Comparison “Age and Gender Classification using Convolutional Neural Networks”

Understanding and Evaluating the Models

Camilo Lagrava*
Universidad Privada Boliviana
La Paz, Bolivia
camilolagrava@gmail.com

Nuria Michel*
Universidad Privada Boliviana
La Paz, Bolivia
nuria.l.michel.l@gmail.com

Lizzeth Vásquez*
Universidad Privada Boliviana
La Paz, Bolivia
lizvasmen@gmail.com

Abstract. - The aim of the following paper is to test the predictive capability of the model used in the paper[1] by comparing it in terms of its accuracy with that of a more up-to-date model[2] using the same CNN(convolutional neural network) approach.

Keywords—age, gender, classification, CNN, dataset, Adience benchmark.

I. Models

- Model 1

Convolutional Layers:

- Layer 1: Input 227 x 227 pixels, 96 filters of size 3x7x7 pixels, followed by a ReLu function, max pooling of 3x3.
- Layer 2: Input 96x28x28 pixels, 256 filters of size 96x5x5 pixels, ReLu, max pooling with the same hyper parameters.

Layers fully connected:

- Layer 3: Input 256x14x14 pixels, 384 filters of size 256x3x3x3 pixels, ReLu, max pooling.
- Layer 4: Consisting of 512 neurons, ReLu, dropout.
- Layer 5: Consisting of 512 neurons, ReLu, dropout.
- Layer 6: Map the layers with the final classes "age" or "gender".

Output Layer:

- Layer 7: softmax that assigns the probability of each class, the prediction is made by taking the class or the highest probability.

- Model 2

Convolutional layers

- Layer 1: Input 227x 227, num_output: 96, kernel_size: 7, ReLu, max-pooling 3x3, Local Response Normalization
- Layer 2: num_output:256, kernel_size: 5, ReLu, max-pooling 3x3, Local Response Normalization
- Layer 3: num_output:384, kernel_size: 3, ReLu, max-pooling 3x3

Layers fully connected:

- Layer 4: num_output:512, ReLu, dropout de 0,5.
- Layer 5: num_output:512, ReLu, dropout de 0,5.
- Layer 6: num_output: 2 (for gender), 8 (for age).
- Last layer: SOFTMAX

II. Experiments

The experiments performed were an evaluation of both models with a small sample. We decided to create our own "dataset", which contains 100 varied images in jpg format.

Why we chose such an unconventional method was because of the limited time we had available, the model [1] is a convolutional network that is trained twice. The paper [1] specifies that it took about four hours of training on a superior, more powerful computer than we had. We trained the convolutional neural network in about 6 hours, which was an obstacle when performing experiments. With this problem it was not possible to retrain the architecture of both models, which made use of the "Adience Benchmark"[1] database, so the aforementioned solution was chosen.

Images of faces were chosen that a human could not easily classify; and also, that the models could not easily match.

III. Results

With our own data set the results will be presented below:

GROUP 6
* The Members worked jointly and equally
Selected Topics in Artificial Intelligence 2021. Professor Gissel Velarde

	MODEL [1]	MODEL [2]
Successful age and gender classification	14	11
Successful classification in gender	62	51
Successful age classification	26	22
Misclassification but approximate by a gender class	21	11
Classified images	100	73

Table 1. Number of classified images out of 100 samples.

The first model presented better results in many aspects, initially, it had a little more advantage in the qualification of both age and gender. It is very proficient in age recognition although a little weak in gender recognition, in addition, it was able to process all the images without size problem surpassing the most recent model [2], which was not able to process many of the images and with a larger margin of error.

IV. Conclusions

We could notice that the literature model [1] presents a better performance compared to the second model [2], despite the fact that both implementations are convolutional neural networks. However, the recognition capability of the model [1] is flawed when predicting images of paintings, people with a lot of makeup and bust sculptures. In addition, it tends to err on the side of gender in the age range of four to twelve years. We believe that this is because both gender and age tasks were trained under the same network, which gives better results for one domain and lowers the prediction quality for the other.

References

[1] Levi, G., & Hassner, T. (2015). Age and gender classification using convolutional neural networks. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 34-42).

[2] Arafat Fahim. (2021). Age & Gender Detection by Python With Google Colab | Bangla 2021 [YouTube Video]. In YouTube. https://www.youtube.com/watch?v=SZOYJq1b_C0

[3] droidadroit. (2021). GitHub - droidadroit/age-and-gender-classification. GitHub. https://github.com/droidadroit/age-and-gender-classification

[4] datasetuyuy - Google Drive. (2011). Google.com. https://drive.google.com/drive/folders/1LZWgGPb6JxiIzNx0tWw98ahdFDxSvXcc