Identifying Humans by Matching their Left Palmprint with Right Palmprint Images using Convolutional Neural Network

Ajay Kumar, Kuo Wang
Department of Computing, *The Hong Kong Polytechnic University*, *Hong Kong*

Abstract—*Palmprint is increasingly adapted as one of the effective modalities for the biometrics identification. There exists high degree of similarity (e.g., spatial arrangement of major flexion creases) between left- and right-hand human palms. However the problem of matching such left palmprint images with the right palmprint images is considered challenging and has not yet attracted the attention of researchers. Such problem arises when only palmprint from one of the hand is available in the registration database while palmprint from the other hands are to be matched or during other forensic analysis. This paper explores on such possibility of matching left with the right palmprint images and attempts to answer the key question on best possible performance. Several promising algorithms are investigated while the best results are obtained from a convolutional neural network (CNN). Therefore, this paper presents first successful usage of CNN for the palmprint matching with outperforming results over most accurate methods available today. We use publicly available palmprint database from 193 different subjects to investigate the performance. Our experimental results using the methods described in this paper systematically validates our approach and achieves encouraging performance.*
Keywords— biometrics, forensics, palmprint, deep learning, convolutional neural network

1. INTRODUCTION

Development of Human identification using biometrics is increasingly used in e-business applications for social welfare programs and also to ensure high degree tradeoff between the security and convenience. Choice of biometric modality often depends on nature of application, user convenience and also on the level of security expected for the intended application. Hand based biometric modalities are easy to present or imaged and have shown to offer high degree of performance which makes them most sought for e-business applications. This is plausible reason for their highest market share. There has been significant advancement in the accuracy of 2D and 3D palmprint matching algorithms available today. Typical matching of palm images is performed using same protocols like for any other biometric modalities and like for fingerprint or iris, the left and right side impressions are popularly considered to be quite unique in establishing identity of the subjects.

Humans are known to accompany left-right symmetry from the early development of embryo [16]. This symmetry is revealed in two eyes, ears, hands and many structures in between. There are a range of survival benefits from left-to-right symmetry, *e.g*. with symmetrical muscles and limbs on both sides, our movement is quick and efficient. However our bodies are not perfectly similar [16] and some earlier work in biometrics has attempted to exploit such asymmetry to aid in the identification, *e.g*. asymmetry of human faces is explored for the face recognition in [11].

A. Motivation

Conventional research and development efforts in the palmprint have been largely focused on the development of accurate and efficient algorithms for the left-to-left or right-to-right palmprint matching [4], [7], [13]. This was reasonable as the research has been typically target for business or commercial applications for which the identity of hands was easy to establish/present and need for left-to-right palmprint matching was uncommon. However there are many cases when automated human identification system mistakenly acquired person's right hand instead of the left hand, or the left hand instead of right hand, which makes an actual genuine matching from the same person to be recognized as an imposter match. In several biometrics applications, it is difficult to impose constraints on how a biometric trait should be acquired [8].In addition to its use in forensics [1], when both left hand samples and right hand samples available from the same person, the cross matching between left hand and right hand can also be achieved to achieve higher confidence in establishing the identity of the person. Therefore the left-to-right palmprint matching can also be of importance as of the left-to left or right-to-right matching.

Among three groups of flexion creases, the major flexion creases are widely considered to be the largest [12]. The similarity in the spatial arrangement of minor finger creases and major flexion creases, *i.e*., distal transverse crease (heart line), proximal transverse crease (head line), and radial transverse crease (life line), between the left and right palms can be explored to establish identity of a person. Figure 1-2 illustrate sample images from publicly available contact-based [3] and contactless [6] palmprint databases. These images indicate correlation among minor finger creases and major flexion creases between the subject's left hand and right hand palmprints. This correlation is reveled in crease length, width and shape while there are also many differences including in minor crease lengths, in small scars, *etc*. Any effort to generate matches from such pair of images should firstly align these two images (like in figure 1-2) and quantify the spatial similarity between these major line patterns.

