

Age-based facial recognition using convoluted neural network deep learning algorithm

Julius Yong Wu Jien¹, Aslina Baharum², Shaliza Hayati A. Wahab³, Nordin Saad⁴, Muhammad Omar⁵, Noorsidi Aizuddin Mat Noor⁶

^{1,2,3,4}User Experience Research Lab (UXRL), Faculty of Computing and Informatics, Universiti Malaysia Sabah, 88400 Kota Kinabalu, Sabah, Malaysia

⁵Faculty of Business Management, Universiti Teknologi MARA, Sarawak Campus, Malaysia

⁶UTM CRES, Faculty of Built Environment and Surveying, Universiti Teknologi Malaysia, Johor Bahru, Johor, Malaysia

Article Info

Article history:

Received Feb 5, 2020

Revised Apr 12, 2020

Accepted Apr 27, 2020

Keywords:

Age detection

Convolutional neural network

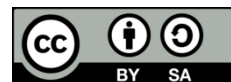
Image processing

Prediction

ABSTRACT

Face recognition is the use of biometric innovations that can see or validate a person by seeing and investigating designs depending on the shape of the individual. Face recognition is used largely for the purpose of well-being, despite the fact that passion for different areas of use is growing. Overall, face recognition innovations are worth considering because they have the potential for broad legal jurisdiction and different business applications. It is widely used in many spaces. How it works is a product of facial recognition processing facial geometry. The hole between the ear and the good way from the front to the jaw are the main variables. This code distinguishes the highlight of the face that is important for your facial separation and creates your facial expression. Therefore, this study gives an overview of age detection using a different combination of machine learning and image processing methods on the image dataset.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Aslina Baharum,
UXRL, Faculty of Computing and Informatics,
Universiti Malaysia Sabah,
88400 Kota Kinabalu, Sabah, Malaysia.
Email: aslina@ums.edu.my

1. INTRODUCTION

Facial recognition (FR) is utilized for the most part for wellbeing purposes, despite the fact that enthusiasm for different territories of utilization is developing [1]. In all actuality, FR innovation has gotten noteworthy consideration as it has the potential for a wide scope of law authorization and different business applications [2]. It has been generally utilized in numerous spaces, for example, ATM, social insurance framework, driving permit framework, train reservation framework, observing assistance, and identification verification. How it functions is the product for facial acknowledgment peruses the face's geometry. The hole between the ears and the good ways from the front to the jaw are key variables [3]. The code distinguishes facial highlights that are crucial to your face separation and produce your facial mark. Because of deep learning techniques, there have been critical advances in FR [4]. In the early stages, exploring the interest, for the most part, focused on FR with a deep system of significant light or picture faces. Stephen [5] given a brief review of the techniques of deep learning and face-to-face learning and compares some of the basic neural formulas based on common convolution neural networks (CNNs). The deep networks used in FR, such as deep belief network (DBN), convolutional neural network (CNN, or ConvNet), autoencoder (AE), and others are analyzed for architecture [6]. Mandal [7] assessed a significant measure of profound learning strategies

for FR. Sepas-Moghaddam [6] studied of FR arrangements dependent on another, all the more incorporating and more extravagant staggered scientific classification. Learned-Miller [8] looked at a variety of surprising inventive strategies in the Labeled Faces in the Wild (LFW) database.

2. RESEARCH METHOD

Figure 1 shows the flow of methodology. This flow consists of phase 1 until phase 3 as below:

- Phase 1
In this phase, data is acquired. In the data preparation stage, samples of facial images are acquired and then undergo pre-processing to enhance the quality of images.
- Phase 2
In this phase, segmented images is then undergone feature extraction process here. These extracted features are then used in the training process.
- Phase 3
The final phase is the prediction and evaluation stage where each built model is used to predict the input image. The accuracy of each model will be calculated and evaluated.

