Age and Gender Classification Using Wide Convolutional Neural Network and Gabor Filter

Sepidehsadat Hosseini	Seok Hee Lee	Hyuk Jin Kwon	Hyung Il Koo	Nam Ik Cho
Seoul Nat'l Univ.	Seoul Nat'l Univ.	Seoul Nat'l Univ.	Ajou Univ.	Seoul Nat'l Univ.
sepid@ispl.snu.ac.kr	seokheel@snu.ac.kr	hjkwon@ispl.snu.ac.kr	hikoo@ajou.ac.kr	nicho@snu.ac.kr

***Abstract*—Age and gender classification has received more attention recently owing to its important role in user-friendly intelligent systems. In this paper, we propose a convolutional neural network (CNN) based architecture for joint age-gender classification, where we use the Gabor filter responses as the input. The weighting of Gabor-filter responses is learned through back-propagation in an end-to-end architecture. The architecture is trained to label the input images into 8 ranges of age and 2 types of gender. Our approach shows improved accuracy in both age and gender classification compared to the state-of-the-art methodologies. We also observe that increasing the width of neural network would increase the accuracy of the overall system.**

***Index Terms*—Convolutional neural network; Classification; Gabor filter; Deep learning**

I. Introduction

With the rapid growth of face recognition algorithm, its sub-research areas such as age/gender classfication researches are also gaining attention. As the aging process is not the same for everyone and it depends on several factors such as gender, race, living habits, etc., it is even difficult even for human to guess a person's age by looking at his picture. The same story is valid for age classifier networks.

The neural network proposed in [13], to the best of our knowledge, is the first work on using CNNs for age and gender classification. Their network is simple (having only 5 layers: 3 convolutional and 2 fully connected layers), and yet achieved a recognizable improvement in performance compared to previous non-CNN based approaches. However, it seems that the result is still far from what is needed in real world tasks. Hence, in this paper, we attempt to improve the performance by adding Gabor filter [11] to the input images in order to use some hand-crafted features in addition to the image itself as in the conventional CNN approaches. Specifically, we first extract Gabor filter responses to the input, and then add image intensities to the Gabor filter output with weights. Also, we use CNN with wider receptive field than the previous model [13], which also contributes to improve the performance.

II. Related work

Gabor filters: 2D Gabor filters are Gaussian kernel functions adjusted by a sinusoidal waves and consist of imaginary and real part [11], where the real part is as below:

$$G(x,y) = \exp\Big(-\frac{x'^2+\gamma y'^2}{2\sigma^2}\Big)\cos\big(2\pi\frac{x'}{\lambda}+\phi\big) \qquad (1)$$

where $x' = x\cos\theta + y\sin\theta \text{and} y' = -x\sin\theta + y\cos\theta$.

From the extensive experiments, we find that the optimal value for λ (wavelength of real part of Gabor filter kernel) is 2, θ (orientation of the normal to the stripes of function) is $n\pi/6$ while n varies from 0 to 5, ϕ (phase offset) is either π or zero, γ(spatial ratio) is equal to 0.3, and optimal value for σ (standard deviation) is 2. Hence in total, we use 12 filters as illustrated in Fig. 1. The Gabor filter has been used in various research areas such as detection [5], facial emotion classification [7], age estimation [4], and some image retrieval systems [1], [8].

Wide CNN: During the last decade, the algorithms based on CNN are now replacing the conventional image processing and computer vision algoirthms. The GPU developments also helped CNNs to get deeper and deeper to achieve better result. But some of recent works have shown that going deeper is not always a good idea, for example [10] proposed a wider CNN with increased size of receptive fields which obtained quite promising results in image denoising. Also Zagoruyko and Komodakis showed in [14] that having wider network can increase the image classification performance.

III. Purposed Method

A. Dataset:

In this paper all experiments are performed with Adience dataset [2]. Unlike other age/gender datasets, the Adience dataset pictures are mainly in wild covering different posses, races, lightning and etc. Adience includes both male and

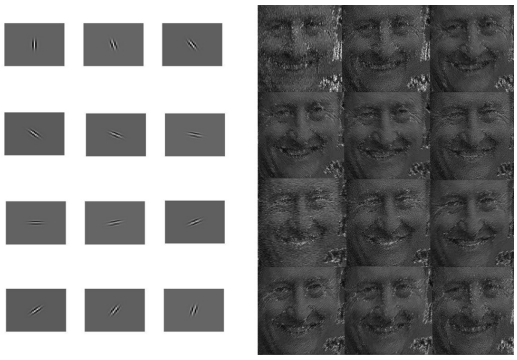

Fig. 1. A Gabor filter sets.

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

TABLE I
ADIENCE DATASET DISTRIBUTION BY AGE AND GENDER [2].

Gender	Age							
	[0:2]	[4:6]	[8:13]	[15:20]	[25:32]	[38:43]	[48:53]	[60:100]
Female	682	1234	1360	919	2589	1056	433	427
Male	745	928	934	734	2308	1294	392	442

female pictures ranging from 0 to 99, which is divided into 8 classes (see Table I).