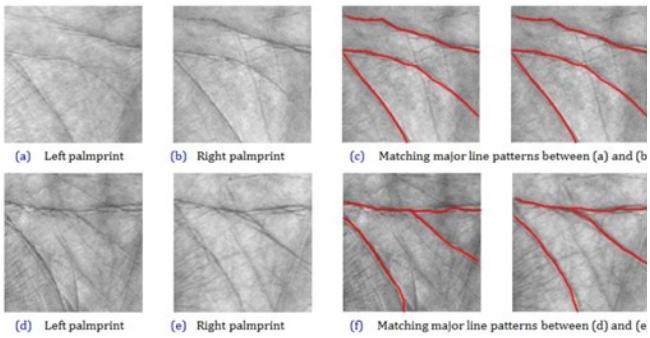

Figure 1: Sample palmprint images of two different subjects in PolyU database. Images are aligned/flipped to ascertain similarity between the palmprint images from two hands of same subject.

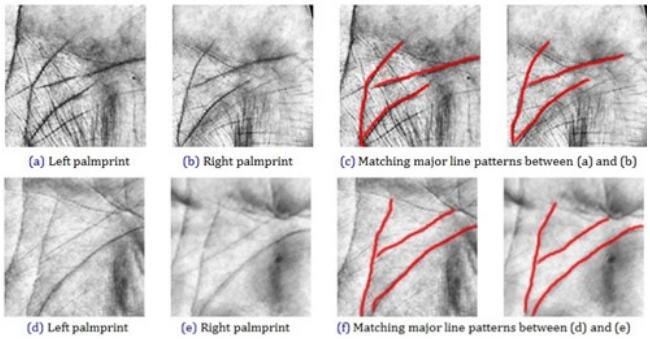

Figure 2: Sample palmprint images of two different subjects in IIT Delhi contactless palmprint database [6]. Images are aligned/flipped to ascertain similarity between the palmprint images from two hands of same subject.

2. PALMPRINT MATCHING ALGORITHMS

Two palmprint images to be matched from the same hands should be firstly aligned to ensure that any possible similarity between the major lines and creases can be revealed for the matching algorithm. If the alignment between such palmprint images is unknown, the best approach is to consider different possibilities for alignment, *i.e.* rotation in steps and/or translation in steps, and then use best match score among all match scores generated from different possibilities of alignments. In our work we used publicly available palmprint databases and course alignment between any two left and right images is achieved by left-to-right flip operation. Image samples in figure 1-2 illustrate images *after* such operations and the aligned images helps to reveal symmetry among major palm lines and crease patterns. The spatial domain approach in [15] was selected as the baseline as it achieves higher accuracy than several methods (*e.g.* using method in [4]) for *same* palmprint matching. The frequency domain approach in [7] has shown to offer highest accuracy, among other approaches available for the *same* palmprint matching. These two algorithms, along with our modifications for [7] to further improve accuracy for the left-to-right palmprint matching, are briefly described in following sections.

A. Neighborhood Matching Radon Transform (NMRT)

The neighborhood matching Randon transform approach firstly extracts the dominant orientation feature template in a palmprint image. For each pixel, the dominant orientation is computed from the dominant localized Randon transform (LRT) orientation [15].

$$Ao(x,y) = arg\min_p\big(LRT\ I(k)\big),\ p = 1,2,\dots Z \quad (1)$$

where $Ao(x,y)$ represents dominant direction of local palm line/creases for k^{th} point in image I while Z is total number of discretized directions. In the matching stage, given or registered palmprint template (T_r) and the unknown palmprint template (T_i) from the other hand is matched. The matching score $\varrho\ (T_r, T_i)$ between the two templates is the minimum of the normalized hamming distance (ψ) between the two different templates with different amounts of alignment or the translation.

$$\varrho\ (T_r, T_i)\ =\ \min_{\forall g\in[-m,m],\forall h\in[-n,n]} \frac{\psi\ (T_r, T_i(g,h))}{size\ (T)} \quad (2)$$

This NMRT scheme is quite similar to LRT in [4] but capable of generating more accurate matches from the local palm regions and was therefore selected for preliminary performance evaluation. We empirically selected the number of discrete orientations as 10, the line width as 4 pixels and the lattice size of 16. We performed preliminary evaluation between left-and-right on the *subset* of 20 subjects in PolyU database [3] to ascertain its effectiveness. This database has palmprint images acquired from the two different sessions. We selected the best performance from different combinations of the sessions as baseline for selecting the matcher. The number of resulting genuine scores were 1,947 and the number of resulting imposter scores were 370, 560. The receiver operating characteristics (ROC) from this preliminary test is shown in figure 3. The minim equal error rate (EER) for ROCs in this figure was 23.04% (from 1st session left to 1st session right palm) while maximum EER was 25.38% (from 2nd session left to 1st session right).

