Histogram Equalized DeepPCA with ELM Classification for Expressive Face Recognition

Kanokmon Rujirakul
Department of Business Computer, Faculty of Management Science, Nakhon Ratchasima Rajabhat University,
Nakhon Ratchasima, Thailand
kanokmon.r@nrru.ac.th

Chakchai So-In
Applied Network Technology (ANT),
Department of Computer Science, Faculty of Science,
Khon Kaen University, Khon Kaen, Thailand
chakso@kku.ac.th

***Abstract*—In this paper, we propose a novel approach for expressive face recognition with deep learning networks. There are three main components: 1) Histogram Equalization (HE), 2) Principal Component Analysis (PCA), and 3) Extreme Learning Machine (ELM). The first module is used for pre-processing to adjust a histogram curve of input images. Then, a deep learning concept with PCA is applied as a feature extraction, and finally, ELM is used as a baseline classification scheme. Two well-known public databases, LFW and KDEF, are selected to evaluate the proposed method with a comparative performance evaluation against a traditional PCA with several classification methods and state of the art facial recognition technique, i.e., PCAnet, where the experimental results demonstrate our superior performance.**

***Keywords*—Deep Learning; Principal Component Analysis; Face Recognition; Expression; Extreme Learning Machine**

I. INTRODUCTION

Nowadays, a facial recognition system is widely used in both research and commercial which is applied to many aspects, e.g., human identifications, robotics, crowd surveillances, and criminal forensics. However, face recognition with variations such as makeup change, expression, and pose has still challenged for real-world use. Thus, many face recognition architectures have been proposed to deal with these issues, one of which is toward deep learning networks.

Deep learning has recently been researched [1] in many areas including image classification and speech recognition including face recognition. In general, deep learning extracts a hierarchical representation of data; one of its key ingredients for image classification is convolutional architecture [2]. Thus, here, as shown in Fig. 1, we propose a novel and simplified method using deep learning as a face recognition architecture for various expression face images. Principal Component Analysis (PCA) filter is also used for data-adapting convolution filter bank. The filter is well-known selected (DeepPCA).

Note that before feature extraction, Histogram Equalization (HE) is applied as a pre-processing used to adjust a histogram curve in each facial image; together with deep learning and PCA, we call HE-DeepPCA. During a classification stage, Extreme Learning Machine (ELM) is adopted. These combination can achieve high accuracy with a comparative time-complexity trade-off (HE-DeepPCA-ELM).

This paper is organized as follows. In section II, we briefly provide an overview of facial recognition background. Then, section III presents a comparative survey of related works. In section IV, we present our approach, HE-DeepPCA-ELM, in particular, an expressive facial image. After that, we discuss the performance of our proposal comparatively. Finally, the conclusions and future work are drawn in section VI.

II. PCA FACE RECOGNITION SYSTEMS

In general, there are two stages in a face recognition system, i.e., training and testing. There are three main steps [3], namely, pre-processing, feature extraction, and image identification in each of these two stages.

In the first stage (training), facial images are acquired as the training images from various sources, such as a video frame, a camera, or photo scanning, and then fed into a pre-processing state. Various pre-processing approaches have been proposed for normalizing the facial images, e.g., illumination normalization, background removal, gray-scale conversion, and HE. The features of the normalized facial images have then extracted for collecting the feature vectors in the training sets. These features are readily prepared for testing (next stage).

Similar to the training, the testing facial image will be pre-processed to derive the normalized (testing) image. Then, the testing feature vectors are extracted. Finally, the classification method (identification) will be performed by comparing the testing image vector against the vectors derived from the training process.

For decades, numerous well-known approaches have been proposed to extract the feature in a face recognition system, such as principal component analysis (PCA), independent component analysis (ICA), and linear discriminant analysis (LDA) [4]. However, PCA is typically employed as a baseline and is widely used for facial recognition [3].

