Adaptive Hyperparameter for Face Recognition

Thanh-Tam Nguyen, Son-Thai LE, Van-Thuy LE

Abstract*: *One of the widely used prominent biometric techniques for identity authentication is Face Recognition. It plays an essential role in many areas, such as daily life, public security, finance, the military, and the smart school. The facial recognition task is identifying or verifying the identity of a person base on their face. The first step is face detection, which detects and locates human faces in images and videos. The face match process then finds an identity of the detected face. In recent years there have been many face recognition systems improving the performance based on deep learning models. Deep learning learns representations of the face based on multiple processing layers with multiple levels of feature extraction. This approach has made sufficient improvement in face recognition since 2014, launched by the breakthroughs of DeepFace and DeepID. However, finding a way to choose the best hyperparameters remains an open question. In this paper, we introduce a method for adaptive hyperparameters selection to improve recognition accuracy. The proposed method achieves improvements on three datasets.

Keywords*: *Face Recognition, Deep Learning, Hyperparameter.

I. INTRODUCTION

Face recognition has attained great attention over the last three decades. The main reasons for this trend are the large variety of commercial and legal requests and the availability of the technologies. After Schroff et al. proposed the FaceNet, a unified embedding for face recognition and clustering [1], Deep Convolutional Neural Networks (CNNs) have made face recognition improved significantly. In previous works [2, 3, 1], fixed hyperparameters are used in face recognition systems. They often choose values of hyperparameters so as they help to distinguish different classes of the testing data. However, the optimal hyperparameters depend on a specific data set. Therefore it is challenging to find optimal hyperparameters for all cases. Besides, input face datasets would frequently change, raising the need for hyperparameter values tuning. We propose a method of adaptive hyperparameters estimation to improve face recognition accuracy. The adaptive hyperparameters help recognize the face regarding the database. Based on the proposed method, we build a deep CNN system with a database. In this system, we use deep CNN to extract the feature vectors of facial images. These feature vectors and hyperparameters are stored in the database. To evaluate the proposed method, we use an evaluation protocol with the same configuration as the proposed system. The experimental results show that our method outperforms the traditional approaches.

The following are the main contributions of this paper:

- We propose a method of adaptive hyperparameters estimation to improve the accuracy of face recognition based on deep CNN.
- We compare our proposed method with the traditional approach on three datasets.

The remaining of this paper is organized as follows: Section II briefly reviews the related works in face recognition. Section III describes the proposed method. In Section IV, we show and analyze experimental results.

The paper ends up with the conclusion and future work in section V.

II. RELATED WORKS

This section presents a brief review of approaches for face recognition. A detail survey on human action recognition can be found in [4]. Methods for face recognition on images often use appearance information. These methods can be divided into four categories: Non-linear, Linear, Deep learning based methods, Sparsity based methods. Among these categories, learning based methods have been successfully used in face recognition systems since 2014 due to the increase in processing power and the compilation of large databases containing labeled samples [5,6,7,8,1].

These works try to achieve a better model by adjusting the objective functions and revising the network structures. The goal of this effort is to use the deep learning model to extract discriminative facial features.

For example, Schroff et al. [1] proposed Face Net, which used GoogleNet-24 as the network architecture with the triplet loss function. The Deep Face method [5] based on DNN employs a nine-layer neural network with over 120 million connection weights. VGG Face [3] CNN descriptors are computed using a CNN implementation based on the VGG-Very-Deep-16 CNN architecture.

However, how to choose the optimal hyperparameter for face recognition in real-world applications is still an open question. In this paper, we introduce to use of adaptive hyperparameter instead of fixed hyperparameters. This approach is more suitable for practical systems.

Revised Manuscript Received on January 15, 2021.
* Correspondence Author

Thanh-Tam NGUYEN*, Faculty of Multimedia, The Posts and Telecommunications Institute of Technologies, Hanoi, Vietnam. Email: ntttam@ptit.edu.vn

Son-Thai LE, Department of Multimedia Communications, The School of Information and Communication Technology - Thai Nguyen University, Thai Nguyen, Vietnam. Email: lsthai@ictu.edu.vn

Van-Thuy LE, the School of Foreign Languages - Thai Nguyen University, Thai Nguyen, Vietnam. Email: levanthuy.sfl@tnu.edu.vn

III. PROPOSED METHOD

A. Overview framework

The framework of the proposed method is shown in Figure 1. There are two phases in our system: Database collection and face recognition.

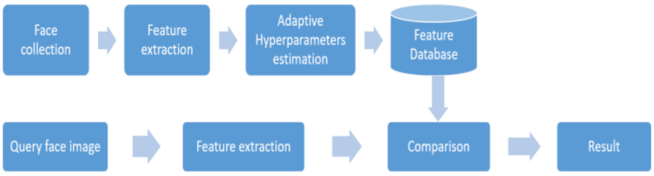

Fig. 1. **Framework of the proposed method.**

For each input face image in the first phase, a feature vector is extracted using a deep CNN. A hyperparameter that is along with each face is also assigned accordingly. The feature vectors and hyperparameters are stored in a database.

