

A technique which is a part of machine learning method that uses a neural network along with feature learning, which is a technique for the system to discover a pattern of each image that detect and classify pattern of features from the data.

2.2 Flatten Layer

Flatten layers are used to convert the multi-dimensional input into 1 dimensional array. Resulting in conversion of fully connected layer to output of convolutional or pooling layer. With this, we can get the single output we expected.

2.3 Activation function

After training the model, we determine the output type and range by using activation function. It adds non-linearity to the model, allowing it to find and learn the complicated pattern and relationship between inputs and outputs. Sigmoid and ReLU activation function are a part of what I used in the project, while the former converts the input a value range between 0 and 1, which is one of the well-known method for predicting age. The latter, linear activation function simply returns the output without much adjustment. It is commonly used in regression tasks where the output can take any real value.

2.4 Model Evaluation

A procedure to observe whether the model prediction performance reaches the expected result or not. The evaluation method is varied by the goal of the project

2.4.1 Confusion Matrix

A table to evaluate the performance of the model, The size of the table will be compared with the predicted result and actual result. By using Confusion matrix. We can evaluate the result of the model by judging from the right and wrong answer. The higher amount of ‘True Positives’ and ‘True Negatives’, the higher the performance of the model is.

	Actually Positive (1)	Actually Negative (0)
Predicted Positive (1)	True Positives (TPs)	False Positives (FPs)
Predicted Negative (0)	False Negatives (FNs)	True Negatives (TNs)

Figure 1 Confusion Matrix

True Positive (TP) - Both prediction and actual result is match with the first condition

True Negative (TN) = Both prediction and actual result is match with the second prediction

False Positive (FP) = The prediction doesn't match with the actual result

False Negative (FN) = The actual result doesn't match with the prediction

The TP,TN,FP,FN can be use to calculate the Accuracy, Precision with the formula mentioned below

Accuracy = (TPs + TNs) / (TPs+TNs+FPs + FNs) (1)

Precision = TP/(TP+FP) (2)

Sensitivity = (TP)/(TP+FN (3)

Specificity = (TN)/(TN+FP) (4)

2.4.2 Accuracy

It is a method I use to evaluate gender prediction problem, it compare whether the predicted value match with predicted value or not, after the it will take all amount of matched value divided by the max number of file and multiply by 100 to get the accuracy of that gender model.

2.4.3 Mean absolute error

In short for MAE. It is a method used to evaluate the average difference between predicted values and the actual values. I used it to evaluate my age model performance because to the ages ranged is very difficult to pin-point the exact values dues to various factor such as Similar face structure of babies, those who have their faces not developed along with their ages, achene, or stain on their faces, etc. I will use the MAE to indicate the overall difference of the prediction result of my model and the actual ages of the Dataset.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|Xi - X| \quad (5)$$

X_i and X can be swapped since after differentiation of each, the value will be convert into positive value before being kept in the formula.

2.4) Percentage difference

A method to calculate the difference between predicted output and actual output in terms of percentage. It is useful to indicate the model output that can't be evaluated by ordinary right or wrong correction method. The output percentage difference inversely varied to the model performance.

$$Percentage\ difference = \frac{(|Value1 - Value2|) * 100}{(Max\_age)} \quad (6)$$

Max_age refers to the maximum number in the age range, the formula computes the percentage difference between the predicted ages by scaling the absolute difference by the maximum age value.

2.5 Parameter initialization

Initialize the model parameter (or weights) to the values. The parameter can be set randomly, or it can be set to recommended value by the pretrained model we used.

3. Methods

3.1 Prepare the dataset

UTK face Dataset contain 23708 images faces with all age range and both male and female I choose this dataset because the content mostly are straight face. I set the shape of both train and test Data set is (None,224,224,1) and (None,224,224,3) depend on the model requirement. I intended to use 25 straight face image from male and female for age range 1-79.

3.2 Image Augmentation

Since UTK face don't have enough file for all age range, I have to perform an augmentation to fill the lacking images files. The augmentation method that I use is Enhance

brightness and contrast. Both of them only affect the light not the face features that might have confused the model otherwise.

3.3 Normalization

Normalization is a process that changes the range of pixel's intensity into different values range but contains the same ratio for other pixels in the images. In my case, I divide the values of image pixels by 255 to rescale the pixel values of an image to a fixed range of 0 to 1. Doing this will result in better contrast detection of a model, lower over fitting problem, and compatibility of the model that require a fixed size of image.

3.4 Split the dataset

The ratio when splitting is 8:2 for training and testing, I got a total of 3160 training dataset and 790 testing dataset for my model

3.5 Neural Network Age and Gender Combine model

.

For age part, Neural network contain an input layer to receive (224,224) gray scale image for both and and gender part. For Age part, it have 2 Dense with ReLU, 2 Droupout layer and 1 output layer of Linear activation. For Gender part, it have 1 Flatten layer, 3 Droupout layer, Dense with ReLU layer and 1 output layer of Sigmoid activation.