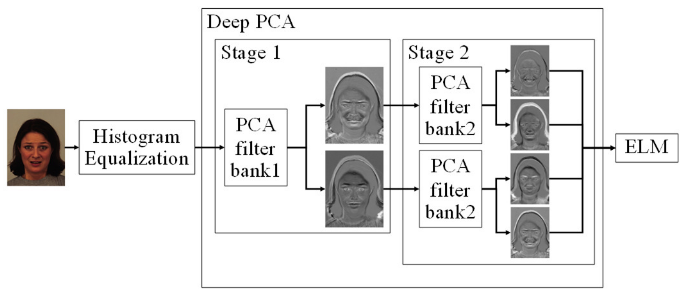

Fig. 1. An overall of the proposed architecture.

It should be noted that PCA is commonly used due to several advantages, e.g., requiring less number of features but with resulting in high accuracy even with a small dataset [5]. With PCA, large-dimensional data can be reduced while retaining only the major features [6]. In general, PCA face recognition systems consist of two main steps, as described below.

1) Feature extraction: this step consists of four components: (1) normalizing all training data by subtracting the mean from each datum; (2) computing a covariance matrix from the normalized data so as to obtain corresponding eigenvalues and eigenvectors; (3) sorting the eigenvectors by their eigenvalues in descending order and eliminating the eigenvectors that have an eigenvalue of zero; and (4) projecting the training image by multiplication of the normalized data with the sorted eigenvectors.

2) Image identification: the testing image will be normalized by subtracting the mean of the training images from the testing image, and then projecting the normalized testing image into the eigenspace via multiplication with the eigenvectors from the training stage. Finally, the projection of training and testing will be compared to determine the minimum difference.

Although PCA expresses many advantages, in particular, with high recognition rate, it is generally applied only to a neutral face, i.e., a face without distinctive noise, such as multi-view face positions (even with occlusion). Note that human expression can also oblique the accuracy in PCA [7].

III. Related Work

There have been several face recognition proposals employing PCA for different purposes and obtained numerous distinctive features, e.g., less memory requirement, simple computation complexity, and high recognition rate [8-10].

Consider multi-expression of human faces. In 2011, H. Mohammadzade *et al.* [11] presented the approach using bivariate empirical mode decomposition integrating with LDA to increase the recognition rate when the dataset contains several expressions. A year later, C. Tripathi and K. P. Singh [12] enhanced the performance while applying multi-dimensional discriminant analysis over multi-dimensional principal components with SVM classifier. However, there is a key limitation of input images, i.e., non-occlusion, which in turn may limit the precision gain.

Similarly, in the same year, P. Marasamy and S. Sumathi [13] comprised Wavelet Transform in LDA facial expression recognition and reported the performance enhancement. However, the key limitation still remains with a small dataset. With a large dataset, PCA yields higher recognition rate but with high computational complexity. Thus, in 2014, K. Rujirakul *et al.* [14] introduced a fast algorithm, expectation maximization (EM), over PCA. This method also applied HE to mitigate the effect of emotion or expression of human faces. The authors reported the performance improvement of their proposed method.

In recent year, deep learning has been widely used in pattern recognition area to increase the accuracy and speed up the computational time. For example, M. A. R. *et al.* [15] used a gated Markov Random Field as the front-end of a DBNs. Here, it can learn better features for facial expressions. It is also robust to occluded images. However, this model is still computationally expensive.

In 2012, G. B. Huang et al. [16] used Convolutional Deep Belief Network (CDBN) to learn hierarchical representations. The concept is to develop a local convolutional restricted Boltzmann machines to exploit a global structure. The networks are trained on the gray images and the Local Binary Patterns (LBP) images to get 59 dimensional uniform LBP features. The test is carried on LFW database, and determined a state-of-the-art accuracy.

In 2015, T. H. Chan et al. [17] proposed a convolutional neural network (CNN) without active function and pooling layers. Instead of BP algorithm, they investigated both PCA and LDA in order to learn the bases and treated the bases as filters in CNN, called PCANet and LDANet, respectively. The experiments illustrated that a two-layer PCANet was superior to the state-of-the-art feature extraction for some image classification tasks.

