Illumination Invariant Face Recognition Using Convolutional Neural Networks

N. Pattabhi Ramaiah, Earnest Paul Ijjina and C. Krishna Mohan
Department of Computer Science and Engineering
Indian Institute of Technology Hyderabad
Andhra Pradesh, India 502205
ramaiah.iith@gmail.com, cs12p1002@iith.ac.in, ckm@iith.ac.in

***Abstract*—Face is one of the most widely used biometric in security systems. Despite its wide usage, face recognition is not a fully solved problem due to the challenges associated with varying illumination conditions and pose. In this paper, we address the problem of face recognition under non-uniform illumination using deep convolutional neural networks (CNN). The ability of a CNN to learn local patterns from data is used for facial recognition. The symmetry of facial information is exploited to improve the performance of the system by considering the horizontal reflections of the facial images. Experiments conducted on Yale facial image dataset demonstrate the efficacy of the proposed approach.**

***Index Terms*—biometrics, facial recognition, convolutional neural networks, non-uniform illumination**

I. INTRODUCTION

Face recognition [1] has gained much attention from the research groups of pattern recognition and machine learning since the early 1990s. Face recognition is a difficult task due to the issues with illumination variance which affects the identification rate [2]. The illumination varies while capturing the face photographs, especially in outdoor environment while creating person's identity. A Fuzzy fisher-face approach was proposed for face recognition [3] using fisher discriminant analysis and principle component analysis. Face recognition is used in applications like access control and surveillance [4]. A new framework for face recognition and feature extraction is proposed with kernel fisher discriminant analysis and fisher linear discriminant analysis [5]. Wiskott et. al. [6] proposed face recognition algorithm by using labeled elastic bunch graphs matching based on gabor wavelet transform. A face verification algorithm was proposed [7] for training a very large and unknown data of different categories.

Variations in ambient lighting produces significant degradation in face recognition performance [4]. Thermal infrared (thermal IR) has been used in facial recognition systems with some success against the ambient illumination. Near infrared (NIR) has the potential to overcome the problems associated with visible [8] and thermal IR face recognition which is more robust against illumination variations and face detection [9]. NIR is useful for face detection as the bright eye effect [10] allows the eyes to be localized and skin reflectance properties at just above and below 1.4 microns which highlights the face regions clearly [9]. Over the last decade, convolutional neural networks [11] is widely used for various computer vision tasks [12].

In this paper, we present an approach for face recognition under non-uniform illumination conditions using deep convolutional neural networks. The rest of the paper is organized as follows: Section II, discusses the details of convolutional neural networks. Experimental results of the proposed face recognition approach are discussed in section III. Conclusions are explained in section IV.

II. FACIAL RECOGNITION USING CNN

A convolutional neural network, capable of learning local features from input data is used to discriminate facial images. A typical CNN classifier [13] consists of a CNN with alternating sequence of convolution and sub-sampling layers for feature extraction and a neural network in the last layer for classification. The architecture of CNN classifier considered in this paper is shown in Figure 1.

The template size considered in convolution layers (C1, C2) and sub-sampling layers (S1, S2) are 5×5 and 2×2, respectively. The facial images are down-sampled to images of size 28×32 and given as input to the CNN. The number of feature maps considered in the first feature map set (F1) and the third feature map set (F3) are 6 and 12, respectively. The second feature map set (F2) and the forth feature map set (F4) will have the same number of feature maps as F1 and F3, respectively. The output of the CNN is traversed in a row major order to obtain a column vector (I) of dimension 240×1 which is used by the neural network to classify the input facial images into one of the 30 output classes (O). The last two layers of the CNN classifier i.e., F4 and the Neural network shown in Figure 1 are fully-connected. The CNN classifier is trained using back-propagation algorithm in batch mode, to learn the convolution masks used in C1, C2 and the connection weights between last two layers of the classifier. The following section explains the experimental setup used to evaluate the performance of the proposed approach.