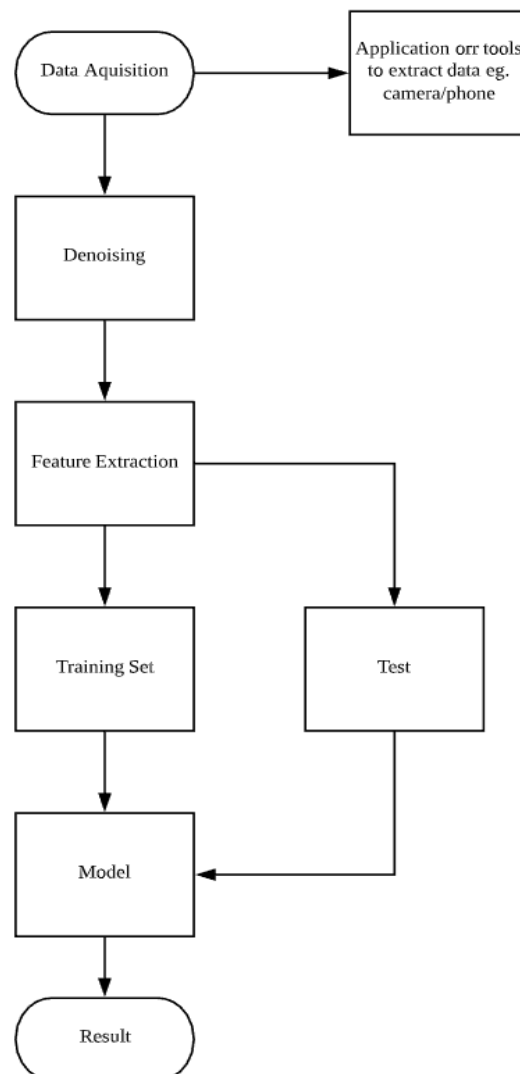


Figure 1. Flow of methodology

3. RESULTS AND DISCUSSION

3.1. Data acquisition

Three databases were utilized in the examination: the BERC database, the PAL maturing database, and the FG-Net maturing database [9]. The BERC database contains the face pictures of 390 people the age extend 3 to 81 years of age. The facial pictures were gotten utilizing a computerized camera at a high resolution of 1600×1200 pixels. The FG-NET developing database is used in this investigation, which contains 1002 face pictures from 82 one of a kind subjects, and each subject has 6–18 face pictures named with ground truth ages. The ages are flowed in a wide range from 0 to 69. The age transport in either the number of pictures or the amount of subjects is especially disproportionate [10-11].

3.2. Image denoising

Numerically, picture commotion is portrayed as a multi-dimensional stochastic process [12]. Therefore, picture commotion can be numerically portrayed by scientific insights, for example through likelihood thickness circulation work [13]. The presentation of a scientific model can accomplish better denoising.

Noise Model

(1) Rayleigh noise

$$P(z) = \begin{cases} \frac{2}{b}(z-a)e^{-(z-a)\frac{2}{b}} & z \geq a \\ 0 & z < a \end{cases} \quad (1)$$

Condition 1 is the likelihood thickness appropriation capacity of Rayleigh commotion. When the gray value is more prominent than or equivalent to the likelihood thickness bend grades to one side and the fundamental region on the correct side is bigger, that is, the dim estimation of commotion focuses is more distributed in the right side of the central axis $a + \sqrt{(b/2)}$

(2) Gaussian noise

$$P(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z-u}{2\sigma^2}} \quad (2)$$

Condition 2 is the likelihood thickness appropriation capacity of the Gaussian commotion. The gray value of the clamor is around the focal grayscale, which is generally dark, and the high contrast commotion circulation is less [14]. Gaussian commotion is likewise called ordinary noise. Its likelihood thickness complies with typical circulation. It is a generally utilized commotion model. Gaussian noise is normally brought about by awful lighting or high temperature during procurement [15].

3.3. Feature extraction

Feature extraction is a strategy of dimensionality decline by which a fundamental game plan of rough data is reduced to logically sensible social affairs for taking care of [16]. Nature of these tremendous enlightening assortments is innumerable elements that require a huge amount of enrolling advantages for the process [17]. Feature extraction is the name for methods that select and additionally merge factors into features, effectively decreasing the proportion of data that must be taken care of, while still absolutely and thoroughly depicting the primary instructive assortment [18-19].

In this study, the highlights can be separated by utilizing CNN. Utilizing locale-based CNN finding key positions, making a sliding window on the picture, and moving the sliding window along the picture to get the potential objective zone, CNN is utilized to remove the standard highlights of the objective region, that is, to get the yield of fixed measurements as per convolution, pooling and different tasks. At that point, the yield vectors of the subsequent stage are grouped (classifiers should be prepared by their highlights); and the facial age is anticipated utilizing the CNN model.