3.5 CNN Age and Gender Combine model

I design for CNN to have a base layer that have an input layer, 3 Conv2D layer, 2 Maxpooling layer, 1 Droupout Layer, 1 Flatten layer which will be the continue point of age and gender part. After flatten the model proceed to age part, having 4 Dense with ReLU Layer and output layer with Linear activation. For Gender Part it consist of 5 Dense with ReLU Layer, 5 Dropout layer and output layer withSigmoid activation

3.6 VGG-16 Pretrained+fine-tuning Age and Gender Combine model

Since I use a pretrained model as a base layer. I freeze its layer and use its output to be the initial point for Age and Gender part. For the age part, it has 2 Dropout layers, 2 Dense with ReLU Layers, Batch Normalization layer, 1 output layer with ReLU activation. For Gender part, it have 1 Flatten layer, 2 2 Dropout layer, 2 Dense with ReLU Layer and 1 output layer with Sigmoid Activation.

Table 1 Comparison of Age model

Model	Train MAE	Test MAE
Neural network	18.4853	17.1355
CNN	10.5386	15.1492
Pretrained	7.4472	9.9541

Table 2 Comparison of Gender model

Model	Train Accuracy	Test Accuracy
Neural network	0.8256	0.7709
CNN	0.8513	0.8468
Pretrained	0.8570	0.8494

Judging from the result, the best model that has lowest test MAE and Highest Test accuracy belongs to Pretrained model that use VGG-16 as a base model while the second best is Convolutional Neural Network. This shows that the amount of epoch and batch size are effective when training then model.

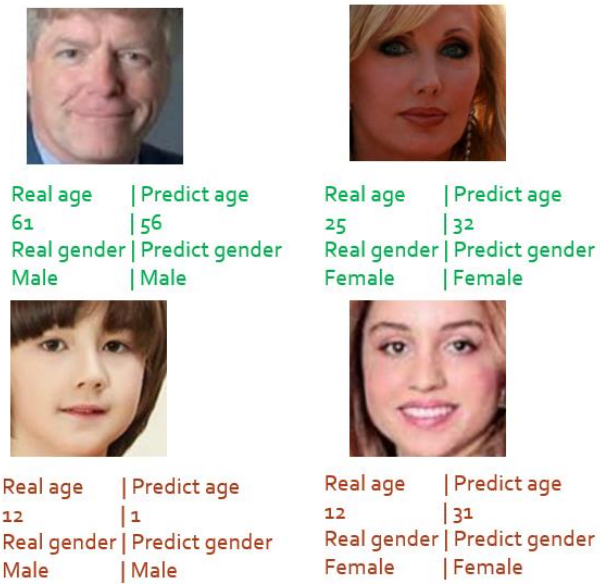

Fig 3 Prediction result

From the example of prediction result shown above, the gender prediction that classify the regression value into ‘male’ or ‘female’ can done the job with out problem, while the age part is struggle quite a bit since there are some case such as youthful face that cause a confusion to the model. While some case, the model are just predicting the value completely difference from the actual value, this represent the variety of how human face feature can be.

5. Conclusion and Suggestion

It is reasonable that VGG-16 Pretrained model get the best result, because unlike the two model, VGG-16 has already been trained beforehand by over million files of image net, so in terms of performance and stability, VGG-16 pretrain model is the most decent one. The second is CNN which has its accuracy slightly lower than pretrained model by a bit while the MAE is clearly higher.

Even though CNN has been introduced since 1980 but its still widely used in various tasks because of its ability to extract the features, finding pattern and can learn the features of the data by itself. Neural networks which have its performance lowest of the three are not mean the model is bad, but due to limited GPU resources in Google Colab, the number of batch size and epoch are being limited to the point that the model can’t even be trained properly when its age and gender part is being combined and give a poor result.

Acknowledgement

First, I would like to express my sincere gratitude for the advice, resources, and comment of my advisor, Assoc. Prof. Dr. Ekachai Phaisangittisagul. I would not have come this far without him.

Second is my project committee, Prof. Dr. Wiroonsak Santipach for his kindness toward my project. And also, the rest of the people who have supported me from time to time.

7. References

[1] Syed Taskeen Rahman, Asiful Arefeen, Shashoto Sharif Mridul, Asir Intisar Khan, Samia Subrina “Human Age and Gender Estimation using Facial Image Processing” January 2020

[2] Jun-Cheng Chen, Amit Kumar, Rajeev Ranjan, Vishal M. Patel “A cascaded convolutional neural network for age estimation of unconstrained faces, December 2016

[3] Koichi Ito, Hiroya Kawai, Takehisa Okano, Takafumi Aoki “Age and Gender Prediction from Face Images Using Convolutional Neural Network”, 07 March 2019

[4] Kamekshi Gaur, Dr. Pawan Kumar “Literature Review on Gender Prediction Model using CNN Algorithm” , Mar 2022

[5] S.A. Khan, Munir, Ahmad, Muhammad Nazir, N.Riaz “A comparative analysis of gender classification techniques”, January 2014