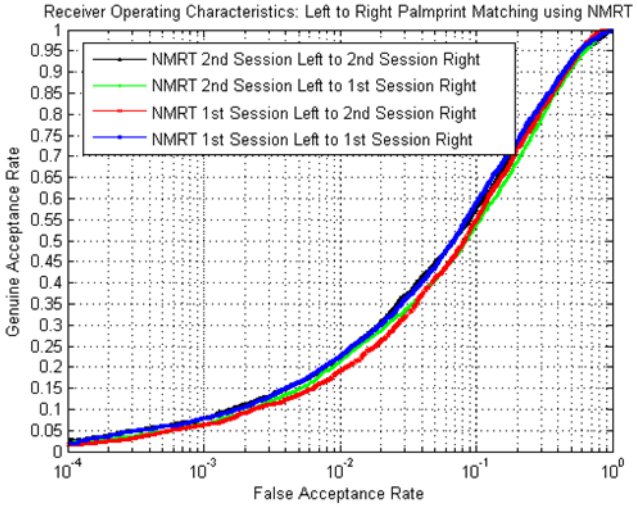

Figure 3: Matching performance using NMRT from test subset with different combinations of sessions for left-right matching.

It was observed from above results that although the 1st session left palm to 1st session right palm achieves slightly better performance when EER is considered as the metric. The ROC in figure 3 however suggests that these performances are quite similar. It is judicious to choose a criterion using test on 1st session left palm to 2nd session right palm (EER was 24.24%) in comparing the competing palmprint methods. Capability to generate local line matches along the dominant palm lines motivated us to select next competing approach which is summarized in next section.

B. Band Limited Phase Only Correlation

This approach for matching two palmprint images considers local correlation of dominant grey-levels in spatial-frequency domain. The accuracy of this method [13] has been improved when only (phase of) correspondence points [7] between two palmprint region is matched to establish the identity of unknown palm. Therefore this method was also considered for the preliminary evaluation of the performance. The phase-based correspondence matching computes the matching score as the maximum peak value of average Band-Limited Phase-Only Correlation (BLPOC) function based on reference points selected by a Difference of Gaussians (DoG) filter.

$$D(n_1, n_2, \sigma_1, \sigma_2) = \frac{1}{2\pi{\sigma_1}^2} exp(-\frac{{n_1}^2+{n_2}^2}{2{\sigma_1}^2}) \frac{1}{2\pi{\sigma_2}^2} exp(-\frac{{n_1}^2+{n_2}^2}{2{\sigma_2}^2}) \quad (3)$$

The DoG filter $D(n_1, n_2, \sigma_1, \sigma_2)$ in above equation is defined by the difference between two Gaussian filters whose standard deviations are σ_1, σ_2. In our work we empirically selected $\sigma_1 = 0.65$ and $\sigma_1 = 0.2$ for two filters. Using such an approach of DoG-based reference point selection, we observed that generally more number of reference point can improve the performance while high number of reference points will also locate in minor creases which can be considered undesirable for matching let-to-right palmprint images. Our experimental results (figure 5, table 1) on a subset of 20 subjects suggest that 144 is the reasonable choice as the number of reference points using this approach.

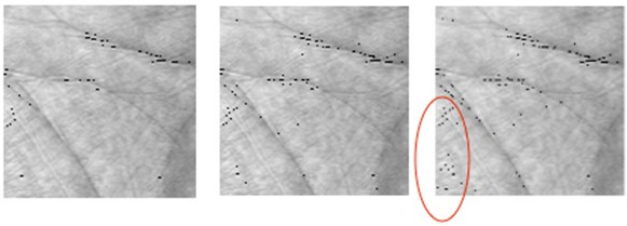

Figure 4: DoG-based reference point selection with (a) 64 reference points; (b) 100 reference points; (c) 144 reference points

Comparative performance from the ROC in figure 3 and 5 also suggests effectiveness and promises from this approach for the cross-palmprint matching. Therefore this approach was further considered and developed for robustly matching the left palmprint and the right palmprint images.