3.4. Training set

3.4.1. Convolutional neural network structure

The classical network structure of convolutional neural system appears in Figure 2. The convolution neural system contains a few "convolution layers" and "examining layers" to process the information signal. At that point, the mapping between the info signal and the yield result is acknowledged in the full association layer. Every convolution extricates the highlights of the information signal through a convolution activity of a convolution filter [20-22]. Examining layer is likewise called the "assembly" layer. Its capacity is to utilize the standard of neighborhood relationship to down example, which lessens information (decreases calculation) and holds valuable data in the system [23].

After the picture goes through all convolution layers and testing layers, the element mapping is normally changed over into highlight vector yield by full association activity, as it were, the full association layer is framed.

The element mapping can be associated persistently for commonly, and the last full association layer is the yield layer. By and by, the yield layer is frequently utilized as the element of picture extraction, and these highlights are generally utilized for relapse characterization preparing (for preparing SVM or Soft Max) [24-25].

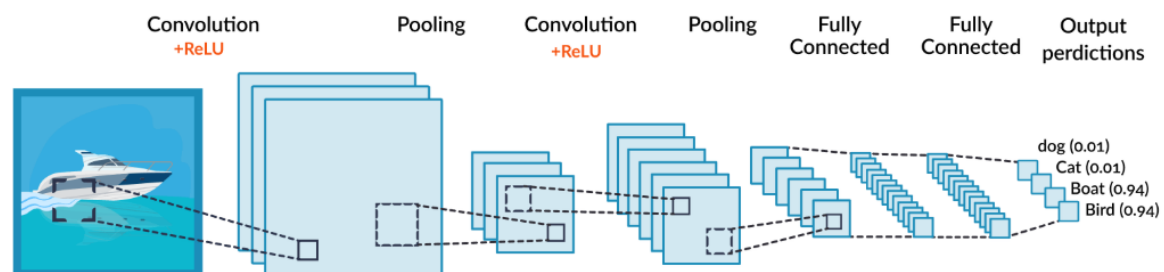


Figure 2. The classical structure of CNN

4. CONCLUSION

It tends to be anticipated that the programmed facial age evaluation strategy dependent on CNN has preferred execution over that dependent on artificial features of SVM. The explanation might be that the CNN can procure progressively inexhaustible and significant highlights in facial pictures by learning, while the falsely planned highlights can just cover some fixed and single highlights in the picture, and the demeanor of highlights is not rich. Overall, this study aims to give an overview of the improvement of the accuracy in detecting the age of facial recognition. Besides, facial images datasets that were proposed to use in this paper for further study. Facial images from a different region in the world such as CXR from Africa or from the western country can be applied too. Furthermore, the total number of datasets use as training and validation suggested to increase as it might improve overall performance.

REFERENCES

- [1] Ioannis A. Kakadiaris, George Toderici, Georgios Evangelopoulos, Georgios Passalis, Theoharis. "3D-2D face recognition with pose and illumination normalization." *Computer Vision and Image Understanding*. 154:137-151, 2017.
- [2] Jiang Jing Lv, Xiao Hu Shao, Jia Shui Huang, Xiang Dong Zhou, Xi Zhou. "Data augmentation for face recognition." *Neurocomputing*. 230: 184-196, 2017.
- [3] Vaibhav Jain, Dinesh Patel. "A GPU based implementation of Robust Face detection System." *Procedia Computer Science*. 87:156- 163, 2016.
- [4] Andrew J.Logan, Gael E.Gordon, Gunter Loffler. "Contributions of individual face features to face discrimination", *Vision Research*. 137: 29-39, 2017.
- [5] Lane, T.M., Niru, M., Aran, N., Surya, G., Stephen, A.B. et al. "Deep Learning Models of the Retinal Response to Natural Scenes," in *Proceedings of the 30th Conference on Neural Information Processing Systems (NIPS 2016)*, Barcelona, Spain.
- [6] A. Sepas-Moghaddam, F.M. Pereira, P.L. Correia. "Face recognition: a novel multi-level taxonomy based survey," *ET Biometrics*. 9(2): 58-67, 2020.
- [7] S.Song, V.Chandrasekhar, B. Mandal et al. "Multimodal Multi-Stream Deep Learning for Egocentric Activity Recognition," in *Proceedings of the 29th IEEE Conference on Computer Vision and Pattern Recognition Workshops, CVPRW 2016*, pp.378–385, 2016.

