Gender and Age Group Prediction using CNN and Transfer Learning

Aarav Madan (1209374462), Laxmikant Patil (1209141566) CIDSE, Arizona State University

Abstract—With an increase in the number of photo blogging platforms that allow users to post pictures, there has been a rise in the quantity of facial images of human subjects that are being uploaded every day. This has paved way to perform automatic recognition of facial features on such images. However, these images are captured in varying environmental conditions which makes detection a challenging task. In this project, we focus our attention specifically to age and gender detection from images of people made available via Adience dataset using Deep Convolutional Neural Network and Transfer Learning. We extend an existing architecture and draw comparison between the results obtained from the original system and the system achieved by incorporating our proposed modifications.

Index Terms—Convolutional Neural Network, Deep Learning, Transfer Learning

1 Introduction

Neural networks in today's world have gained a lot of importance to perform complex calculcations. Aided by the increase in the compute power at cheaper costs, neural networks have found widespread applications in the domain of computer vision. Tasks such as image segmentation, 3D image processing, object recognition and motion detection are now being performed with great accuracy.

The idea of a machine being able to recognize facial features and perform complex calculations is relatively new. It is hence not surprising to know that this field has gained a lot of attraction leading to some great strides in building systems to detect, understand and reconstruct different parts of human facial features. One such task is to accurately detect the gender of a subject and provide an age estimate of the subject. It is a fairly easy task for a human to identify the gender of a person in an image and guess the probable age group he/she may belong to. However, making a machine automatically understand the same is not trivial.

An interesting applications of above is in the domain of advertising. Knowing the age and gender of a profile on a social media platform, efficient advertising strategies can be used to target the right audience. Furthermore, since age information about people is generally kept confidential, this system can help in providing an estimate of age of subjects where the personal information is not available.

In this project, we direct our focus to achieve this task of gender detection and age group prediction by using deep Convolutonal Neural Networks and Transfer Learning.

2 Research Problem

As a part of this project we plan to extend the work of Levi, Gil, and Tal Hassner [1] to build a system to detect age and gender of human subjects from their images. We intend to replicate their existing system and attempt to improve its ability to classify human subjects by incorporating the idea of transfer learning.

3 Motivation

While training the system, the authors in [1] have used the same architecture for age and gender prediction with the difference being in the ground truth for prediction. For this, they have resorted to random weight initialization for different layers in the network.

Since the architecture is same, training the gender model helps in capturing the gender specific features from the dataset. We believe that these weights should be reflecting crucial gender based features and should be able to enhance the predictions of the age estimation system, by conveying the semantics of gender related information during age group estimation. Furthermore, since now trained system would be holding gender related feature information, we expect the new system to perform with similar level of efficiency but with reduced convolutional layers.

Hence, rather than randomly initializing the weights while training the CNN layers for age estimation system, we plan to train the same architecture for gender detection and use the weights trained for this system to seed the weights of the age estimation system.

We derive our inspiration for this potential improvement from the correlation between features of males and facial features such as beard and moustache as highlighted in [12]. These features are present in mainly males only and a system that jointly learns features for males and females needs to go undergo more complicated learning to understand the distinction. Further, the presence of such features provides sort of an informed 'prior' since males below the age of 15-18 cease to show such facial features.

4. Literature Survey

Neural networks were used earlier for gender detection, in system called SexNet [2] in early 90's. Various methods were later developed for extracting various facial features

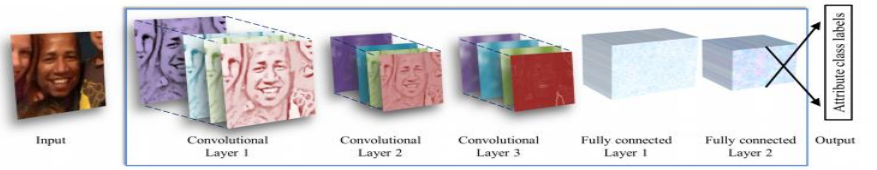

Figure 1. CNN Architecture for gender and age prediction [1]

in order to use them in classifiers. O'toole et al. in his research used the 3D structure and graylevel images of the human head to determine gender [4]. Various traditional classifiers - SVM, AdaBoost [5] etc were used on the image features and intensities. I. Ullah, M. Hussain, et. al. [6] also introduced new approach of calculating a texture descriptor (WLD) to locate the spatial features, setting up a benchmark on FERET database [8]. Caifent Shan [7] did experiments to find discriminative Local Binary Features using AdaBoost for gender classification.