Consider classification stage. Several approaches have been investigated, such as Euclidian Distance (ED), Manhattan Distance, Mahalanobis Distance, Nearest Neighbor, and SVM [18]. Recently, ELM has attracted more attention in image processing with reliable performance and fast learning speed [19]. For example, W. Zong and G. B. Huang [20] studied multi-label face recognition performance using ELM classifier. Discussions and comparisons on four benchmark face databases showed that ELM based classifier was able to achieve a comparable recognition rate to SVM.

In summation, PCA can achieve high recognition rate but with high computational time complexity. Here, our proposal is to investigate PCA with deep learning approach for a purpose of computational complexity reduction with significantly accuracy for multi-expressive face recognition systems. To mitigate the effect of expression, HE and deep learning are applied including the use of ELM classifier to increase the recognition rate and speed up the recognition system.

IV. Proposed Method

The method proposed in this paper follows as a pre-processing of the face image, then, feature extraction. The extracted features are used to construct classifiers for each subject with ELM classification. Thus, person identification will be the final outcome.

A. *Pre-processing*

With regards to our previous experiment [14], HE was a key identifier to enhance the recognition precision. HE was applied to adjust the intensity and histogram curve in each facial image. Thus, this research also adopted HE as one of the pre-processers of the proposed method.

The details are as follows: (1) create a counting table of the color image in range of 0-255, (2) compute cumulate color values as well as its least frequencies, and (3) apply equation below to achieve the equalization.

$$h[v] = round\left(\frac{cd[v]-cd_{min}}{(M\times N)-cd_{min}}\right)\times(L-1) \quad (1)$$

Here, *cd* is a cumulative frequency of the size of images in terms of width (*M*) and height (*N*). *v* denotes a color value; *L* is an entire color space, i.e., from 0 to 255.

B. Deep PCA Feature Extraction

To reduce the impact of the poor-quality regions in the face image caused by variations, e.g., pose, illumination, and facial expressions, we apply deep learning for improving the traditional PCA feature extraction. The features extracted from the first level of DeepPCA are then used as the input of the second level of DeepPCA, as shown in Fig. 2.

The PCA features are learned from the *N* input images - $\{I_i\}_{i=1}^{N}$. In the first level, the $p \times p$ patch are collected from each image. Each patch is then subtracted by its mean; we then compute a covariance matrix from the normalized data to determine an eigendecomposition in order to obtain the eigenvalues and eigenvectors from the covariance matrix. The leading eigenvectors with non-zero eigenvalues are selected. The second level follows the same process as the first level. The feature of output layer (f_i) is retrieved after these two levels.

C. ELM Classification

The proposed method, DeepPCA, is applied for feature extraction along with ELM classification. The ELM proposed by G. B. Huang *et al.* [19] applies a single hidden layer feed-forward network (SLFN). There is no need to adjust the weight of hidden nodes in ELM. Thus, ELM can result in fast training.

Fig. 3 shows an overall architecture of ELM. The *N* input neurons are given in the format of (x_i, t_i) such that *i* = 1, 2, 3, …, *N*. The input $x_i = (x_{i1}, x_{i2}, \dots, x_{in})$ is fed into the network with a given target $t_i = (t_{i1}, t_{i2}, \dots, t_{in})^T$. To convert the non-linearity, the output weight (β) and the bias (b) are used. Both are derived from the following equation used for training.

$$\beta = H^{\dagger}T \tag{2}$$

The output weight β can be obtained by resolving the least-square solution, as stated in the following.

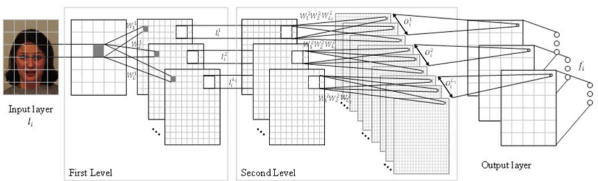

Fig. 2. Detailed architecture of the proposed two-stage DeepPCA.