- [8] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller et al. "Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments," in "Technical Report 07-49," University of Massachusetts, Amherst. 2007.
- [9] G. Cybenko, Approximation by superpositions of a sigmoidal function, *Mathematics of control, signals and systems*. 2(4): 303-314, 1989.
- [10] O. Delalleau, Y. Bengio, "Shallow vs. deep sum-product networks," in: *Advances in Neural Information Processing Systems*, pp. 666–674, 2011.
- [11] Z. Lu, H. Pu, F. Wang, Z. Hu, L. Wang, "The expressive power of neural networks: A view from the width," in: *Advances in Neural Information Processing Systems*, pp. 6231-6239, 2017.
- [12] M. Usman (Akram), Aasia Khanum, "Retinal images: blood vessel segmentation by threshold probing," in: *IEEE Symposium on Industrial Electronics & Applications (ISIEA)*, 3–6th, October, 2010, Penang, Malaysia, pp. 493-497, 2010.
- [13] H. N. Mhaskar, T. Poggio, "Deep vs. shallow networks: An approximation theory perspective," *Analysis and Applications*, 14(06): 829-848, 2016.
- [14] H. Mhaskar, Q. Liao, T. Poggio, "Learning functions: when is deep better than shallow," arXiv preprint arXiv:1603.00988.
- [15] G. F. Montufar, R. Pascanu, K. Cho, Y. Bengio, "On the number of linear regions of deep neural networks," in: *Advances in neural information processing systems*, pp. 2924–2932, 2014.
- [16] Gasah, M., Baharum, A., Zain, N.H.M., Halamy, S., Hanapi, R., Noor, N.A.M. "Evaluation of Positive Emotion in Children Mobile Learning Application," *Bulletin of Electrical Engineering and Informatics* 9(2): 818-826, 2020.
- [17] Yin Bi, Mingsong Lv, Yangjie Wei, Nan Guan, Wang Yi, "Multi-feature fusion for thermal face recognition," *Volume 77*: 366-374, 2016.
- [18] Ayan Seal, Debotosh Bhattacharjee, Mita Nasipuri. "Human face recognition using random forest based fusion of à-trous wavelet transform coefficients from thermal and visible images," *AEU-International Journal of Electronics and Communications*, 70(8):1041-1049, 2016.
- [19] Li-Fang Zhou ,Yue-WeiDu, Wei-Sheng Li, Jian-XunMi, XiaoLuan. "Pose-robust face recognition with Huffman-LBP enhanced by Divide-and-Rule strategy", *Pattern Recognition*, 78: 43-55, 2018.
- [20] Thibault Napoleon, Ayman Alfalou. "Pose invariant face recognition: 3D model from single photo," *Optics and Lasers in Engineering*, 89: 150-161, 2017.
- [21] Michele Nappi, Stefano Ricciardi, Massimo Tistarelli. "Deceiving faces: When plastic surgery challenges face recognition," *Image and Vision Computing*, 54:71-82, 2016.
- [22] Yan Liang, Yun Zhang, Xian-Xian Zeng. "Pose-invariant 3D face recognition using half face" *Signal Processing: Image Communication*, 57: 84-90, 2017.
- [23] Chenfei Xu, Qihe Liu, Mao Ye. "Age invariant face recognition and retrieval by coupled auto-encoder Networks," *Neurocomputing*, Volume 222: 62-71, 2017.
- [24] Sung Eun Choi, Jaeik Jo, Sanghak Lee, Heeseung Choi, Ig-Jae Kim, Jaihie Kim. "Age face simulation using aging functions on global and local features with residual images," *Expert Systems with Application*, Volume 80: 107-125, 2017.
- [25] Q. Xu, M. Zhang, Z. Gu, G. Pan, "Overfitting remedy by sparsifying regularization on fully-connected layers of CNNs," *Neurocomputing*, 328: 69-74, 2019.