Multi-Siamese Networks to Accurately Match Contactless to Contact-based Fingerprint Images

Chenhao Lin, Ajay Kumar
Department of Computing, The Hong Kong Polytechnic University, Hong Kong
csclin@comp.polyu.edu.hk, ajay.kumar@polyu.edu.hk

Abstract

Contactless 2D fingerprint identification is more hygienic, and enables deformation free imaging for higher accuracy. Success of such emerging contactless fingerprint technologies requires advanced capabilities to accurately match such fingerprint images with the conventional fingerprint databases which have been developed and deployed in last two decades. Convolutional neural networks have shown remarkable success for the face recognition problem. However, there has been very few attempts to develop CNN-based methods to address challenges in fingerprint identification problems. This paper proposes a multi-Siamese CNN architecture for accurately matching contactless and contact-based fingerprint images. In addition to the fingerprint images, hand-crafted fingerprint features, e.g. minutiae and core point, are also incorporated into the proposed architecture. This multi-Siamese CNN is trained using the fingerprint images and extracted features. Therefore, a more robust deep fingerprint representation is formed from the concatenation of deep feature vectors generated from multi-networks. In order to demonstrate the effectiveness of the proposed approach, a publicly available database consisting of contact-based and respective contactless fingerprints is utilized. The experimental evaluations presented in this paper achieve outperforming results, over other CNN-based methods and the traditional fingerprint cross matching methods, and validate our approach.

1. Introduction

Fingerprints are widely considered to be the most reliable biometrics and widely used in arrange of human identification systems [20]. In order to address the limitations of conventional contact-based fingerprint acquisition technology, contactless fingerprint acquisition technologies have been introduced [12, 14, 18, 30]. These systems are more hygienic, offer higher user convenience and therefore increasingly deployed for e-governance and business applications. With the development of contactless fingerprint recognition technology and large-scale use of different fingerprint systems, contactless to contact-based fingerprint sensor interoperability has drawn increased attention.

There have been some promising efforts [3, 6, 14, 15] to investigate the sensor interoperability problem which have indicated significantly degradation in the performance for matching fingerprint images, using the conventional minutiae-based matching method, that are acquired from the contactless and contact-based sensors. The deformation of contact-based fingerprint and perspective distortion of contactless fingerprint make it very difficult to achieve accurate minutiae extraction and matching using conventional methods. Therefore this problem deserves more attention and is also important for advancing the deployments of contactless fingerprint technologies as the legacy databases from billions of subjects have been acquired using contact-based fingerprint technologies.

Recently, deep learning technologies [5, 11, 25] have been introduced and have shown remarkable performance for the image classification and recognition. Comparing with recognizing images from different categories problem, identifying the images from same category problem, like fine-grained problem and biometrics recognition, is more challenging. Convolutional neural network based approaches also have accomplished success in such problems [2, 26, 27] and have outperformed many conventional hand-crafted feature based matching methods. For the cross-fingerprint matching problem, because the hand-crafted features, i.e. minutiae, cannot be accurately extracted, it's meaningful to investigate such problem using the CNN-based approach.

1.1. Related Work

The cross-sensor fingerprint matching problem has attracted the attention of many researchers [14, 23]. Among such cross-sensor matching problems, the contactless to contact-based fingerprint matching problem is more challenging and the related research has shown such difficulties and challenges. Two related technical reports [3, 6] have dis-

978-1-5386-1124-1/17/$31.00 ©2017 IEEE

cussed several popular commercial contactless and contact-based sensors and have investigated their interoperability. The experimental results have indicated significant degradation in matching performance using the fingerprints from different sensors. The main reason for performance degradation can be the contact-based fingerprint deformation and inaccurate minutiae extraction. In order to address these problems, reference [15] introduced a contact-based fingerprint deformation correction model and illustrated improved performance on contactless to contact-based fingerprint matching. However, the performance still cannot meet the expectations of the biometrics community for the real application deployments.