III. EXPERIMENTAL RESULTS

The experiments were conducted on Extended Yale Face Database B [14], [15] which consists of 168×192 gray scale facial images of 38 subjects under 9 poses and 64

978-1-4799-1823-2/15/$31.00 ©2015 IEEE

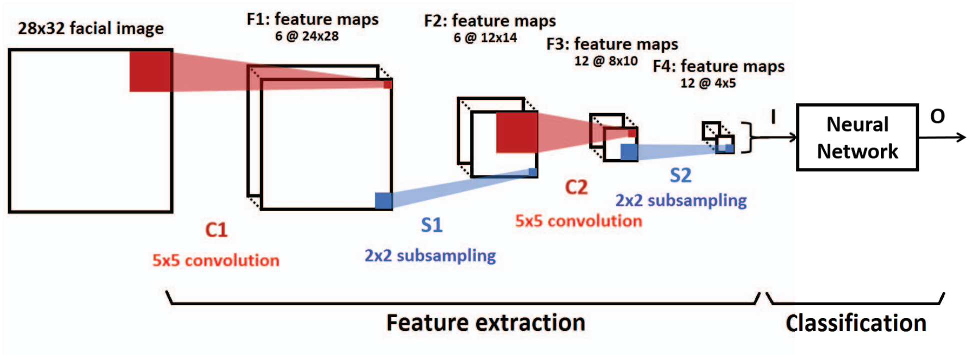

Fig. 1. The deep convolutional neural network architecture considered in the proposed approach

illumination conditions. Five-fold cross validation is used to evaluate the performance of the proposed approach on facial images captured under varying illumination conditions. Similar to other existing approaches, the angle between the direction of light source and the camera axis are considered in grouping the facial images into 5 sets as shown in Table I. We consider 62 illumination images of 30 subjects (a total 1860 frontal face images) from Extended Yale B database, for experimental evaluation of the proposed approaches. The typical distribution of facial images across the five set are shown in Figure 2, where facial images in one row correspond to one subset.

TABLE I
FIVE SETS OF EXTENDED YALE DATABASE B

Set #	1	2	3	4	5
Lighting angle (deg)	0-12	13-25	26-50	51-77	>77

Fig. 2. Typical distribution of Extended Yale Face dataset B facial images across the five sets (only 6 instances per subset are shown)

The CNN is trained using back-propagation algorithm in batch mode with a batch size of 2 for 500 epochs and evaluated using 5-fold cross validation. The variation of classification error with the number of training epochs is given in Table II. The plot of classification performance against number of training epochs is shown Figure 3. An average classification performance of 89.05% is obtained for the proposed approach. The recognition performance of sets #1 and #2 is better when compared to the other sets and relatively low for #3, #4 and #5, which could be due to the inadequacy of training data to capture the necessary discriminative information to recognize the faces in test dataset.

TABLE II
MISCLASSIFICATION ERROR OF THE PROPOSED APPROACH FOR THE FIVE SETS

Set #	5 epochs (%)	50 epochs (%)	250 epochs (%)	500 epochs (%)
1	81.1	0.88	0.4	0.35
2	64.1	3.33	1.96	1.21
3	84.4	18.8	14.72	12.93
4	95	23.8	18.05	17.2
5	88.05	28.3	25.5	23.05

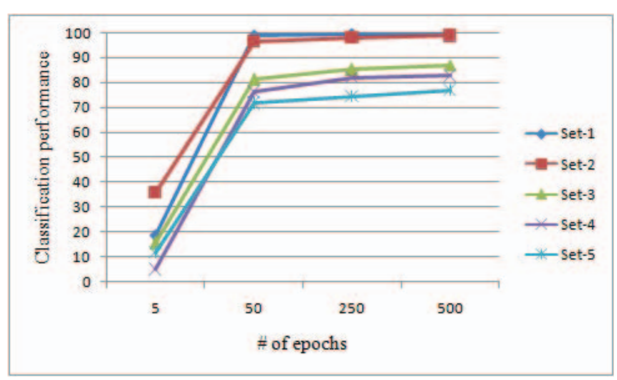

Fig. 3. Classification performance on five sets of Extended Yale Face dataset B

The symmetry of facial information is considered along the vertical line passing through the center of the face image inorder to address the issue. The training dataset is enhanced by including the horizontal reflection of facial images. This enhancement will provide additional information to the classifier especially when there is a shadow on one side of the face. During testing, for a given facial image the maximum value is considered among the evidences generated for each class, to determine the output class. The misclassification error of the proposed approach with this enhancement is given in Table III. Plot of classification performance against training epoch is shown in Figure 4.