Age classification has also been dealt with as a problem of facial feature extraction. One of the studies [3] calculates the wrinkle index to estimate the age from facial image. The focus on only one aspect for the age determination may not work in all cases. Various approaches presented till now do not use only local facial features. One of the methods used Gaussian Mixture Models [17] to extract distribution of facial patches. Extracting the specific features still limit the capabilities of the classifiers, since the features extraction might not be perfect due to varying facial arrangements of these features. Also, there might be more features necessary which might not be captured by the traditional feature extraction techniques.

Deep learning showed significant improvement in image classification. Deep convolutional network presented by Krizhevsky et al. [10] was able to classify the images successfully using convolutional layers followed by max-pooling and fully connected layers. They also employed dropout regularization to avoid overfitting. Ting Liu et al. [11] proposed new novel approach to estimate age using multiple face subregions. It achieved the superior performance using this multi-region convolution neural network.

5. Dataset

To carry out this research, we will be using the Adience dataset which has been made available via Face Image Project [13] at The Open University of Israel.

6. Architecture

With regards to the architecture proposed by Gil Levi [1], this network also comprises of three convolutional layers along with two fully-connected layers. The advantage of using small network is that it avoids overfitting. The classification probabilities of gender or age are obtained through softmax layer that is connected to last fully connected layer. The maximum probability class is a prediction of age group for the input image.

7. Experiments

As an evaluation task, we intend to compare the accuracy of the system with the case where gender information is already known as ground truth. This will involve training the age estimation system with random weights but on separate male and female data and train the model independently on each of them to compare performance.

8. Timeline

No.	Task	Member	Date of Completion
1	AWS setup	Laxmikant	Feb 25
2	Implementation of CNN architecture	Aarav, Laxmikant	March 15
3	Training benchmark models	Aarav	March 30
4	Training of Gender and proposed Age model	Laxmikant	April 10
5	Validation and testing of models	Aarav	April 17
6	Detailed analysis of results	Laxmikant	April 25

References

[1] Levi, Gil, and Tal Hassner. "Age and gender classification using convolutional neural networks." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2015.

[2] B. A. Golomb, D. T. Lawrence, and T. J. Sejnowski. Sexnet: A neural network identifies sex from human faces. In Proceedings of the 1990 Conference on Advances in Neural Information Processing Systems 3, NIPS-3, pages 572–577, San Francisco, CA, USA, 1990. Morgan Kaufmann Publishers Inc.

[3] Y. H. Kwon and N. da Vitoria Lobo. Age classification from facial images. In Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on, pages 762–767, Jun 1994

[4] A. J. O'toole, T. Vetter, N. F. Troje, H. H. Bulthoff, et al. Sex¨ classification is better with three-dimensional head structure than with image intensity information. Perception, 26:75–84, 1997.

[5]]S. Baluja and H. A. Rowley. Boosting sex identification performance. Int. J. Comput. Vision, 71(1):111–119, 2007

[6] I. Ullah, M. Hussain, G. Muhammad, H. Aboalsamh, G. Bebis, and A. M. Mirza. Gender recognition from face images with local wld descriptor.

[7] C. Shan. Learning local binary patterns for gender classification on real-world face images. Pattern Recognition Letters, 33(4):431–437, 2012

[8] P. J. Phillips, H. Wechsler, J. Huang, and P. J. Rauss. The feret database and evaluation procedure for face-recognition algorithms. Image and vision computing, 16(5):295–306, 1998

[9] K. Fukunaga. Introduction to statistical pattern recognition. Academic press, 1991

[10] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, pages 248–255. IEEE, 2009

[11] Liu, Ting, et al. "Age Estimation Based on Multi-Region Convolutional Neural Network." Chinese Conference on Biometric Recognition. Springer International Publishing, 2016.

[12] Gudi, Amogh. "Recognizing semantic features in faces using deep learning."arXiv preprint arXiv:1512.00743 (2015).D.S. Coming and O.G. Staadt, "Velocity-Aligned Discrete Oriented Polytopes for Dynamic Collision Detection," *IEEE Trans. Visualization and Computer Graphics*, vol. 14, no. 1, pp. 1-12, Jan/Feb 2008, doi:10.1109/TVCG.2007.70405. (IEEE Transactions)

[13] Unfiltered faces for gender and age classification, The Face Image Project, http://www.openu.ac.il/home/hassner/Adience/data.html