In the face recognition phase, given a query image, a feature vector is extracted. The similarity scores between the query and all feature vectors in the database are then computed. We determine the query image's identity based on the similarity scores and regarding hyperparameter as a threshold to give the result.

B. Feature extraction using Deep CNN

Given an image, we first use the Multi-task Cascaded Convolutional Networks (MTCNN) [9] to detect the face. To detect faces, firstly, MTCNN produces candidate windows through a simple CNN. It then uses an involved CNN to reject non-faces windows. Finally, a strong CNN is utilized to give face location, Figure 2.

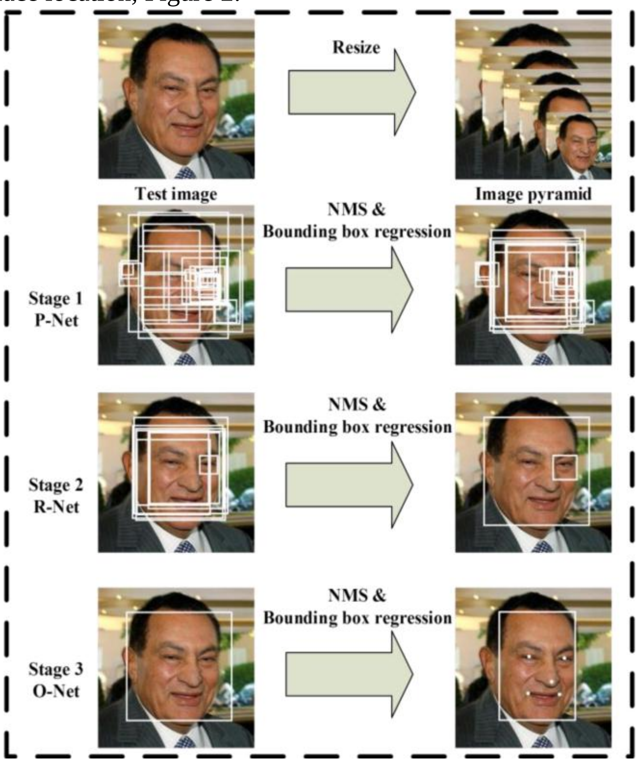

Fig. 2. **Pipeline of MTCNN face detection framework [9]**

After archiving facial landmarks positions, a pre-trained model FaceNet [1] is employed to extract high-quality features from faces. FaceNet uses a deep CNN trained to directly optimize mapping from face images to a compact Euclidean space. The training use triplets of roughly aligned matching or non-matching face patches generated using a new online triplet mining method, Figure 3.

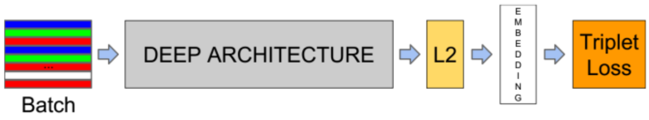

Fig. 3. **Model structure of FaceNet [1]**

C. Similarity score and adaptive hyperparameters

Because feature vectors of faces computed using FaceNet are vectors in a compact Euclidean space, we use the dot product between two feature vectors to compute face similarity. The similarity between two faces are defined as (1):

$$SL(v_1, v_2) = v_1.v_2 \qquad (1)$$

In the database collection phase, for each face image f belong to class p, we compute a feature vector v using the FaceNet model and a hyperparameter h. The hyperparameter h is computed as the maximum value among similarities between v and facial feature vectors in the database not belonging to class c (Eq.2):

$$h = \underset{c_i \neq c}{Max}(SL(v, v_i)) \qquad (2)$$

The hyperparameters will be updated once occur a new face image in the dataset.

D. Face recognition

Given a query facial image f_q, the corresponding feature vector v_q is computed using the FaceNet model. We then compute the similarity scores between v_q and all the feature vectors stored in the database. The index i_m of the most similar class is found as Eq.3:

$$i_m = \arg\max_i (SL(v_q, v_i)) \qquad (3)$$

After that, the query facial feature vector v_q is defined to belong to c_{i_m} class if the highest similarity is greater than or equal to the associated hyperparameter of the most similar class. Otherwise, the query facial image f_q does not belong to any known class.

IV. EXPERIMENTS

A. The dataset and evaluation protocol

We evaluate the effectiveness of the proposed method on three datasets: Labeled Faces in the Wild (LFW) [10], Adience [11] and Color FERET [12].

LFW is a public benchmark for face verification which was created by the University of Massachusetts, Amherst. It has 13,233 images of 5,749 people, Figure 4.

Fig. 4. **Some facial image in LFW dataset.**

The Adience dataset contains the challenges of real-world imaging conditions. It has 19,339 images with variations in appearance, noise, pose, lighting, Figure 5.

Fig. 5. **Example images from Adience dataset.**

The FERET is a popular dataset for face recognition evaluation created at George Mason University and the Army Research Laboratory. It has 14,126 images of 1199 people. The images were taken in a different situation. The testing dataset contains 2,413 face images captured from 856 people, Figure 6.