TABLE III
PERFORMANCE OF THE PROPOSED APPROACH ON THE ENHANCED FACIAL DATA

Set #	5 epochs (%)	50 epochs (%)	250 epochs (%)	500 epochs (%)
1	41.01	0.88	0.4	0.35
2	96.43	5.21	4.38	0.76
3	64.96	4.08	2.11	1
4	53.16	12.89	5.73	4.89
5	70.11	34.43	24.28	22.95

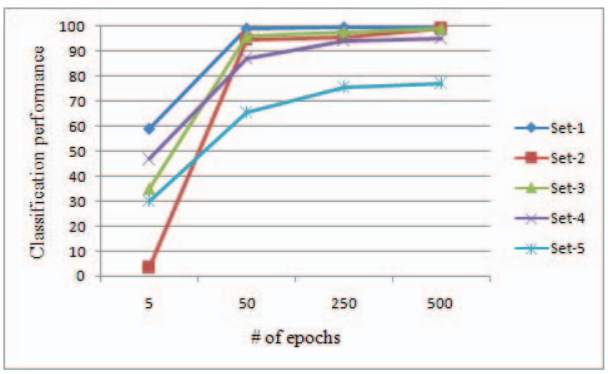

Fig. 4. Classification performance including the horizontally reflected facial images on five sets of Extended Yale Face dataset B

An average classification accuracy of 94.01% is obtained by the proposed approach on the enhanced facial data. It can be observed that the performance of the proposed approach improves by 4.96% when horizontal reflections of the facial data are considered in training.

The performances of existing approaches DT-CWT (dual-tree complex wavelet transform) [16]-[17], LBP (local binary patterns) [18], SQI (self quotient image) [19], LTV (logarithmic total variation) [20], LWT (logarithmic wavelet transform) [21], LNSCT (logarithmic nonsubsampled contourlet transform) [22] along with the two proposed approaches for facial recognition are given in Table IV. The classification performance of the existing methods and the proposed approach are shown in figure 5.

The low classification accuracy for facial images in set-5 is due to the extreme illumination conditions as shown in the last row of Figure 2. We assume that the performance of the proposed facial recognition approaches is unaffected by the inclusion of additional subjects, due to the uniform distribution of subjects facial images across all the five

TABLE IV
COMPARISON OF CLASSIFICATION PERFORMANCE WITH EXISTING APPROACHES

	Set 1 (%)	Set 2 (%)	Set 3 (%)	Set 4 (%)	Set 5 (%)	Avg. (%)
LBP	100	100	62.28	10.34	6.65	55.85
Curvelet	91.66	100	55.92	14.85	5.95	53.67
Contourlet	76.31	98.68	52.19	22.36	9.83	51.87
DT-CWT	98.68	99.34	76.75	38.91	13.85	65.49
SQI	100	98.68	71.27	69.37	63.98	80.66
LTV	100	99.78	78.51	75.75	82.41	87.29
LWT	100	100	82.01	81.95	70.77	86.95
LNSCT	100	100	83.33	87.96	84.34	91.126
Proposed CNN	99.65	98.79	87.07	82.81	76.95	89.05
Proposed CNN (reflection)	**99.65**	**99.24**	**99**	**95.11**	**77.05**	**94.01**

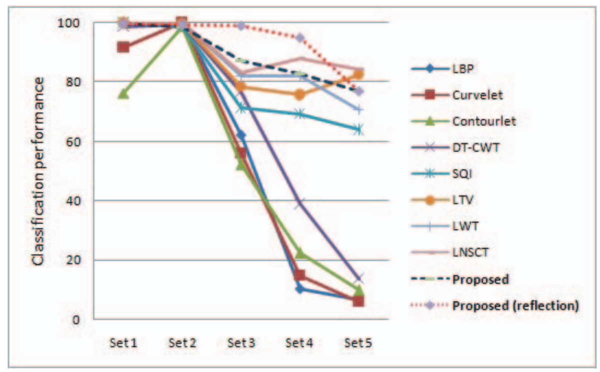

Fig. 5. Comparison of classification performance with existing methods

sets, thereby enabling the comparability of the proposed approach with the existing approaches as shown in Table IV. Table IV shows the effectiveness of the proposed approaches for facial recognition. It is time consuming for training the convolution neural network if the input image size increases.

IV. CONCLUSION

In this paper, we propose two approaches for facial recognition under varying illumination conditions using deep convolutional neural networks. These approaches discriminate the human subjects from the local patterns in their facial information. Experimental studies suggest that there is a significant improvement (4.96%) in the performance of CNN classifier when horizontal reflection of facial data is also considered during training. Assuming performance remains unaffected by the inclusion of additional subjects, the proposed approaches recognize facial images effectively. Future work includes the use of horizontally reflected images in testing and the exploration of various approaches to combine evidences for facial recognition.

REFERENCES

[1] Christophe Garcia and Manolis Delakis, "Convolutional face finder: A neural architecture for fast and robust face detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 26, no. 11, pp. 1408–1423, 2004.

