Apparent Age Estimation

Wallace S. Loos

March 13, 2019

1 Introduction

The apparent age estimation is a regression problem established in order to estimate the age apparent and not the real (or biological) age. In the real age estimation, each image has only one age associated to the image. In the apparent age estimation, each age can have more than one age associated to the same image [4]. This is a regression problem since the age is from a continous domain. Some of the applications of the apparent age estimation are: customer profiling, security, and entertainment[1, 4]. In this report, we present a simple deep convolutional neural network to estimate the apparent age given the image of the person.

2 Network Architecture

Deep Convolutional Neural Network (CNN or ConvNet) has been used widely in computer vision tasks [6, 8]. Our CNN use the inceptionv3 architecture [7], followed by a convolutional layer (3x3x64), dropout (rate = 0.5), bacth normalization and two fully connected layers (FC), one with 512 neurons and the other with 2 neurons. The first and second FC layer are followed by a rectified linear unit (ReLU). For a good performance, usually deep CNN's need large training datasets. Since our dataset is small, we cannot train a deep CNN from scratch. However, we can pretrain our CNN on a very large dataset. This is called *transfer learning*. To overcome the limited dataset we pretrain our CNN on Imagenet dataset.

Imagenet contains more than 2 million images with 1000 categories [2].

The dataset used is from ChaLearn LAP competition on Apparent Age Estimation. Each age is the averaged opinion by many people, the standard deviation is also provided for each image [5]. The validation and test set were from ChaLearn LAP competition on Apparent Age Estimation 2015 and the training set from ChaLearn LAP competition on Apparent Age Estimation 2016. The size of the trainning, validation, and test set were 4113, 1136, and 1079 respectively. For each image, the haar-cascade detection provided for OpenCV is used for face detection. After the detection of the face, the image is cropped and resized to 224×224. We train this architecture for about 20 epochs on the training and validation datasets. The input image resolution was fixed in (224×224×3). Throughout training we use a batch size of 64 and the loss function mean square error (MSE). The optimizer used was *Adam* [3]. Our final layer predicts the age and the standart deviation of the person.

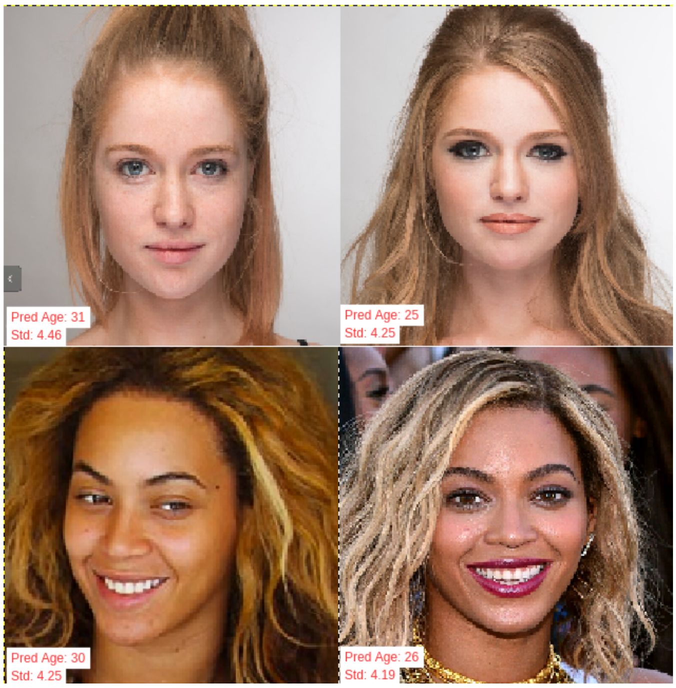

Figure 1: Makeup off (left column) and makeup on (right column).

3 Result

To measure the performance of the model the mean absolute error (MAE) and the ϵ–*error* were used. The ϵ–*error* is given by the formula

$$\epsilon - error = 1 - e^{-\frac{(\hat{y}-y)^2}{2\sigma^2}} \tag{1}$$

where $\hat{y}$ is the predicted age and y and σ are the annotated age and standart deviation. The ϵ–*error* value is between 1 (worst) and 0 (best). This metric was used in the ChaLearn LAP competition on Apparent Age Estimation. For the test set the model got MAE = 9.3368 and ϵ–*error* = 0.6524. The winner of the ChaLearn LAP competition on Apparent Age Estimation 2016 got ϵ–*error* = 0.2411.

