

Improving Deep Learning for Face Verification Using Color Histogram Equalization Data Augmentation

Yi-Quan Li, Daw-Tung Lin*, Zhe-Wei Yeh

Department of Computer Science and Information Engineering
National Taipei University

151, University Rd., San-Shia, New Taipei City, Taiwan

Email: superjerry1984@gmail.com; dalton@mail.ntpu.edu.tw; junkinsane20@gmail.com

Abstract - This paper proposes a new method of improving face verification learning using color histogram equalization by incorporating the results of deep convolutional neural networks (CNNs). The entire process of face verification using deep learning and color histogram equalization is described in detail. This research uses advanced deep learning methods for face verification tasks. Because the CNN achieves the best results for larger datasets, the main challenge is to increase the smaller dataset enhancements and validate in environments different from the training datasets. This paper presents a new training enhancement method. When the face-image datasets were small, we could use our enhanced method to expand the dataset and improve the accuracy of face verification to adopt different environments. Consequently, the accuracy reached 99.7996%, i.e., approximately 7% higher than the result trained on the original dataset.

Keywords: Deep Learning; Face Verification; Color Histogram Equalization; VGG-Face.

1. Introduction

Owing to the advent of modern technology, an individual's safety is of utmost importance. Problems such as campus terrorist attacks, commercial espionage, or impersonation of households are still emerging. To correctly and accurately recognize the identity of a company, home, school, etc., the research and application of human face; regardless of face recognition or verification, has become an important research topic. The face contains the primary information for the recognition of a person's identity, in addition to their gait and body contours; therefore, the development of face recognition technology is receiving a significant amount of more attention. This technology can be applied to various scenarios, for example, in determining whether a criminal appears in surveillance video and confirm whether the suspect is the same person in different scenarios, and in the field of forensics demanding strict technical requirements. The primary steps of face recognition technology typically include face detection [1], feature positioning of face key points, and face alignment [2]. The geometric topological relationship between face feature points such as eyebrows, pupils, noses, and mouths is used to represent face information in the geometric feature method.

Face verification is an application based on face recognition, and its primary goal is to determine whether two face images belong to the same person. The similarity between an image obtained from a camera and an image in a gallery is calculated to confirm the identity of the image. Because a face collection can be obtained easily, face verification has been widely used in security systems and e-commerce systems. However, many limitations still exist in the application of face verification. This method relies on images, and many factors affect the results of face verification, for example, lighting changes, skin color, and aging effects [4,5]. However, since 2012, owing to the re-emergence of deep convolutional neural networks, the performance of sampled face feature points has been improved significantly. Further, because of the combination of deep learning methods and a large amount of training data, face recognition applications have demonstrated significant progresses and breakthroughs.

Currently, deep convolution neural networks are applied to face verification and face recognition. Training a face verification system begins with selecting a dataset of face images, detecting faces in the image, cropping and aligning those faces, and subsequently training the deep network on the face images of the cropped and possibly aligned faces. Each step of the process involves many design issues and choices. These problems have received particular attention from researchers.

For deep learning, the number of training samples is important; therefore, the method to increase the limited sample size is emphasized in the current study.

2. Related Work

With the progress in face verification, many computerized monitoring and attendance systems are now available for companies, homes, and campuses. These systems use different core technologies. In [6], Arsenovic et al. proposed a face recognition system to verify work attendance. Wang *et al.* [7] proposed a classroom attendance rate evaluation system for colleges based on deep learning. These methods utilized deep learning networks for face verification and have been applied successfully in company and school attendance systems, and were inspired by their preprocessing of datasets. The method to enhance the dataset to adopt different environments using face verification was studied.

Face alignment is an important preprocessing step for face analysis systems [8]. Particularly, the performance of face recognition systems can be improved using aligned face images [9]. Facial feature point positioning, that is, face alignment is important in face recognition; the positioning accuracy and positioning speed directly affect the face recognition system. Face alignment in a real scene becomes extremely difficult owing to factors such as different poses, expression, illumination, and partial occlusion in face images. Zhang *et al.* [10] presented a face alignment method based on the SURF of the scale-invariant feature transform that can quickly detect and characterize the key points of a face image.