Table 1: The EER with different numbers of reference points on DoG based BLPOC scheme for left to right palmprint matching

Number of reference points	EER
64	19.96%
100	17.68%
144	18.3%

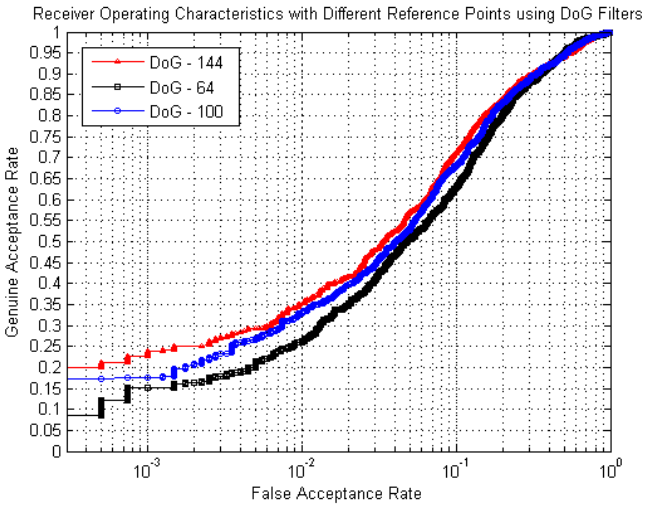

Figure 5: Performance using DoG-based reference point selection on test subset with different number of reference points.

C. Localizing Key Reference Points between Left and Right Palm Images

Our experiments in previous section indicated that correlation of phase based local matching can offer better performance and this approach was further considered in this work. In order to avail most of the prevailing correlation between the left palmprint and right palmprint, DoG-based reference point selection approach center's the reference points on the major palm lines. However as can also be observed from figure 4, when we increase the number of reference points, there are some spurious reference points that deviate from the major lines and these are undesirable or noisy reference for accurate left-to-right palmprint matching. As also observed in [10], it is due to the nature of DoG filter that makes it more suitable for (left-to-left or right-to-right) palmprint texture analysis.

In order to suppress the noisy reference points, we introduce a new reference point selection approach using the combination of efficient edge detectors. There are several edge detection filters in the computer vision literature [17]. Among several edge detection filters considered in this work, we observe that three filters, *i.e.*, Prewitt, Roberts and Sobel, can recover the major palm lines in more effective manner. Given a gray level palmprint image, the edge points that can localize the palm lines are defined as those pixels whose gradient magnitude exceeds a certain predefined threshold. Different filters use different convolution kernels to acquire gradient information from different orientations.

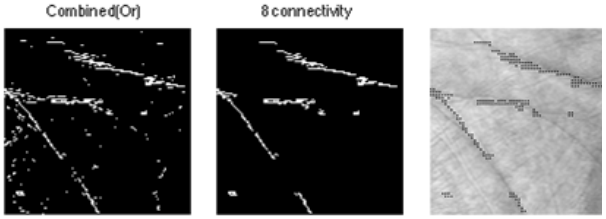

Figure 6: (a) Combined binary response from the three filters, (b) (b) eliminating responses from minor lines using post-processing, and (c) selected reference points from combined edge detection.