In recent years, CNN based approaches outperform the conventional state of art methods in many fields of image classification and recognition [5, 25, 28]. In terms of biometrics recognition and feature representation, CNN-based approaches have also achieved significant improvement [17, 19, 26, 27]. For example, in reference [17], the authors used deep CNN to learn the discriminative representations on challenging face images and demonstrated their method achieve remarkable better performance than existing commercial products using conventional method. Damaged fingerprint identification was discussed in [33]. A CNN-based approach using center images of damaged fingerprint was proposed for fingerprint identification. Their reported results showed that fingerprint recognition based on CNN had a higher robustness than conventional minutiae-based matching. Reference [31] learned two Siamese CNNs for cross-domain sketch retrieval problem and demonstrated their method was better than state-of-art methods and other CNN-based approaches. Such Siamese structure is also suitable for fingerprint cross-sensor matching problems.

More recently, researchers have investigated fine-grained recognition problem using CNN rather than traditional hand-crafted features. In reference [32], the authors proposed a deep ranking model using a triplet sampling algorithm to learn the fine-grained image similarity. The results showed that their method outperformed the method based on hand-crafted visual features. Reference [9] integrated specific sematic part of object into CNN for object detection and fine-grained classification. The experimental results indicated better performance by using their methods on small part detection and fine-grained classification. A CNN-based bilinear model is proposed in reference [16] to generate fine-grained image descriptor. The network was firstly fine-tuned on ImageNet dataset to generate initialized parameters. The results demonstrated the effectiveness of their methods on various fine-grained recognition datasets. Fingerprint recognition is similar to fine-grained recognition problem because they both need to identify the images from same category.

Figure 1: Illustration of *high intra-class variation* between contactless fingerprints and respective contact-based fingerprints due to different acquisition approaches and user habits.

1.2. Open Challenges and Our Work

Accurate cross-sensor fingerprint image matching using the conventional CNN is challenging, especially for the fingerprints acquired from contactless and contact-based sensor. For fingerprint recognition problem, the images have *low inter-class variation* because they are from same category. Therefore, matching such fingerprints is more difficult than recognizing images from different categories. On the other hand, as compared to the fingerprint recognition using the same sensor, fingerprint crosssensor matching is more challenging because the cross-sensor images have high intra-class variance due to the different acquisition approaches and user habits. Figure 1 illustrates the fingerprint image samples with *high intra-class variation*. In either case, it's difficult to achieve good performance by directly using conventional CNN on fingerprint cross matching problem.

Traditional fingerprint matching highly relies on the features from the fingerprint ridges i.e. fingerprint minutiae. Matching fingerprints using conventional CNN with only fingerprint images is challenging because different fingerprints, even from different subjects, illustrate similar ridge structure. In addition, conventional convolution and pooling operations can result in the loss of fingerprint texture details, which is likely to generate more false accept pairs during the test phase for the CNN-based approach. In order to address this challenge, we propose to extract specific feature, i.e. minutiae of fingerprint and incorporate it with ridge images as the inputs of our network.

In this work, we present a novel CNN-based approach for accurately matching contactless and respective contact-based fingerprints. We adopt the Siamese CNN structure to more accurately learn the similarities between the contactless and contact-based fingerprints from same subject. The proposed multi Siamese CNNs are trained not only using fingerprint ridge map but also incorporating the reliable fingerprint features, like minutiae and core point region.