Human face detection has been challenging in image processing and pattern recognition. Cuimei *et al.* [11] proposed a new human face detection algorithm using the primitive Haar cascade algorithm combined with three additional weak classifiers and demonstrated excellent results. Shu *et al.* [12] proposed a new face detection method that demonstrated excellent high accuracy. We introduce a novel facial region network to detect faces in unconstrained conditions, in which faces with large pose variations and severe occlusion can be detected correctly.

Facial recognition poses many difficulties because two images of the same person can vary considerably according to time, pose, facial expression, and illumination conditions. Azimifar *et al.* [13] presented an approach based on sparse discriminative multimanifold embedding that used feature extraction rather than intensity and normalization for preprocessing to reduce the effects of uncontrolled condition such as illumination.

Extracting facial feature is considered as the most important step in face recognition systems. Ammour *et al.* [14] proposed a hybrid feature (global and local) based on the Gabor filter bank and double coding local binary pattern. It performed excellently in extracting facial features.

The remainder of this paper is organized as follows. Section 3 describes the proposed method, Section 4 presents the experimental results and discussion, and Section 5 draws the conclusions.

3. Methodology

This section details the proposed method for improving deep learning for face verification using color histogram equalization. The process is divided into several important stages, including face detection, face alignment, sharpness justification, obtaining training datasets and data augmentation methods, training VGG-Face [15], and finally using the trained VGG-Face model for testing and reach the output results. Figure 1 depicts the block diagram of the proposed face verification system.

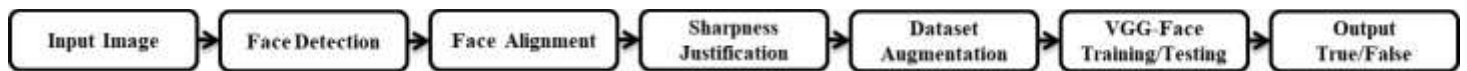


Fig. 1: Block diagram of the proposed face verification system.

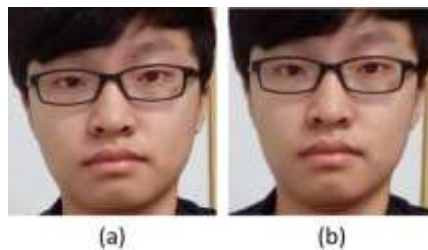


Fig. 2: Face detection and alignment: (a) original image; (b) aligned face image.

3.1. Face Detection and Alignment

Face detection is necessary and prioritized in applications such as face verification. In face detection and alignment pre-process, the eyes must be placed in the same horizontal position. Therefore, we adopted the Dlib C++ [16] algorithm function to obtain the position coordinates of 68 anchor points of the facial features, and applied posture correction and scaling. The rotation angle was calculated using eye coordinates for posture correction. The face alignment results are compared as shown in Fig. 2. After face alignment was performed, the eyes were corrected to the horizontal position.

3.2. Image Sharpness Justification

Because blur and motion blur are typical problems in face verification [17], it is necessary to justify whether the image is blurred when collecting the datasets. In this study, the variance calculation algorithm was used to detect fuzziness. The grayscale difference value of the image with sharp focus is larger than the grayscale difference value of the image with blurred focus; therefore, its variance is relatively large. Therefore, image clarity can be measured by calculating the variance of the image grayscale data. The variance value is greater, and the sharpness is better. The method was used to filter the collected data to prevent excessively blurred data from hindering deep learning and training, hence affected the accuracy of face verification. The image clarity test results are shown in Fig. 3.



Fig. 3: Image sharpness detection: (a) clear image detection with variance 40.8; (b) blurred image detection with variance 30.62.

3.3. Data Augmentation with Color Histogram Equalization

To increase the number of training samples, typical data augmentation technology was adopted including mirror image conversion (as shown in Fig. 4), or object simulation with accessories [18]. In this study, we proposed a data augmentation scheme using color histogram equalization. The proposed equalization step ensures that face verification does not affect the test result due to environmental lighting changes between the training samples and the test samples.

In the proposed color histogram equalization method, we convert images from red, green, blue to hue, saturation, and intensity value. Then, the histogram equalization is performed for the intensity value. Figure 5 shows an example of conversion image. The resultant conversion training images with better color histogram distribution help to verify positive samples with uneven foreground or background brightness distribution during testing and adopt environmental differences.

