Age and Gender Estimation using Deep Residual Learning Network

Seok Hee Lee
Seoul National University
seokheel@snu.ac.kr

Sepidehsadat Hosseini
Seoul National University
sepid@ispl.snu.ac.kr

Hyuk Jin Kwon
Seoul National University
hjkwon@ispl.snu.ac.kr

Jaewon Moon
Korea Electronics Technology Institute
jwmoon@keti.re.kr

Hyung Il Koo
Ajou University
hikoo@ajou.ac.kr

Nam Ik Cho
Seoul National University
nicho@snu.ac.kr

***Abstract*—In this paper, we propose a deep residual learning model for age and gender estimation. Our method detects faces in input images, and then the age and gender of each face are estimated. The estimation method consists of three deep neural networks where we adopt residual learning methods. We train the model with IMDB-WIKI database [4]. However, since the database has only a small number of face images under the age of 20, we augment the set by collecting the images on the Internet. Experimental results show that the proposed model with residual learning yields improved performance.**

***Keywords*— Age estimation; Gender estimation; Deep learning; Residual learning; Machine learning; Image processing**

I. Introduction

Age and gender information is very important for advertising and marketing, where companies need to use different strategies for different groups. Among numerous methods in age and gender estimation, the work in [1] was the first method adopting deep neural networks and they showed improved performance compared with traditional feature based methods [2]. The proposed method is based on this work, however, we extend the work by using residual learning blocks from Resnet [3].

II. Proposed Model Architecture

The block diagram of our model is shown in Fig. 1. The face detectors such as [7-9] find faces in input images and the detected faces are cropped and resized to 224 x 224, and these pre-processed images are fed to the network having residual learning blocks.

We also adopt an idea in [4], i.e., our network with residual connections consists of one gender estimation network and two gender specific age estimation networks. All the three networks have the identical architecture, i.e. 50 layers with the residual connections. The output scores from the gender network are used as the weights when estimating the final age calculated by weighted sum of the outputs of the two age networks. The method in [1] divided the range of age into 8 classes and approached the problem as a classification task. However, since the aging is continuous process rather than discrete, grouping the range of age into 8 classes can be debatable. For these reasons, our model estimates age with regression and uses mean absolute error for the loss function of the model.

We train the proposed model with the images in IMDB-WIKI database [5], which is the largest public image database with age and gender labels. However, the database lacks face images under the age of 20, and so we have augmented the set by collecting over 14,000 images on the Internet.

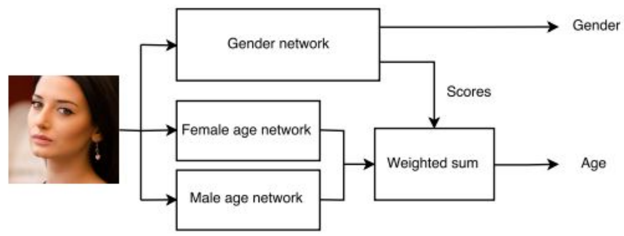

Fig 1. Architecture of age and gender estimation model

III. Experiments

We analytically performed four experiments to show the improvement resulted from the proposed learning strategies. First, we compare the residual connection in the deep network to the identical network without the connection. Second, our data augmentation is tested whether it mitigates the age imbalance problem of IMDB-WIKI database [4]. Finally, the influences of the last two strategies, the regression model and the hybrid architecture are also examined. All the experiments are blind-tested on FG-NET database [6] instead of using leave-one-person-out (LOPO) test, and all the experimental results performed in this section are shown in Table 1.

IV. Results

A. Baseline

First of all, we design the baseline model in Fig 2 to draw the comparison with the other models. The baseline model is a simple regression model with the residual learning block trained on our augmented dataset, and it outputs estimated age and gender simultaneously.