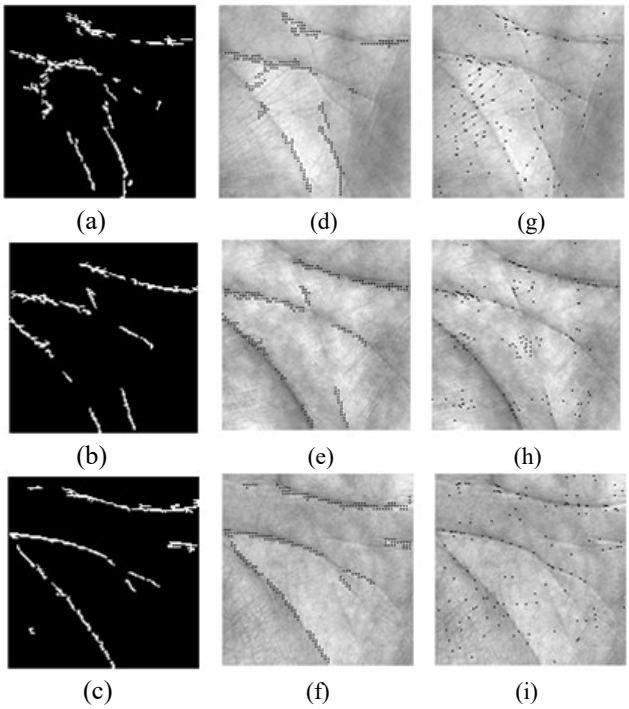

Figure 7: (a)-(c) Binary responses from the edge detectors, (d)-(f) selected reference points for matching, (g)-(i) DoG based 144 points on the same palmprint image for the comparison.

The Prewitt filter is used for detect the edges in two vertical and horizontal orientation. The Roberts filter is known to respond maximally to the edges running at 45° to the pixel grid. The Sobel filter responds maximally to edges running vertically and horizontally relative to the pixel grid.

One plausible approach to generate more effective response from major palm lines is to combine responses from (above) multiple filters and use some heuristics to discard noisy key points from the combined responses. The binary response from the respective filters are expected to have complimentary information for the detection of *curved* major lines as the trajectory of such major palm lines changes several times all along its way. The combination of binary response can enhance the detection of major lines by increasing the pixel connectivity along these palm lines. This approach is designed to ensure higher pixel connectivity for the major palm lines, *i.e.*, the region where the similarity between two palms to be matched is expected/suspected, while smaller pixel connectivity for the noisy minor palm lines that can be discarded in post processing operation.

In order to reduce the number of reference points to a comparable size to those from the DoG-based reference point selection scheme, the final steps of combined edge detection approach are down-sampling and up-sampling operations. Every four neighboring pixels in the combined binary response will translate to a single reference point. Therefore the number of reference points selected using this approach varies between 160 to 220, most of these points were observed to be located along the major palm lines. Although high number of reference points along the major palm lines are expected to generate superior performance, the DoG-based approach with or over 144 reference points already starts generating noisy reference points that are off the major palm lines (as shown in the 3rd column in figure 7 or from red circle in figure 4). The pseudocode in the following summarizes the steps employed for the key point selection scheme while matching left-to-right palmprint images.

Function $[p_0, p_1, p_2] = \text{KeyPointDetection}(A)$

$A' \rightarrow$ Enhancement of aligned palmprint image A

$[A_1\ A_2\ A_3] \rightarrow$ Binary responses from A' using Prewitt, Roberts and Sobel filters

$A_c = A_1 \parallel A_2 \parallel A_3$

$A'_c \rightarrow$ Morphologically open binary image of A_c (less than 8-connected pixels)

Locate reference points in $A'_c \rightarrow$ Get p_1 by down sampling to the first layer

Get $p_2 \rightarrow$ Down sampling p_1, Get $p_0 \rightarrow$ Up sample p_1 *i.e.*, $p_l = \left\lfloor \frac{1}{2}\, p_{l-1} \right\rfloor,\ l = 1, 2$

3. PALMPRINT MATCHING USING CNN

In order to explore wide correlation of features among two palmprint images, a CNN [5] based approach was also investigated. The CNNs have so far not been explored for the palmprint matching but it is judicious to explore the remarkable capabilities of CNNs in automatically learning the effective features for more accurate palmprint identification. The CNN is essentially a kind of neural network [2] which uses multiple layers (convolution pattern) to connect each neuron. The architecture of this network consists of combination of different layers, like as convolution layer, pooling layers, normalization layer, fully connected layers, loss layers, *etc.*, and the network can