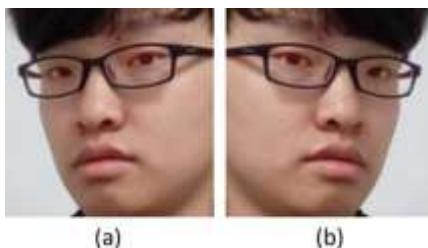


Fig. 4: Image mirror conversion: (a) original image; (b) mirror image.

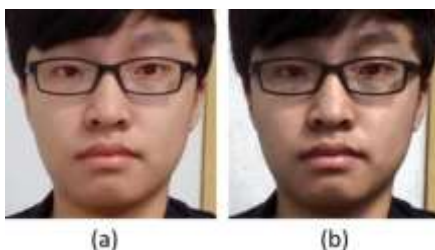


Fig. 5: Color histogram equalization: (a) original image, (b) conversion result.

3.4. Deep Learning Architecture

This study utilized VGG-Face as training/testing network architecture [15]. VGG-Face applies the VGG-16 convolutional network architecture [19]. It is trained on a dataset of approximately 22.6 million images of 2622 people.

Because this study focuses on personal face verification, the output will be either true or false. Therefore, we modified the output of the last layer of convolutional network architecture and changed the number of output nodes to two. Table 1 presents the modified network configuration and parameters setup for the last layer. Both the image data adopted from the VGG-Face training dataset and our collected test dataset must be scaled to the same size 224×224 , and training was performed using the VGG-Face model.

Table 1: VGG-Face last layer network configuration.

Type Name	Support	Filt dim	Num filts	stride	Pad
Convolutional fc8	1	4096	2	1	0

4. Experimental Results

4.1. Dataset Collection and Preparation

The proposed method was tested in a collaboration company with a dataset collected from 15 people's photographs including two students and 13 volunteer employees from the company. The dataset contains photographs captured at different times and different angular positions as shown in Fig. 6(a). Totally, 1850 face images were collected. Furthermore, 1850 face images were divided into training set and test set containing 1422 and 428 face images, respectively. The details number of training samples and test samples for each person is listed in Table 2.

Table 2: The details number of training samples and test samples for each person collected in the experimental dataset.

Person ID	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Total
Training Samples	64	85	71	23	169	119	65	42	113	148	73	172	90	156	32	1422
Test Samples	7	7	6	10	53	29	34	28	46	45	15	54	27	50	17	428

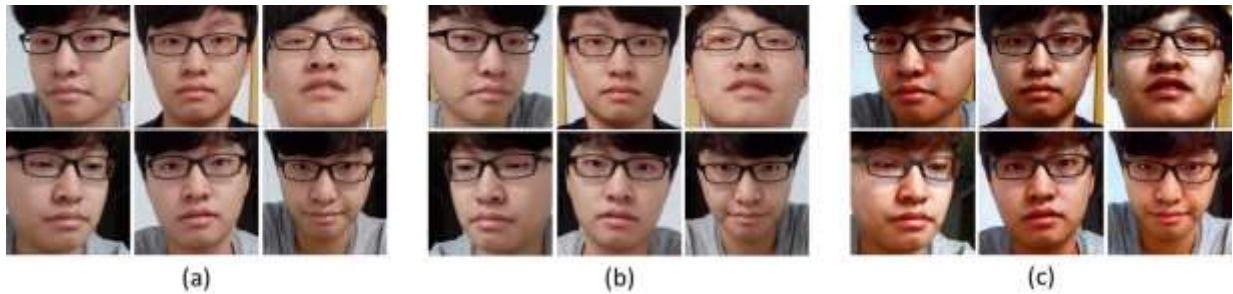


Fig. 6: Various types of training data: (a) original photographs captured with different angles and at different times; (b) aligned face images; (c) images after performing color histogram equalization.

Next, according to the proposed data augmentation method, alignment, mirror conversion and color histogram equalization were performed and created an enhanced dataset of total 5688 face images. Figures 6(b) and 6(c) demonstrates the sample images of face alignment and color histogram equalization, respectively. In addition, to reduce false alarm, 1071 negative samples were acquired and further separated into 1000 negative training examples and 71 negative test samples. Finally, the VGG-Face [15] was used to train the collected dataset and evaluated on the test dataset.