4 Discussion

The low performance of the model is associated with the size of the dataset and the umbalanced of it. The model fails to predict ages outside of the range 20–30 as presented in Figure 2, and has more success to predict ages between 20–30, as presented in the cover figure. The use of fine-tuning and other dataset, like IMDB-Wiki dataset, can improve the performance of our model. Combining diferent models can also result in a better performance. Some works treated the apparent age estimation like a classification problem and got good results. The output layer has 100 neurons, 0–99, and the model give the probability of the age. Figure presents how a makeup can change the apparent age of the person, so, makeups makes the difference.

5 Conclusion

We introduced a simple architecture to solve the apparent age estimation. Despite the low performance of the model, it helped to understand how to use

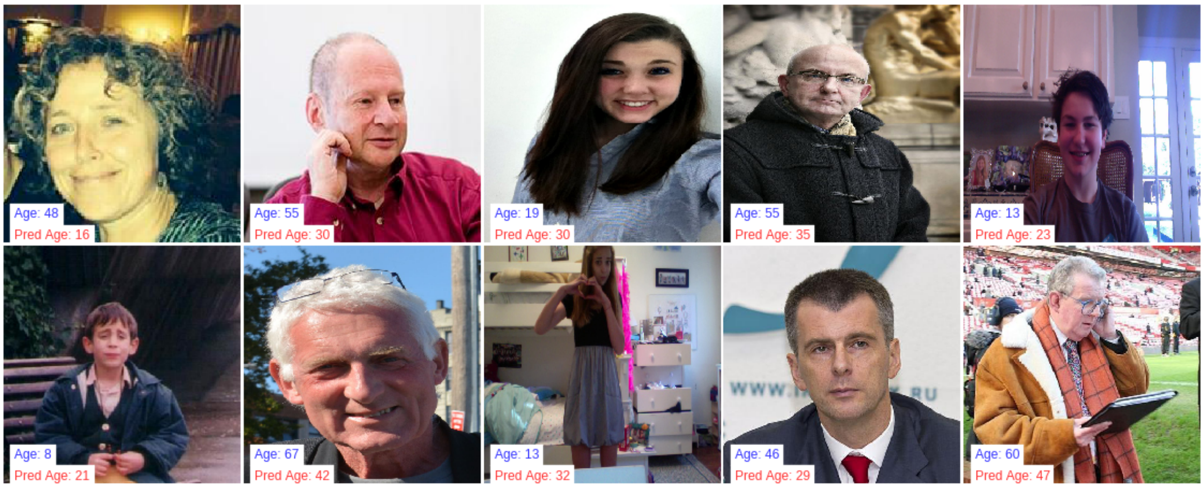

Figure 2: The model fail to predict ages from 0–19 and 40–80.

deep learning to solve a linear regression problem. The codes are available on github.

References

[1] Grigory Antipov, Moez Baccouche, Sid Ahmed Berrani, and Jean Luc Dugelay. Apparent Age Estimation from Face Images Combining General and Children-Specialized Deep Learning Models. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pages 801–809, 2016.

[2] J. Deng, W. Dong, R. Socher, L. Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, June 2009.

[3] Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. pages 1–15, 2014.

[4] Refik Can Malli, Mehmet Aygun, and Hazim Kemal Ekenel. Apparent age estimation using ensemble of deep learning models. *CoRR*, abs/1606.02909, 2016.

[5] Rasmus Rothe, Radu Timofte, and Luc Van Gool. Deep expectation of real and apparent age from a single image without facial landmarks. *Int. J. Comput. Vision*, 126(2-4):144–157, April 2018.

[6] Karen Simonyan and Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. pages 1–14, 2014.

[7] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, and Zbigniew Wojna. Rethinking the Inception Architecture for Computer Vision. 2015.

[8] Songtao Wu, Shenghua Zhong, and Yan Liu. Deep residual learning for image steganalysis. *Multimedia Tools and Applications*, pages 1–17, 2017.