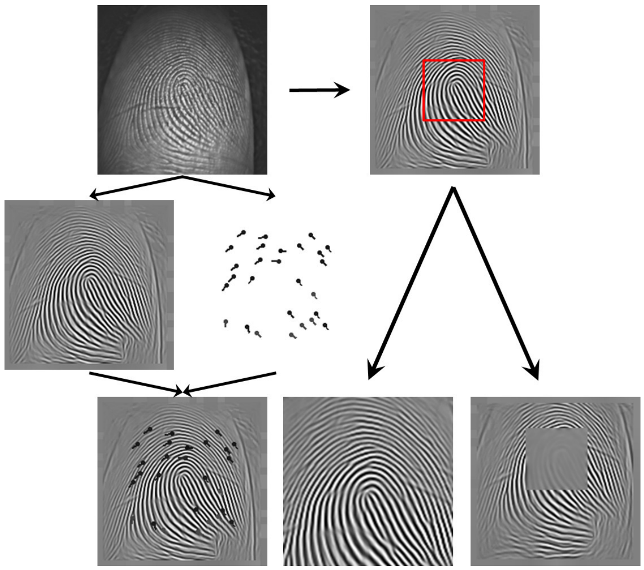

Figure 2: The process for generating input images of proposed multi Siamese networks. The first row shows contactless fingerprint image sample and respective ridge map with core point region. The second row shows fingerprint ridge map and respective minutiae maps. The last row shows the input for the network, i.e. combined fingerprint ridge map and minutiae map, fingerprint core point region and fingerprint with blurred core point region.

The feature vectors generated from three sub-networks are concatenated to form more robust fingerprint deep feature representation with rich descriptors for matching cross fingerprint images. The experimental results detailed in Section 3 of this paper achieve outperforming results over several popular CNN-based approaches and demonstrate significantly improvement over traditional minutiae-based cross fingerprint matching methods on a publicly available dataset.

2. Methodology

In this section, the proposed fingerprint cross-matching approach is detailed in different subsections. Firstly fingerprint image preprocessing approach is described. Secondly we introduce the approaches for generating multiple input pairs of proposed networks. Then the proposed framework for incorporating fingerprint ridge map, with respective minutiae feature and specific regions is elaborated. The architecture of the proposed network and the deep feature representation generation process are described in the last subsection.

2.1. Fingerprint Image Preprocessing

The conventional convolutional neural network approach on image classification and recognition problems highly relies on the quality of input images. It is essential for the success of cross-fingerprint matching that the similarities between input contactless and contact-based fingerprint image pairs is identified/localized. The contactless and contact-based fingerprint images acquired from different sensors are of different sizes. In addition, the different acquisition technologies make these two kinds of fingerprints, even from same subject, appear quite differently. Therefore, fingerprint preprocessing including ROI extraction and image enhancement is necessary, particularly when the size of training samples available is small. Firstly, the contact-based and contactless fingerprint images are subsampled and respectively cropped to normalize them to be of same scale and size. Secondly, in order to enhance the contrast of contactless fingerprint ridge-valley pattern, histogram equalization operation is utilized. Then the conventional fingerprint enhancement method [7] is incorporated to increase the similarity between the two kinds of fingerprints. Lastly the fingerprint ridge map is generated by applying adaptive histogram equalization filter [21]. Figure 2 illustrates contactless fingerprint image sample and its corresponding ridge map.

2.2. Fingerprint Ridge Map and Extracted Features

Minutiae-based fingerprint matching algorithms are widely used in different fingerprint identification systems and have shown better performance than image-based fingerprint matching methods. However, for the contactless to contact-based cross-fingerprint matching problem, because of the deformation of contact-based fingerprint and perspective distortion of contactless fingerprint, it's very difficult to *accurately* extract minutiae features from these two kinds of fingerprint. The missing or false minutiae can also easily result in false minutiae correspondences. On the other hand, convolution operation is a kind of spatial filter that works efficiently on images. Several references have also shown that the by incorporating hand-crafted features along with the original images can benefit the CNN-based image recognition/learning. Therefore, we firstly combine fingerprint ridge image with traditional minutiae as one input for our networks, to help the networks effectively learn the common features corresponding to the fingerprint images from two different sensors.