4.2. Training Scenarios and Test Results

To facilitate and adjust the deep learning performance, various training scenarios have been engaged. Initially, we applied the collect training dataset (1422 face images) and performed face verification on the test dataset with 499 face images (428 positive samples and 71 negative samples). The experimental results are summarized in Table 3. As can be observed from Table 3, the overall face verification accuracy is 92.5852%, however, the test results showed that a false negative can occur easily, thus deteriorating the accuracy. To reduce the false negatives, we further acquired 1000 negative sample and included them into the training dataset. Then, the total number of face images in the training dataset extended to 2422. The experimental results after joining face photographs of people not on the training list as a negative sample are presented in Table 4. The experimental results indicate that the accuracy in face verification increased significantly after increasing the negative sample data, from 92.5852% to 97.5952%, i.e., an increase of 5.01%. Both false negatives and false positives were reduced.

Table 3: Experimental results of face verification using 1422 positive training samples.

	True	False
Positive Test Samples	395	33
Negative Test Samples	67	4

Table 4: Experimental results of face verification using training set containing 1422 positive samples and 1000 negative samples.

	True	False
Positive Test Samples	416	12
Negative Test Samples	71	0

Moreover, different lighting conditions generated from capturing images will also deteriorate the accuracy. Next, we increase the training samples by using the color histogram equalization method and extend the dataset size to 2844 positive samples and 1000 negative samples. Table 5 shows the experimental results after using the color histogram equalization method. The face verification accuracy was improved significantly and reached 99.1984%. Compared with the face verification results of any enhanced training samples, the accuracy rate increased by 6.6132% from 92.5852%. Compared with the face verification results after training with the negative sample data, the accuracy rate increased by 1.6032% from 97.5952%. Our experiment indicates that the color histogram equalization method can effectively improve the accuracy of face verification. This is because the number of training samples are increased and the color are equalized to become less-random changed.

Table 5: Experimental results of adding color histogram equalization positive training samples.

	True	False
Positive Test Samples	426	2
Negative Test Samples	69	2

Finally, we combined the aforementioned data augmentation strategies including mirror images, color histogram equalization, and negative training samples to extend the dataset to 6688 face images and trained on the VGG-Face network. The experimental results of deep learning face verification on the test dataset are summarized in Table 6. Compared with the results of previous experiment scenarios, the accuracy is again improved significantly. After using our method, the face verification accuracy reached 99.7996%. Among the 499 test data, only one error occurred, which is better than the other experimental results. Table 7 shows the comparison of test results of various training strategies presented in this study.

Table 6: Experimental results using the proposed data augmentation method.

	True	False
Positive Test Samples	427	1
Negative Test Samples	71	0

Table 7: Comparison of experimental results obtained from various type of datasets.

Dataset Combination	Accuracy
Original positive samples	92.5852%
Positive samples + Negative samples	97.5952%
Positive samples + Color Histogram Equalization + Negative samples	99.1984%
Positive samples + Mirror + Color Histogram Equalization + Negative samples	99.7996%

5. Conclusion

We herein proposed a training data augmentation method for improving deep learning for face verification using the color histogram equalization method. We described the entire process of face verification using color histogram equalization and advances in deep learning. When the face-image datasets were small, we could use our enhanced method to expand the dataset and improve the accuracy of face verification to adopt different environments. Consequently, the accuracy reached 99.7996%, i.e., approximately 7% higher than the result trained on the original dataset. The results indicated that using color histogram equalization to expand the dataset could reduce the face verification error caused by lighting condition changes. Although face images were aligned for the photographs at different angles in this study, differences still remained in the features. This resulted in verification errors. Hence, the accuracy of face verification could be improved using face data generated from different angles. We intend to improve the overall accuracy of face verification in the future.

Acknowledgements

This work is partially supported by the Ministry of Science and Technology under grants MOST 105-2221-E-305 -006 -MY3, 106-2622-E-305-003-CC3, and 107-2622-E-305-003-CC3. We also would like to thank the partial support by the National Taipei University under grant 107E19101.