978-1-5386-2615-3/18/$31.00 ©2018 IEEE

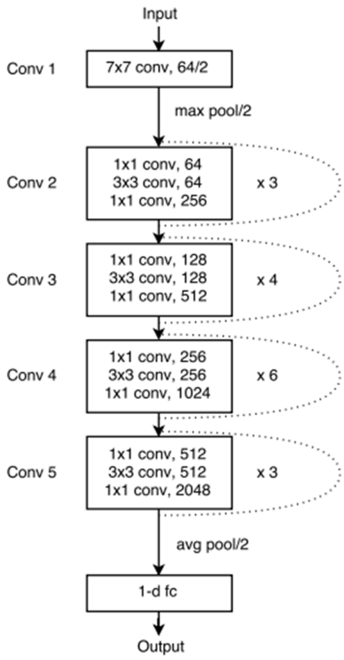

Fig 2. Structure of deep residual age estimation model

B. Residual Learning

To examine the effectiveness of residual learning, the baseline model is compared to the network modified from the baseline model by taking out all the residual connections. The result in Table 1 (5.91 vs 16.81) clearly shows that the residual learning block in the deep model enhances the performance, and enables the model to be deeper without degradation.

C. Data Augmentation

The train set without our data augmentation is used to train the model with the same architecture as the baseline model. With the data augmentation, the results in Table 1 show that the imbalance problem in the dataset is alleviated successfully.

D. Regression over Classification

Similarly, the classification model is also based on the baseline model, but its output of age is classified into 100 classes from age 0 to age 99. As shown in Table 1, the regression model yields better performance over the classification model.

E. Hybrid Architecture

As mentioned previously, our proposed hybrid network consists of three simple baseline models. This ensemble-like architecture makes the network robust by incorporating the results, and helps improve the performance.

F. Adience Benchmark

We also conduct an experiment on Adience benchmark [1]. The regression output of the baseline model is classified into 8, because the Adience set provides only age ranges as the ground truth. As shown in Table 2, our method outperforms the method in [1].

TABLE I. Experimental Result

Method	**MAE**
Baseline	5.91
w/o Residual block	16.81
w/o Augmentation	10.69
Classification	7.37
Hybrid	5.56
Rasmus [5]	8.65

TABLE II. Adience Benchmark Result

	Age Accuracy	**1-off Accuracy**	**Gender Accuracy**
Levi[1]	50.7%±5.1	84.7%±2.2	86.8%±2.2
Baseline	52.2%±6.1	92.0%±2.4	88.5%±1.4

V. CONCLUSION

In this paper, we have investigated the various strategies of improving age and gender estimation network with the residual learning. It has been shown that the regression model is better than the classification model in real age estimation. Moreover, we have shown that the residual connection helps to improve the performance in the deep network by minimizing the degradation, and our proposed model yields improved performance.

ACKNOWLEDGMENT

This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIP) (No.2017-0-00385, Development of User-Context Responsive & Interactive Digital Signage Platform) and supported by the Brain Korea 21 Plus Project in 2017.

REFERENCES

[1] Gil Levi and Tal Hassner, “Age and Gender Classification using Convolutional Neural Networks”, IEEE Workshop on Analysis and Modeling of Faces and Gestures (AMFG), at the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), Boston, June 2015

[2] Eidinger, Eran, Roee Enbar, and Tal Hassner. "Age and gender estimation of unfiltered faces," IEEE Transactions on Information Forensics and Security 9, no. 12 (2014): 2170-2179.

[3] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, pp. 770-778.

[4] X. Junliang, L. Kai, H. Weiming, Y. Chunfeng and L. Haibin. “Diagnosing deep learning models for high accuracy age estimation from a single image,” Pattern Recognition. 66. . 10.1016/j.patcog.2017.01.005.

[5] Rothe, Rasmus, Radu Timofte, and Luc Van Gool. "Deep expectation of real and apparent age from a single image without facial landmarks," International Journal of Computer Vision (2016): 1-14

[6] A. Lanitis. The FG-NET aging database, 2002. Available: www-prima.inrialpes.fr/FGnet/html/ benchmarks.html.

[7] P. Viola and M. J. Jones, “Robust real-time face detection,”International journal of computer vision, vol. 57, no. 2, pp. 137–154, 2004.

[8] H. I. Kim, S. H. Lee, and N. I. Cho, “Rotation-invariant face detection using angular projections,” Electronics letters, vol. 40, no. 12, pp. 726–727, 2004.

[9] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, “High-speed tracking with kernelized correlation filters,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 37, no. 3, pp. 583–596, 2015.