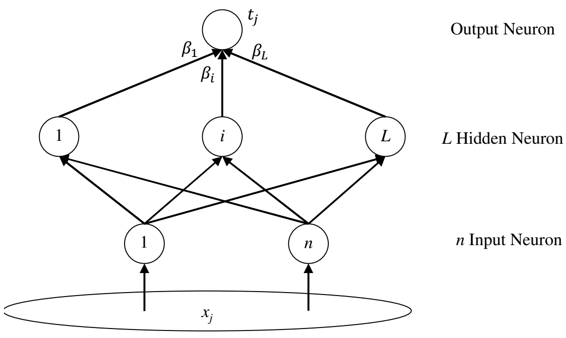

Fig. 3. Extreme Learning Machine Architecture.

$$min \sum_{i=1}^{N} \|\beta_i \cdot h_i - t_i\| \tag{3}$$

where, T denotes the target and H is the hidden node function, i.e., $\{h_{ij}\} \mid (i = 1, 2, \dots, N$ and $j = 1, 2, \dots, K)$ that is derived from $h_{ij} = g(w_j x + b)$. $H^{\dagger}$ is Moore-Penrose matrix. For example, the hidden nodes' input weights are $w_j = (w_{j1}, w_{j2}, \dots, w_{jn})^T$ with a bias of b_i.

$g()$ denotes an activation function, i.e., Linear Kernel, Radial Basis Function (RBF) Kernel, Polynomial Kernel, and Wavelet Kernel [21]. In testing stage, the unknown input *x* will be fed into $h_{ij} = g(w_j x + b)$ with the defined weight (w) before applying the activation function and then employing the reverse equation (4) to calculate the predicted target.

$$T = H\beta \tag{4}$$

where $H(w_1, w_2, \dots, w_K, b_1, b_2, \dots, b_K, \dots, x_1, x_2, \dots, x_N) =$

$$\begin{bmatrix} g(w_1 \cdot x_1 + b_1) & \cdots & g(w_K \cdot x_1 + b_K) \\ \vdots & \ddots & \vdots \\ g(w_1 \cdot x_N + b_1) & \cdots & g(w_K \cdot x_N + b_K) \end{bmatrix} N \times K \tag{5}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_K^T \end{bmatrix} \text{ and } T = \begin{bmatrix} T_1^T \\ \vdots \\ T_K^T \end{bmatrix}. \tag{6}$$

From related works, the Kernel-ELM classifiers present a better performance than traditional ELM, Constrained ELM (CELM), and SVM, in both classification accuracy and training speed [12]. Therefore, we have carried out the simulation using the Kernel-ELM classifiers (ELM-Kernel) from Nanyang Technological University, Singapore [21].

V. PERFORMANCE EVALUATION

In this section, we performed the evaluation process in order to assure the performance of our efficient mechanism for Histogram Equalization Deep PCA ELM (HE-DeepPCA-ELM) face recognition.

A. Experimental Setup

Our testbed is a standard configuration on personal computer Windows 7 Ultimate operating systems (64bits): CPUIntel® Core(TM) i-3770K 8-Cores 3.50GHz (8MB L3 Cache), 8192×2 MB DDR3-SDAM, and 500 GB 5400 RPM disk. The experimental testbed was implemented in Matlab R2016a programming environment and compared with traditional method including state of art method - PCAnet [17]. Two public face datasets, i.e., KDEF and LFW, including Middle East, Asian, and Caucasoid, were selected [22].

In these datasets, a set of grayscale images includes multi-expression, afraid, angry, disgusted, happy, sad, and surprised as well as normal images were collected with size of 77×57 pixels (KDEF 140 images) and 64×64 pixels (LFW 926 images). We performed 5-fold cross validation using the 5 subsets of each dataset for all experiments.

