

Development Study of Deep Learning Facial Age Estimation

Puspaningtyas Sanjoyo Adi

*Department of Informatics, Faculty of Science and Technology,
Sanata Dharma University, Yogyakarta, Indonesia
Corresponding Author: puspa@usd.ac.id*

(Received 28-05-2019; Revised 29-05-2019; Accepted 31-05-2019)

Abstract

Human age estimation is one of the most challenging problem because it can be used in many applications relating to age such as age-specific movies, age-specific computer applications or website, etc. This paper will contribute to give brief information about development of age estimation researches using deep learning. We explore three recent journal papers that give significant contribution in age estimation using deep learning. From these papers, they selected classification methods and there is gradual improvement in result and also in selected loss function. The best result gives MAE (mean average error) 2.8 years and VGG-16 is the most selected CNN architecture.

Keywords: age estimation, facial analysis

1 Introduction

Human age can be estimated by facial appearance. Our faces show a special pattern in every lifetime so that our faces will have a huge difference every lifetime such as in childhood and adulthood. For the same person, the photo taken at different years indicate the aging process on their faces. The longer the interval is, the more obvious

changes there will be. Facial age estimation has potential application such as age-specific movies, age-specific products vending machine like tobacco, alcohol, and other age-specific computer applications or websites.

Estimating age from images is one of the most challenging work in facial analysis. It is hard to accurately predict human age because human facial aging is a slow and complicated process effected by many factors. With rapid advances in computer vision and pattern recognition, this problem becomes an interesting topic.

A typical pipeline of the existing methods for age estimation usually consists of two modules: age image representation and age estimation techniques [1]. Recently, deep learning schemes, especially Convolutional Neural Networks (CNNs), have been successfully employed for many tasks related to facial analysis. This paper aims to provide a brief description about some papers that have done age estimation research using CNN or deep learning. We will limit discussion to only a few paper published in journals or conferences in the last 5 years and became an important milestone of age estimating work. This paper is organized as follows: in section 2, age estimation algorithm will be explained and in section 3, we will explain about CNN architecture.

2 Age Estimation Algorithm

There has been a significant volume of research done for age estimates. This paper will focus on some papers that contributed significant development. We will explain these researches together with the estimation algorithm used. For age estimation, there are three methods that have been worked on, namely, classification, regression and ranking. In classification method, human age is assumed to be classified according to age-groups. The weakness of classification method is the sharing of important information between adjacent age groups. This is addressed by regression methods which appear to perform better. A different approach to deal with this challenge is to adopt ranking methods.

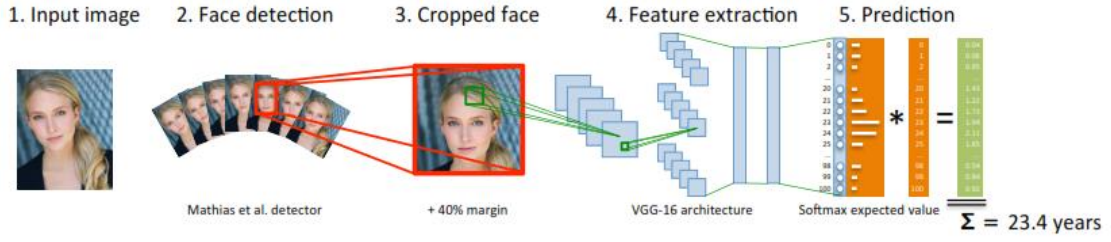


Figure 1. Pipeline of DEX method [2]

We choose Rothe's work [2] as first paper examined and the winner of the LAP 2015 challenge [3] on apparent age estimation. Age estimation done by Rothe is a classification method. They use VGG-16 [4] as base CNN architecture called DEX (Deep Expectation). Fig. 1 shows pipeline of DEX method. System will get face image and then, it will be classified using CNN into 101 classes. These classes describe possible age groups from face image samples. They train CNN for classification and at test time, they compute expecting value over the softmax-normalized output probabilities of $|Y|$ neurons.

$$E(O) = \sum_{i=1}^{|Y|} y_i o_i, \quad (1)$$

where $O = \{1, 2, \dots, |Y|\}$ is the $|Y|$ -dimensional output layer and $O_i \in O$ is the softmax-normalized output probability of neuron i . Their research result a MAE (mean average error) 3.09 years with using IMDB-WIKI [2] as training dataset and FG-NET as testing dataset [5].

The same research was also conducted by Antipov [6]. They also use VGG-16 as base CNN architecture. They did the research with 3 kind age encoding, Fig. 2,: (1) pure year classification, called 0/1 Classification Age Encoding (0/1 CAE), (2) pure regression, called Real Value Age Encoding (RVAE), (3) soft classification, called Label Distribution Age Encoding (LDAE). Each encoding has its loss function but LDAE gives the best result.

$$L_{LDAE} = -\frac{1}{N} \sum_{k=1}^N \sum_{i=1}^{100} (t_i^k \log p_i^k + (1 - t_i^k) \log(1 - p_i^k)) \quad (2)$$

where N denotes number of images in batch, k denotes number of age class, t denotes targets, p denotes to prediction. Loss function refer to Gaussian distribution.

Lost function become differentiator between Rothe [2] and Antipov [6], but they are still in classification method. Antipov research result MAE 2.84 years using FG-NET as testing dataset.