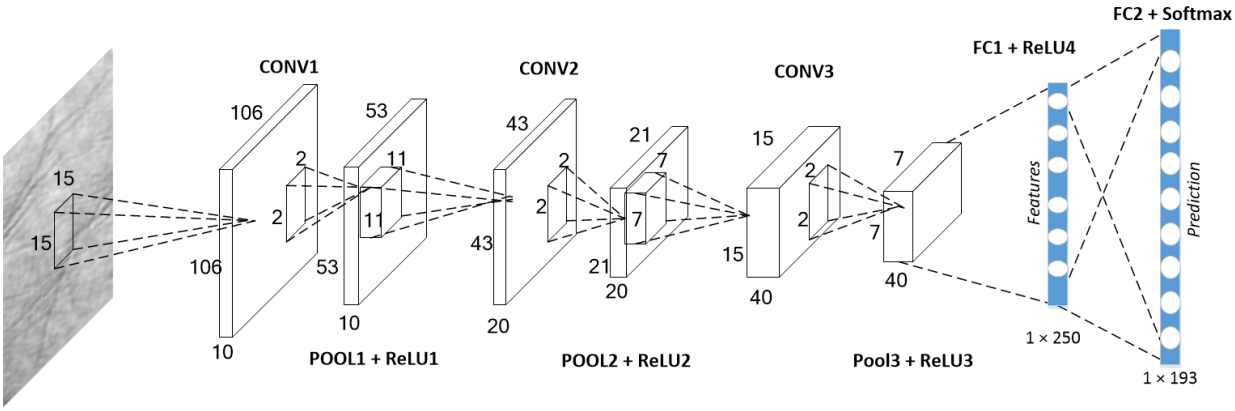

Figure 8: The architecture of convolutional neural network employed for the palmprint matching in our experiments.

automatically learn best set of parameters through forward propagation and backward propagation during the training phase. The CNN architecture employed in our work to ascertain matching performance for the palmprints images is shown in figure 8. Our CNN architecture includes three convolutional units and two fully connected layers. The input x^j from the convolution units firstly enters the convolutional layer.

$$y^i = \sum_j (w^{ij} * x^j) + b^{ij} \quad (4)$$

where x^j is the input for the j^{th} channel of convolutional layer, y^i is the output from the i^{th} channel, w^{ij} is the weights for the convolutional layer and b^{ij} represents bias for the respective neuron. The parameters w^{ij} and b^{ij} are learned during the learning process. In the polling layer maximum of inputs is extracted from the kernel. The output layer uses rectified linear unit [18] nonlinear as the activation function:

$$y^i = \max\ (y^i, 0) \quad (5)$$

The fully connected layer processes the output from the three convolutional units and connects all the outputs from the current layer.

$$y^i = b^i + \sum_j x^j w^{ij} \quad (6)$$

The final output is achieved from the last fully connected layer by using a *softmax* function.

$$\hat{y}_i = \frac{e^{y_i}}{\sum_{j=1}^{N} e^{y_j}} \quad (7)$$

The output is essentially a 1 × N vector (figure 8) whose values represents the probability of a predefined class. Cross entropy loss function is minimized during the learning phase to ensure higher probability for correct class labels.

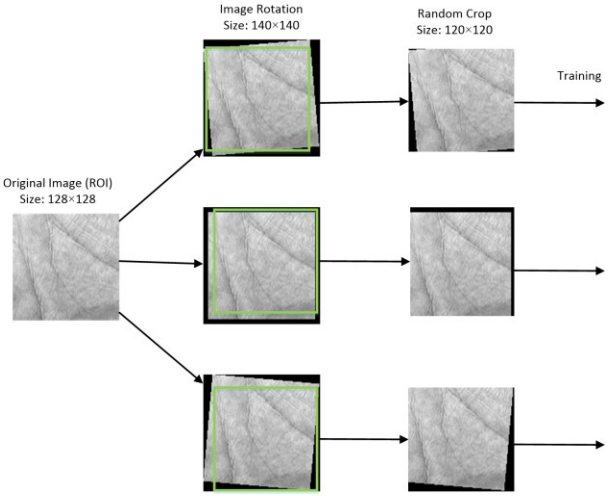

Figure 9: Augmentation and enrichment of palmprint images during the training of CNN employed in figure 8.

4. Experiments and Results

The effectiveness of the key point based left-to-right palmprint matching scheme detailed in section 2-3 was experimentally evaluated on entire PolyU dataset. This PolyU palmprint database consists of 7,753 palmprint images from the two sessions. We evaluated the possible combinations of palmprint pairs between left hand and the right hand for all 193 different subjects in this database. The numbers of genuine scores for the left to right palmprint matching using second session or the test data were 19,550 and the numbers of imposter scores were 7,497829. The ROC from the experiments is shown in figure 9. Table 2 summarizes the EER achieved from these two methods. It can be observed that the key point based approach detailed in section 2-C achieves superior performance than DoG scheme for the left-to-right palmprint matching.