Our proposed method, HE-DeepPCA-ELM was evaluated against the existing facial recognition methods, i.e., traditional PCA-ED, HE-PCA-ED, DeepPCA-ED, HE-DeepPCA-ED, PCAnet [17], HE-PCAnet, and DeepPCA-ELM. With DeepPCA and PCAnet, the parameters were set as following:

- **PCAnet:** NumStages = 2, PatchSize = [7 7], NumFilters = [8 8], HistBlockSize = [7 7], BlkOverLapRatio = 0.5;

- **ELM-Kernel:** Regularization coefficient = 1, Kernel type = Linear Kernel, Kernel Para = 100.

Here, ED, LIBSVM [13], and ELM kernel (Linear) were applied to compare the performance in classification stage. The performances were measured by averaging the 5-fold cross validation with also 5 trials in order to evaluate our model.

B. Experimental Results

From Figs. 4 and 5, the result shows that the improving of face recognition in that using HE and deep learning approaches over PCA incurs the performance enhancement.

Our proposed method, HE-DeepPCA-ELM, is superior, i.e., in average of 83% in precision (54, 60, 79, 81, 80, and 81) with 208 seconds in computational time. (21, 21, 436, 410, 287, and 285) for PCA-ED, HE-PCA-ED, DeepPCA-ED, HE-DeepPCA-ED, PCAnet, and HE-PCAnet, respectively.

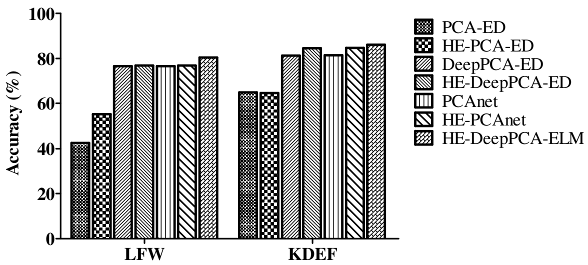

Fig. 4. Percentage of recognition rate of LFW and KDEF databases.

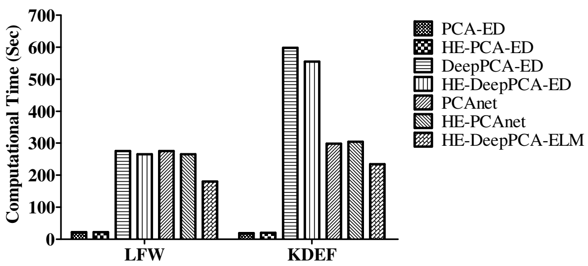

Fig. 5. Computational Time of LFW and KDEF databases.

VI. CONCLUSION

This paper proposed DeepPCA for feature extraction along with ELM as classification in a face recognition system. The experimental conclusion indicated that HE-DeepPCA-ELM can provide a satisfactory classification rate on well-known publicly available face image databases - KDEF and LFW.

The experimental results demonstrated that the proposed method can achieve high performance in terms of speed and accuracy. Note that HE-DeepPCA-ELM outperforms the other approaches: PCA-ED, HE-PCA-ED, DeepPCA-ED, HE-DeepPCA-ED, PCAnet, HE-PCAnet, and DeepPCA-ELM, by the factors of 1.5, 1.4, 1.1, 1.0, 1.1, and 1.0, respectively. In addition, the speed-up of HE-DeepPCA-ELM over the other approaches is in order of 0.1, 0.1, 2.1, 2.0, 1.4, and 1.4, respectively

In sum, the proposed method shows its superiority over other methods not only in the accuracy but also the robustness to expression. However, the utilization of deep learning is still at the beginning; therefore, a lot of works are required to be well-investigated, such as verify a persons' face on low-resolution images and transform expressions to emotions.

REFERENCES

[1] Z.P. Fu and Y.N. Zhang, "Survey of Deep Learning in face recognition," in *Proc. of the ICOT 2014*, pp. 5-8, 2014.

[2] S. Balaban, "Deep learning and face recognition: the state of the art," in *Proc. of the SPIE 9457*, vol. 9457, 2015.