Fig. 6. **Examples of Color FERET dataset.**

To compare with the fixed hyperparameter approach, we conduct the fixed-hyperparameter experiments following the way in [FaceNet] that selected optimal fixed-hyperparameter. We use 10-fold cross-validation all three datasets.

B. Experimental results

Overall, shown in Table I, our proposed method outperforms traditional approach on all three datasets.

Table- I: Comparison with the traditional approach in accuracy (%)

Dataset	Proposed method (adaptive hyperparameter)	Traditional approach (fixed hyperparameter)
LFW	**76.5%**	54.0%
Adience	**84.5%**	80.5%
Color FERET	**84.0%**	81.1%

As we can see in Table I, the adaptive hyperparameter approach's advantages are shown the most clearly on the dataset LFW. On this dataset, the proposed method's recognition rate is 22.5% higher than that of fixed hyperparameter on the dataset LFW. With the dataset Adience, the traditional approach contains an accuracy of 80.5%, while the adaptive hyperparameter approach obtains 84.5% (4% higher). The proposed method is also slightly better than the fixed-hyperparameter on the Color FERET.

V. CONCLUSION

This paper proposed a method for hyperparameter selection for face recognition using deep learning feature vectors. The proposed method is an adaptive approach for hyperparameter estimation. The experimental results show that the proposed method is better than the traditional approach. The proposed technique of adaptive hyperparameter selection can help us realize face recognition systems in practical applications.

ACKNOWLEDGMENT

This research is funded by the School of Foreign language - Thai Nguyen University under grant number NNCS 2044.

REFERENCES

1. Schroff, Florian, Dmitry Kalenichenko, and James Philbin. "Facenet: A unified embedding for face recognition and clustering." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 815-823. 2015.
2. Chen, Sheng, Yang Liu, Xiang Gao, and Zhen Han. "Mobilefacenets: Efficient cnns for accurate real-time face verification on mobile devices." In *Chinese Conference on Biometric Recognition*, pp. 428-438. Springer, Cham, 2018.
3. Parkhi, Omkar M., Andrea Vedaldi, and Andrew Zisserman. "Deep face recognition." (2015).
4. Adjabi, Insaf, Abdeldjalil Ouahabi, Amir Benzaoui, and Abdelmalik Taleb-Ahmed. "Past, present, and future of face recognition: a review." *Electronics* 9, no. 8 (2020): 1188.

5. Y. Taigman, M. Yang, M. Ranzato and L. Wolf, "DeepFace: Closing the Gap to Human-Level Performance in Face Verification," *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, 2014, pp. 1701-1708, doi: 10.1109/CVPR.2014.220.
6. Sun, Yi, Yuheng Chen, Xiaogang Wang, and Xiaoou Tang. "Deep learning face representation by joint identification-verification." *Advances in neural information processing systems* 27 (2014): 1988-1996.
7. Y. Sun, X. Wang and X. Tang, "Deep Learning Face Representation from Predicting 10,000 Classes," *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, 2014, pp. 1891-1898, doi: 10.1109/CVPR.2014.244.
8. Y. Taigman, M. Yang, M. Ranzato and L. Wolf, "Web-scale training for face identification," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, 2015, pp. 2746-2754, doi: 10.1109/CVPR.2015.7298891.
9. Zhang, Kaipeng, Zhanpeng Zhang, Zhifeng Li, and Yu Qiao. "Joint face detection and alignment using multitask cascaded convolutional networks." *IEEE Signal Processing Letters* 23, no. 10 (2016): 1499-1503.
10. Huang, Gary B., Marwan Mattar, Tamara Berg, and Eric Learned-Miller. "Labeled faces in the wild: A database forstudying face recognition in unconstrained environments." 2008.
11. 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.
12. Phillips, P. Jonathon, Hyeonjoon Moon, Syed A. Rizvi, and Patrick J. Rauss. "The FERET evaluation methodology for face-recognition algorithms." *IEEE Transactions on pattern analysis and machine intelligence* 22, no. 10 (2000): 1090-1104.

AUTHORS PROFILE

Thanh-Tam NGUYEN has received her degree of bachelor in Information Technology from Thai Nguyen University in 2004 and degree of Master in Information System from the University of Engineering and Technology, Ha Noi National University, in 2017. She has 11 years of teaching experience. She is currently a lecturer at the Faculty of Multimedia - The Posts and Telecommunications Institute of Technologies (PTIT), Hanoi, Vietnam. Her research interests include Machine Learning, Multimedia Application Development.

Son-Thai LE is a lecturer at the Department of Multimedia Communications, The School of Information and Communication Technology, Thai Nguyen University. His research interest is Multimedia.

Van-Thuy LE works at the School of Foreign Languages - Thai Nguyen University. He graduated from Information Technology in 2008 and Master of Computer Science in 2014 at the University. His research interest is Face Recognition.