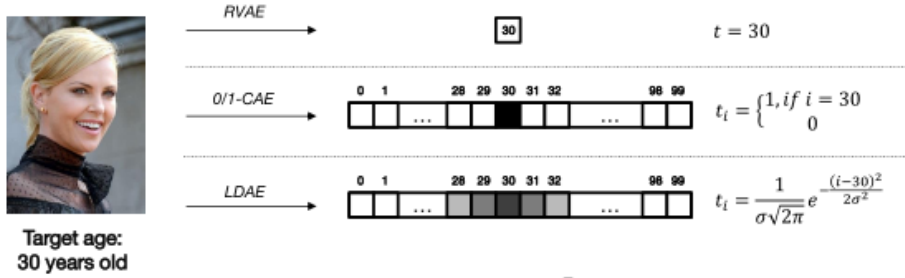


Figure 2. Example of encoding [6]. t denotes encoding result and σ is a hyper parameter of LDAE.

Two papers before showed evolution of classification methods especially in loss function development. Hu et al [7] made improvement with adding age difference estimator. This component is built to overcome limitation of ground-truth age label dataset. Making ground-truth age label dataset is a costly effort so Hu et al propose a new dataset consisted of pair face images of same person with different taken time. This new dataset is used to make age difference loss function.

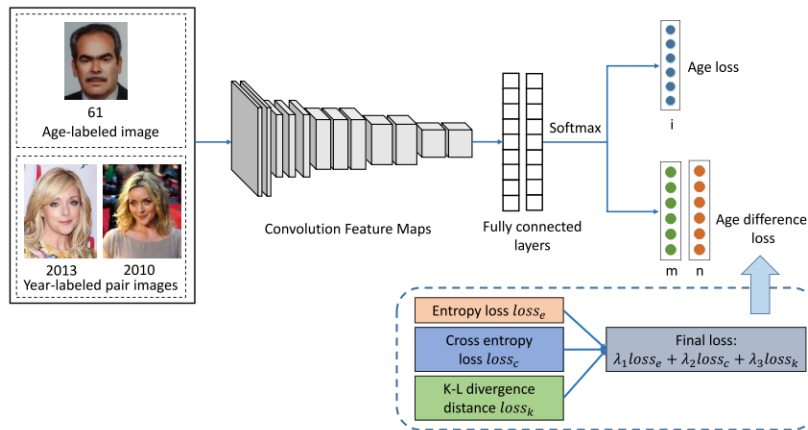


Figure 3. An overview of proposed method by Hu [7]

Fig. 3 shows an overview of proposed method by Hu. The system will give 2 outputs: age estimation and age difference. After training phase, CNN architecture will have values that will be tested with testing dataset. Initial probability distribution of age classes is set to Gaussian distribution. The age difference information with three kinds of loss functions, i.e. entropy loss, cross entropy loss and Kullback-Leibler (K-L) divergence distance. These loss functions can not only force the probability distribution of age classes to have one single peak value but also make the probability distribution locate within the correct range. This research result MAE 2.8 years using FG-NET as testing dataset.

3 Discussions

From three recently age estimation researches [2], [6], [7], we know that CNN architecture give good result in age estimation. Estimation method using classification approach gives good result. The challenge in CNN architecture is to find the best loss function which Gaussian distribution is the most choice used by researchers.

Based on three papers above, CNN architecture giving best prediction is VGG-16. This architecture is basically designed for face recognition but based on these papers, this architecture give good result for age estimation.

The other challenge is to find effective and efficient training dataset. From 3 papers, 2 papers [2], [7] contribute new dataset that is used by another similar paper for its training. IMDB-WIKI dataset [2] is not only used by [2] but it is also used by [6] for training phase. The big challenge is to make age label automatically. Further implication is hard to label age for its facial image.

4 Conclusions

In this paper, we have reviewed a few milestone papers in age estimation using deep learning. All papers result significant improvement in age estimation. Significant development component for these problem solving are loss function and dataset. From the recent researches, the task still open for improvement especially using deep learning.

In the future, this problem still gives challenging because aging process is a complex process influenced by many internal and external factor such as gene, environment, etc.

Acknowledgements

Authors wishing to acknowledge assistance or encouragement from colleagues, special work by technical staff or financial support from organizations should do so in an unnumbered. Acknowledgments section immediately following the last numbered section of the paper.

References

- [1] Y. Fu, G. Guo, and T. S. Huang, “Age synthesis and estimation via faces: a survey,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **32** (11), 1955–1976, 2010.
- [2] R. Rothe, R. Timofte, and L. V. Gool, “Deep expectation of real and apparent age from a single image without facial landmarks,” *International Journal of Computer Vision*, **126** (2-4), 144–157, 2018.
- [3] S. Escalera, J. González, X. Baró, P. Pardo, J. Fabian, M. Oliu, H. J. Escalante, I. Huerta, and I. Guyon, “ChaLearn looking at people 2015: apparent age and cultural event recognition datasets and results,” *Proceedings of the IEEE International Conference on Computer Vision*, 243–251, 2015.
- [4] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv*, 1–10, 2014.
- [5] G. Panis, A. Lanitis, N. Tsapatsoulis, and T. F. Cootes, “Overview of research on facial ageing using the FG-NET ageing database,” *IET Biometrics*, **5** (2), 37–46, 2015.
- [6] G. Antipov, M. Baccouche, S. A. Berrani, and J. L. Dugelay, “Effective training of convolutional neural networks for face-based gender and age prediction,” *Pattern Recognition*, **72**, 15–26, 2017.
- [7] Z. Hu, Y. Wen, J. Wang, M. Wang, R. Hong, and S. Yan, “Facial age estimation with age difference,” *IEEE Transactions on Image Process*, **26** (7), 3087–3097, 2017.