[3] A.K. Bansal and P. Chawla, "Performance Evaluation of Face Recognition Using PCA and N-PCA," *Int. J. Comput. Appl.*, vol. 76, no. 8, pp. 14-20, 2013.

[4] R. Jafri and H.R. Arabnia, "Survey of Face Recognition Techniques," *J. Info. Process Syst.*, vol. 5, no. 2, pp. 41-68, 2009.

[5] K. Delac, M. Grgic, and S. Grgic, "Independent Comparative Study of PCA, ICA, and LDA on the FERET Data Set," *Int. J. Imag. Syst. Tech.*, vol. 15, no. 5, pp. 252-260, 2005.

[6] L. Min, L. Bo, and W. Bin, "Comparison of Face Recognition Based on PCA and 2DPCA," *J. Adv. Info. Sci Service Sci*, vol. 5, pp. 545-553, 2013.

[7] G. Givens, J.R. Beveridge, B.A. Draper, P. Grother, and P.J. Phillips, "How features of the human face affect recognition: a statistical comparison of three face recognition algorithms," in *Proc. of the CVPR 2004*, vol. 2, pp. 381-388, 2004.

[8] N. Rathore, D. Chaubey, and N. Rajput, "A survey on face detection and recognition," *Int. J. Computer Architecture and Mobility*, vol. 1, no. 5, 2013.

[9] D. Dandotiya, R. Gupta, S. Dhakad, and Y. Tayal, "A survey paper on biometric based face detection techniques," *Int. J. Software and Web Sciences*, vol. 2, no. 4, pp. 67–76, 2013.

[10] P. Shamna, A. Paul, and C. Tripti, "An exploratory survey on various face recognition methods using component analysis," *Int. J. Advanced Research in Computer and Communication Engineering*, vol. 2, no. 5, pp. 2081–2086, 2013.

[11] Y. Bengio, "Learning Deep Architectures for AI," *Foundations and Trends in Machine Learning,* vol. 2, no. 1, pp. 1-127, 2009.

[12] L. Duan, B. Yuan, C. Wu, J. Li, and Q. Guo, "Text-Image Separation and Indexing in Historic Patent Document Image Based on Extreme Learning Machine," in *Proc. of the ELM-2014*, vol. 2, pp. 299-307, 2015.

[13] "LIBSVM Tools," 2017. Avaiable at [https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/]

[14] K. Rujirakul, C. So-In, and B. Arnonkijpanich, "Weighted Histogram Equalized PEM-PCA Face Recognition," in *Proc. of the ICSEC 2014*, pp. 144-150, 2014.

[15] M. Ranzato, J. Susskind, V. Mnih, and G. Hinton, "On deep generative models with applications to recognition," in *Proc. of the CVPR 2011*, pp. 2857-2864, 2011.

[16] G.B. Huang, H. Lee, and E. Learned-Miller, "Learning hierarchical representations for face verification with convolutional deep belief networks," in *Proc. of the CVPR 2012*, pp. 2518-2525, 2012.

[17] T.H. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng, and Y. Ma, "PCANet: a simple deep learning baseline for image classification?," *IEEE Trans. Image Process.*, vol. 24, pp. 5017-5032, 2015.

[18] R. Jafri and H.R. Arabnia, "Survey of Face Recognition Techniques," *J. of Info. Process. Syst.*, vol. 5, no. 2, pp. 41-68, 2009.

[19] G.B. Huang, D.H. Wang, and Y. Lan, "Extreme learning machines: a survey," *Int J Mach. Learn. Cybern.*, vol.2, pp. 107-122, 2011.

[20] W. Zong and G.B. Huang,. "Face recognition based on extreme learning machine," *Neurocomputing*, vol. 74, no. 16, pp. 2541-2551, 2011.

[21] "MATLAB Codes of ELM Algorithm (for ELM with kernels)," 2017. Available at [http://www.ntu.edu.sg/home/egbhuang/elm_kernel.html]

[22] G.B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: a database for studying face recognition in unconstrained environments," *University of Massachusetts, Amherst, Technical Report* 07-49, Oct., 2007.