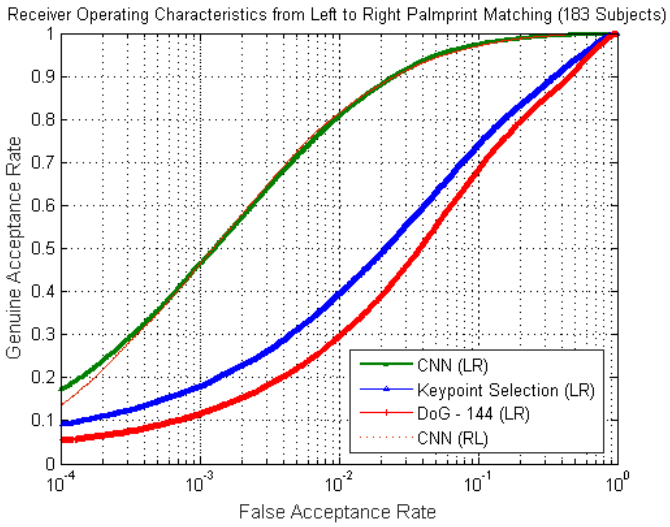

Figure 10: Comparative Left-to-Right palmprint matching performance from different methods using PolyU database.

The CNN detailed in section 3 was also used to evaluate performance. The training set from the first session images was also enriched by rotating each of the images by $\pm$ 5 degrees as shown in figure 9. Therefore total of 3863 training images from 193 subjects first session database were enriched to 11589 images (three times) for the training. Our CNN model was trained with Caffe [19] on a single NVIDIA GTX670 platform. We assigned same class labels to the left and right palmprint images to let CNN learn features among these image pairs from first session data. The comparative ROC using the trained model on unseen second session database is illustrated in figure 10. It can be observed from this ROC that the matching using CNN achieves outperforming results. Table 2 also indicates significant reduction in EER using the CNN based approach.

Table 2: The EER from Left-to-Right Palmprint matching.

Approach	EER
BLPOC – DoG based reference point selection	20.34%
BLPOC – Keypoint detection and selection	18.01%
Convolutional Neural Network	**9.25%**

It may be noted that left to right palmprint maching can generate *different* results than right to left palmprint matching. Therefore we also evaluated the matching performance from the best matcher or CNN when the right palmprint images from the second session images are matched with left palmprint images. The corresponding performance using the ROC is also displayed in figure 10.

The EER in this case was 8.94% and was slightly better than those from left-to-right palmprint matching The distribution of genuine and impostor matching scores from the best performing matcher using the test data is shown in figure 11.

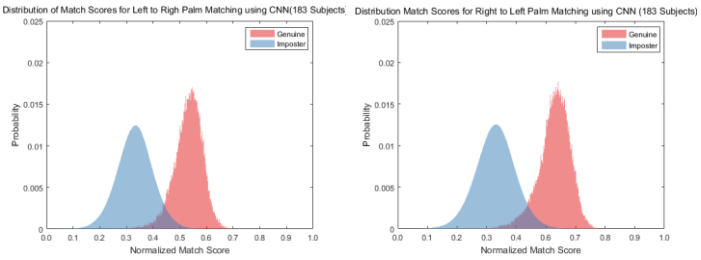

Figure 11: Distribution of genuine and impostor matcing scores from the CNN using second session PolyU database.

5. Conclusions and Further Work

This paper has investigated on the possibility of matching left palmprint images with the right palmprint images to establish the identity of an individual. In order to achieve more accurate matching for the left-to-right matches, we explored a superior scheme to establish key points along the palmer creases and generate more accurate matches. Our experimental results, on the database from 193 different subjects, presented in section 4 (ROC in figure 10 and table 2) suggest significant improvement in the performance as compared to the BLPOC approach using reference points. However the most accurate results for the left to right palmprint matching were obtained from a trained CNN detailed in section 3. The plausible reason for superior performance from CNN lies in its remarkable capability to automatically learn best set of features, among the two palmprints, during the training stage.