References

- [1] Y. Zhou, D. Liu, T. Huang, "Survey of face detection on low-quality Images," in *Proceedings of the IEEE Conference on Automatic Face & Gesture Recognition*, Xi'an, China, May, 2018, pp. 769-773.
- [2] V. Kazemi, J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, June, 2014, pp. 1867-1874.
- [3] J. Zeng, J. Zeng, X. Qiu, "Deep learning based forensic face verification in videos," in *Proceedings of the IEEE Conference on Progress in Informatics and Computing*, Nanjing, China, Dec, 2017, pp. 77-80.
- [4] S. Bijarnia, P. Singh, "Age invariant face recognition using minimal geometrical facial features," In *Advanced Computing and Communication Technologies*, Springer, Singapore, June, 2016, pp. 71-77.
- [5] N. Ramanathan, R. Chellappa, "Face verification across age progression," *IEEE Transactions on Image Processing*, Oct, 2006, vol. 15, pp. 3349-3361.
- [6] S. Florian, D. Kalenichenko, J. Philbin, "Facenet: A unified embedding for face recognition and clustering," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 815-823.
- [7] R. Fu, D. Wan, D. Li, Z. Luo, "University classroom attendance based on deep learning," in *Proceedings of the IEEE Conference on Intelligent Computation Technology and Automation*, Changsha, China, Oct, 2017, pp. 128-131.
- [8] P. Viola, M. J. Jones, "Robust real-time face detection," *IJCV*, vol. 57, no. 2, pp. 137-154, 2004.
- [9] M. Yalçın, H. S. Yavuz, "Comparison of matrix decomposition and SIFT descriptor based methods for face alignment," In *Signal Processing and Communication Application Conference*, Zonguldak, Turkey, May, 2016, pp. 1841-1844.

- [10] K. Cui, H. Cai, Y. Zhang, H. Chen, "A face alignment method based on SURF features," in *International Congress on Image and Signal Processing, BioMedical Engineering and Informatics*, Shanghai, China, Oct, 2017, pp. 1-6.
- [11] L. Cuimei, Q. Zhiliang, J. Nan, W. Jianhua, "Human face detection algorithm via Haar cascade classifier combined with three additional classifiers," in *Proceedings of the IEEE Conference on Electronic Measurement & Instruments*, Yangzhou, China, January, 2018, pp. 483-487.
- [12] H. Shu, D. Chen, Y. Li, S. Wang, "A highly accurate facial region network for unconstrained face detection," in *Proceedings of the IEEE Conference on Image Processing*, Beijing, China, Sept, 2017, pp. 665-669.
- [13] Z. Azimifar, A. Nazemi, F. Shahali, "Single sample face identification utilizing sparse discriminative multi manifold embedding," in *Proceedings of the IEEE Conference on Artificial Intelligence and Signal Processing Conference*, Shiraz, Iran, Oct, 2017, pp. 129-133.
- [14] B. Ammour, T. Bouden, L. Boubchir, S. Biad, "Face identification using local and global features," in *Proceedings of the IEEE Conference on Telecommunications and Signal Processing*, Barcelona, Spain, July, 2017, pp. 784-788.
- [15] O. M. Parkhi, A. Vedaldi, A. Zisserman, "Deep face recognition," In *Proceedings of the British Machine Vision Conference*, vol. 1, no. 3, 2015, p. 6.
- [16] V. Kazemi, J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, June, 2014, pp. 1867-1874.
- [17] K. Knežević, E. Mandić, R. Petrović, B. Stojanović, "Blur and Motion Blur Influence on Face Recognition Performance," in *Symposium on Neural Networks and Applications*, Belgrade, Serbia, Nov, 2018, pp. 1-5.
- [18] M. Arsenovic, S. Sladojevic, A. Anderla, D. Stefanovic, "FaceTime—Deep learning based face recognition attendance system," in *Proceedings of the IEEE International Symposium on Intelligent Systems and Informatics*, Subotica, Serbia, Sept, 2017, pp. 53-58.
- [19] N. Crosswhite, J. Byrne, C. Stauffer, O. Parkhi, Q. Cao, A. Zisserman, "Template adaptation for face verification and identification," in *Proceedings of the IEEE International Conference on Automatic Face & Gesture Recognition*, Washington, DC, USA, June, 2017, pp. 1-8.