Fingerprint minutiae features are extracted using the method in [7]. It can be represented as $m = [x, y, \theta]$, where $[x, y]$ represents minutiae position and θ is the minutiae direction. Based on minutiae position and direction, we directly mark the minutiae on corresponding fingerprint ridge map. We represent each minutia using a solid circle and short line along the direction of fingerprint ridge. In such way, the response of the marked minutiae can help the network learn the texture details of the same subject (fingerprint) from different sensors. The contactless fingerprint

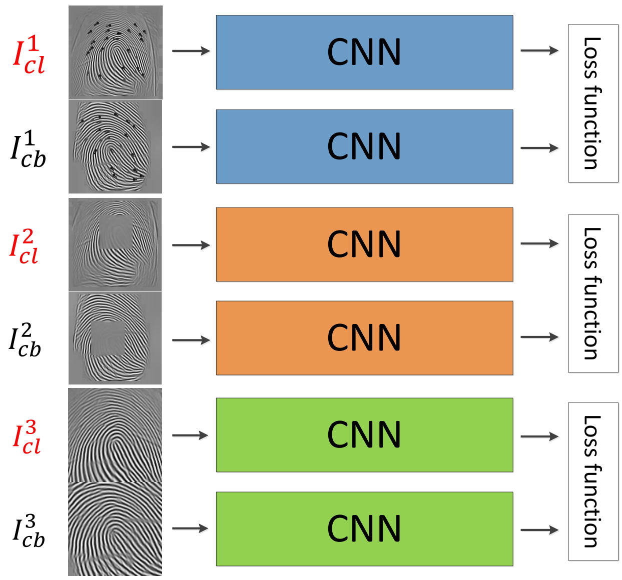

Figure 3: Architecture for proposed multi-Siamese convolutional neural network, where I_{cl} (red color) and I_{cb} (black color) represent contactless and contact-based fingerprint respectively. I^1 represents fingerprint ridge map, I^2 represents fingerprint ridge map with blurred core point region and I^3 represents fingerprint core point region of ridge map

image sample for the combined fingerprint ridge and minutiae map is shown in Figure 2.

In addition to using fingerprint ridge map and corresponding minutiae feature, the specific regions of the fingerprint ridge map based on fingerprint core point are also extracted for the input of proposed network. The specific regions include fingerprint core point region and blurred core point region. This operation can augment the size of training dataset which can help to avoid overfitting to some extent. More importantly, incorporating the specific regions can make the network learn relative texture details that may be lost due to the convolution and pooling operation on fingerprint ridge map.

Fingerprint core point is automatically detected using the method in [8]. For the fingerprints, which have two core points, we always select the point with smaller horizontal value as the core point. Based on the core point, we crop the $120 * 120$ pixels surrounding region as the core point region from the fingerprint ridge map ($192 * 192$ pixels). The ridge map with blurred core point region ($164 * 164$ pixels) is generated by applying Gaussian blur (filter size equals to 13 and sigma equals to 2) to the resized ridge map image on $60 * 90$ pixels region surrounding the core point position. The contactless fingerprint image samples for core point region and the ridge map with blurred core point region are illustrated in Figure 2. The process for generating multiple input pairs is also shown in this Figure 2.

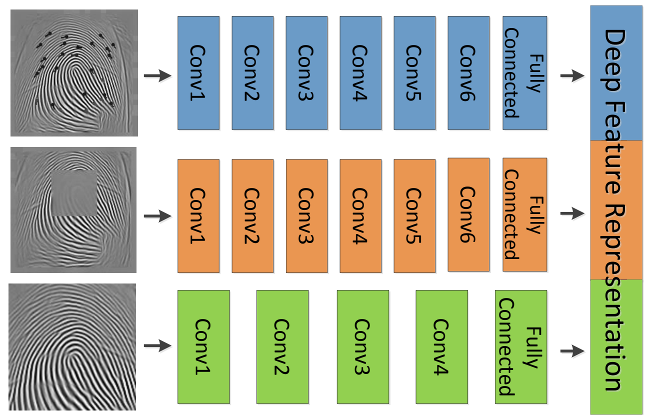

Figure 4: Architecture for each of the single CNN shown in *Figure 3* for contactless fingerprint deep feature representation generation.