Matching of left palmprint to right palmprint patterns can be used in forensics as a supporting tool to establish identity of suspects. Therefore tolerance for the errors can be typically higher for such applications as the law-enforcement officers can use some of the contextual information (non-biometrics) to eliminate some of the false matches [8]. In this context, the error rates shown in table 2 or figure 10 from 193 different subjects, although high as compared to those from same palm matches in the literature, are quite encouraging and suggests promises from this research. The objective of our investigation was not to verify or support claims on the similarity of minor finger creases and/or major flexion creases from the two hands but to scientifically evaluate such possibility for making identifications using publicly available database. The performance shown in section 4 is expected to further improve by using superior CNN architecture, such as in [14], and/or by developing new ways to accommodate deformations in the palms that are matched. Further performance evaluation on other publicly available palmprint databases is part of our ongoing work in this area.

References

[1] D. R. Ashbaugh, "Palmar flexion crease identification," *Journal of Forensic Identification*, vol. 41, no.4, pp. 255-273, Jul./Aug. 1991.

[2] A. Kumar, "Neural network based defection of local textile defects," *Pattern Recognition*, pp. 1645-1659, July 2003.

[3] *PolyU Palmprint Database*, Available Online: http://www.comp.polyu.edu.hk/~biometrics, 2015.

[4] W. Jia, D.-S. Huang, and D. Zhang, "Palmprint verification based on robust line orientation code," *Pattern Recognition*, vol. 31, no. 5, pp. 1504–1513, May 2008.

[5] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed and D. Anguelov, "Going deeper with convolutions," *Proc. CVPR 2015*, pp. 1-9, Boston, June 2015.

[6] IIT Delhi Touchless Palmprint Database (Ver. 1.0), Available Online, 2015, http://www.comp.polyu.edu.hk/_csajaykr/IITD/Database_Palm.htm

[7] S. Aoyama, K. Ito, T. Aoki, "Similarity measure using local phase features and its application to biometric recognition," *Proc. CVPR 2013 Biometrics Workshop*, pp. 180-187, 2013.

[8] A. K. Jain, K. Nandakumar, A. Ross, "50 years of biometric research: accomplishments, challenges, and opportunities," *Pattern Recognition Letters*, vol. 79, pp. 80-105, Aug. 2016.

[9] J. Zbontar and Y. LeCun, "Stereo matching by training a convolutional neural network to compare image patches," *J. of Machine Learning Research*, pp. 1-32, vol. 17, 2016

[10] X. Wu, K. Wang, and D. Zhang. "Palmprint texture analysis using derivative of Gaussian filters." *Proc. IEEE Intl. Conf. Comp. Intelligence and Security*, pp. 751-754, Nov. 2006.

[11] S. Gutta, H. Wechsler, "Face recognition using assymetrical faces", *Proc. ICBA 2004*, pp. 162-168, 2004.

[12] A. K. Jain and M. Demirkus, "On Latent Palmprint Matching," *MSU Technical Report*, Michigan State University, MSU-CSE-08-8, May 2008.

.[13] K. Ito, T. Aoki, H. Nakajima, K. Kobayashi, T. Higuchi, "A palmprint recognition algorithm using phase-only correlation," *IEICE Trans. Fundamentals of Electronics, Communications and Computer Sciences*, vol. E91-A, pp. 1023-1030, April 2008.

[14] Z. Zhao and A. Kumar, "Accurate periocular recognition under less constrained environment using semantics-assisted convolutional neural network," *IEEE Trans. Info. Forensics and Security,* available online, Dec. 2016.

[15] Y. Zhou and A. Kumar. "Human identification using palm-vein images." *IEEE Trans. Info. Forensics and Security*, vol. 6, pp. 1259-1274, 2011.

[16] C. Zimmer, The big similarities and quirky differences between our left and right brains, *Discover*, Apr. 2009. http://discovermagazine.com/2009/may/15-big-similarities-and-quirky-differences-between-our-left-and-right-brains

[17] R. C. Gonzalez and R.E. Woods, *Digital Image Processing,* Pearson, 3rd Ed., 2009.

[18] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," *Proc. ICML*, pp. 807-814, 2010.

[19] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," *Proc. 22nd ACM Int. Conf. Multimedia*, pp. 675-678, Nov. 2014.