[2] Athinodoros S. Georghiades, Peter N. Belhumeur, and David Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 23, no. 6, pp. 643–660, 2001.

[3] Keun-Chang Kwak and Witold Pedrycz, "Face recognition using a fuzzy fisherface classifier," *Pattern Recognition*, vol. 38, no. 10, pp. 1717–1732, 2005.

[4] Wenyi Zhao, Rama Chellappa, P Jonathon Phillips, and Azriel Rosenfeld, "Face recognition: A literature survey," *Acm Computing Surveys (CSUR)*, vol. 35, no. 4, pp. 399–458, 2003.

[5] Jian Yang, Alejandro F Frangi, Jing-yu Yang, David Zhang, and Zhong Jin, "Kpca plus lda: a complete kernel fisher discriminant framework for feature extraction and recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 2, pp. 230–244, 2005.

[6] Laurenz Wiskott, J-M Fellous, N Kuiger, and Christoph Von Der Malsburg, "Face recognition by elastic bunch graph matching," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 19, no. 7, pp. 775–779, 1997.

[7] Sumit Chopra, Raia Hadsell, and Yann LeCun, "Learning a similarity metric discriminatively, with application to face verification," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. IEEE, 2005, vol. 1, pp. 539–546.

[8] Diego A Socolinsky, Andrea Selinger, and Joshua D Neuheisel, "Face recognition with visible and thermal infrared imagery," *Computer Vision and Image Understanding*, vol. 91, no. 1, pp. 72–114, 2003.

[9] Jonathan Dowdall, Ioannis Pavlidis, and George Bebis, "Face detection in the near-ir spectrum," *Image and Vision Computing*, vol. 21, no. 7, pp. 565–578, 2003.

[10] Carlos Hitoshi Morimoto, David Koons, Arnon Amir, and Myron Flickner, "Pupil detection and tracking using multiple light sources," *Image and vision computing*, vol. 18, no. 4, pp. 331–335, 2000.

[11] AR Syafeeza, M Khalil-Hani, SS Liew, and R Bakhteri, "Convolutional neural network for face recognition with pose and illumination variation.," *International Journal of Engineering & Technology (0975-4024)*, vol. 6, no. 1, 2014.

[12] Yann LeCun, Koray Kavukcuoglu, and Clément Farabet, "Convolutional networks and applications in vision," in *IEEE International Symposium on Circuits and Systems (ISCAS 2010)*, 2010, pp. 253–256.

[13] R. B. Palm, "Prediction as a candidate for learning deep hierarchical models of data," M.S. thesis, Technical University of Denmark, Asmussens Alle, Denmark, 2012.

[14] Peter N. Belhumeur, João P Hespanha, and David Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 19, no. 7, pp. 711–720, 1997.

[15] Kuang-Chih Lee, Jeffrey Ho, and David Kriegman, "Acquiring linear subspaces for face recognition under variable lighting," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 5, pp. 684–698, 2005.

[16] Nick Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *Applied and computational harmonic analysis*, vol. 10, no. 3, pp. 234–253, 2001.

[17] Aryaz Baradarani, Q. M. Jonathan Wu, and Majid Ahmadi, "An efficient illumination invariant face recognition framework via illumination enhancement and dd-dtwt filtering," *Pattern Recognition*, pp. 57–72, 2013.

[18] Timo Ahonen, Abdenour Hadid, and Matti Pietikainen, "Face description with local binary patterns: Application to face recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 12, pp. 2037–2041, 2006.

[19] Haitao Wang, Stan Z Li, and Yangsheng Wang, "Generalized quotient image," in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*. IEEE, 2004, vol. 2, pp. II–498.

[20] Terrence Chen, Wotao Yin, Xiang Sean Zhou, Dorin Comaniciu, and Thomas S Huang, "Total variation models for variable lighting face recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 9, pp. 1519–1524, 2006.

[21] Taiping Zhang, Bin Fang, Yuan Yuan, Yuan Yan Tang, Zhaowei Shang, Donghui Li, and Fangnian Lang, "Multiscale facial structure representation for face recognition under varying illumination," *Pattern Recognition*, vol. 42, no. 2, pp. 251–258, 2009.

[22] Xiaohua Xie, Jianhuang Lai, and Wei-Shi Zheng, "Extraction of illumination invariant facial features from a single image using nonsubsampled contourlet transform," *Pattern Recognition*, vol. 43, no. 12, pp. 4177–4189, 2010.