B. Network:

The overall structure of our network is depicted in Fig 2. As can be seen in Fig 1, the Gabor filter output orientations are perfectly matched with face wrinkles so that it makes much easier for the network to focus on them. To maintain the extra information which is obtained from the Gabor filters along with the original image, the weighted sum of image and Gabor responses is used as the input to the CNN. It is noted that the optimal weights might be different for different types of images. Hence we also let the weights to be learned from the network through back propagation.

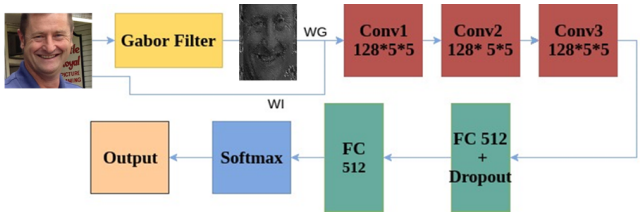

Fig. 2. A Network structure.

C. Result:

For the training and testing, the input images are preprocessed and resized to 227×227. Experiments have been done using five-fold subject exclusive protocol. The learning rate started with 0.01, and decreased exponentially. The network is implemented on Nvidia GeForce GTX 1060 6G 192 GPU and it took about 100 minutes to get the result(10000 iteration). Table II shows our result for age and gender. Our system is shown to improve up to 7%p in age-accuracy and 2%p in gender accuracy compared to the state-of-the art methods.

IV. CONCLUSION:

Researches on age and gender estimation have been divided into two main groups: one is to devise appropriate features that reflect the age and gender correctly, while the other is to use deep CNN which automatically learn the features from the massive training data. In this paper, we have proposed a method to get the benefits of both methods by enforcing the CNN to use appropriate hand-crafted features. We believe that the advantage of our scheme is to let the network to focus on useful features, which improves the performance as demonstrated in the experiments.

TABLE II
AGE/GENDER ESTIMATION RESULT ON ADIENCE DATASET.

Method	Age accuracy	Gender accuracy
LBP [12]	41.1	N.A.
Eidinger [3]	45.1	77.8
Levi [9]	50.7	86.8
DAPP [6]	54.9	N.A.
Ours	**61.3**	**88.9**

ACKNOWLEDGMENT

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

REFERENCES

[1] Y. L. Choi, N. I. Cho, and J. W. Kwak. Image retrieval method based on combination of color and texture features, 06 2006.

[2] E. Eidinger, R. Enbar, and T. Hassner. Age and gender estimation of unfiltered faces. *IEEE Transactions on Information Forensics and Security*, 9(12):2170–2179, 2014.

[3] E. Eidinger, R. Enbar, and T. Hassner. Learning pixel-distribution prior with wider convolution for image denoising. *IEEE Transactions on Information Forensics and Security*, 9(12):2170–2179, 2014.

[4] G. Guo, G. Mu, Y. Fu, and T. S. Huang. Human age estimation using bio-inspired features. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, 2009.

[5] L.-L. Huang, A. Shimizu, and H. Kobatake. Robust face detection using gabor filter features. *Pattern Recognition Letters*, 26(11):1641–1649, 2005.

[6] M. T. B. Iqbal, M. Shoyaib, B. Ryu, M. Abdullah-Al-Wadud, and O. Chae. Directional age-primitive pattern (dapp) for human age group recognition and age estimation. *IEEE Transactions on Information Forensics and Security*, 12:2505–2517, 2017.

[7] M. J. Lyons, S. Akamatsu, M. G. Kamachi, and J. Gyoba. Coding facial expressions with gabor wavelets. In *Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on*. IEEE, 1998.

[8] J. W. Kwak and N. I. Cho. Relevance feedback in content-based image retrieval system by selective region growing in the feature space. *Signal Processing: Image Communication*, 18:787–799, 2003.

[9] G. Levi and T. Hassner. Age and gender classification using convolutional neural network. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2015 IEEE Conference on*, pages 34–42. IEEE, 2015.

[10] P. Liu and R. Fang. Learning pixel-distribution prior with wider convolution for image denoising. *CoRR*, abs/1707.09135, 2017.

[11] N. Petkov. Biologically motivated computationally intensive approaches to image pattern recognition. *Future Generation Computer Systems*, 11(4–5):451,465, 1995.

[12] T. Wu and R. Chellappa. Age invariant face verification with relative craniofacial growth model. In *European Conference on Computer Vision ECCV 2012: Computer Vision, ECCV 2012*, volume 7577, pages 58–71, 2012.

[13] S. Yang, P. Luo, C. Change Loy, and X. Tang. From facial parts responses to face detection: A deep learning approach. In *IEEE International Conference on Computer Vision*, pages 3676–3684, 2015.
[14] S. Zagoruyko and N. Komodakis. Wide residual networks. In *BMVC*, 2016.