2.3. Cross-Fingerprint Match Framework

2.3.1 Siamese Network

Siamese convolutional neural network [1] consists of two CNNs that share the same network structure and parameter weights. Usually pairs of images from same subject and different subjects are generated as the input of the Siamese network. The similarity metric is learned from the distance of each input pairs using a contrastive loss function. In comparison with conventional CNN, Siamese based CNNs have shown its superior performance on different problems like, face recognition [1, 26] and image patch recognition [34]. Especially for cross-domain problem like sketch retrieval [22, 31], using Siamese based CNN structure also achieve better performance than using conventional CNN. In addition, because of the high intra-class variation between the contact-based and contactless fingerprint images, conventional CNN structure may be not suitable for our particular case. Building on the successes of Siamese CNN and limitations of conventional CNN, we use Siamese based structure to train proposed network.

2.3.2 Three Sub-networks

In order to avail the advantages from the Siamese CNN architecture and existing hand-crafted fingerprint features, we investigate a framework consisting of three sub-networks to generate more robust feature representations for the cross fingerprint matching. The inputs of the three sub-networks are fingerprint ridge map with minutiae feature (sub-net1), ridge map with blurred core point region (sub-net2) and core point region of ridge map (sub-net3) respectively. Figure 3 illustrates the whole architecture of this network with the cross fingerprint image pairs same subject. Each subnetwork is a Siamese CNN including two single CNNs and

TABLE 1: Configurations of each single CNN for each of three sub-networks

	Sub-net1			Sub-net2			Sub-net3		
Operation	Filter Size	Output Num	Stride	Filter Size	Output Num	Stride	Filter Size	Output Num	Stride
Conv1	3	48	1	3	48	1	5	64	1
Max Pool1	3	-	2	3	-	2	3	-	2
Conv2	3	64	1	3	64	1	3	256	1
Max Pool2	3	-	2	3	-	2	3	-	2
Conv3	3	96	1	3	96	1	3	512	1
Max Pool3	3	-	2	3	-	2	3	-	2
Conv4	3	128	1	3	128	1	3	512	1
Max Pool4	3	-	2	3	-	2	3	-	2
Conv5	3	256	1	3	256	1	-	-	-
Max Pool5	3	-	2	3	-	1	-	-	-
Conv6	3	512	1	3	512	1	-	-	-
FC	-	1024	-	-	1024	-	-	1024	-

contact-based and contactless fingerprint images are the input pairs. For training ridge maps with minutiae and blurred core point region images, each single CNN includes six convolution layers and five pooling layers. After each convolution layer and last fully connected layer, we use ReLU as the activation function. For training core point region images, each single CNN includes four convolution layers and pooling layers. After the last fully connected layer, ReLU is used as the activation function. Each sub-network is trained separately during the train phase and the fully connected layers are concatenated to form the robust feature representation during the test phase. The sample architecture of each single CNN and deep feature representations generation process, for the contactless fingerprint, is shown in Figure 4. The configurations of each single CNN employed in each of the three sub-networks are detailed Table 1.

Contrastive loss function is mainly used in Siamese network. Let (I_{cl}, I_{cb}) be the input image pairs of the proposed network, which represents one contactless and one contact-based fingerprint image respectively. Let Y be the binary label of input images pairs, $Y = 0$ represents the input pairs from the same subject and $Y = 1$ represents that they are belong to different subjects. Then in the train phase, the positive and negative pair can be represented as $(I_{cl}, I_{cb}, Y = 0)$ and $(I_{cl}, I_{cb}, Y = 1)$. Let W represent the shared parameters learned from the network. The similarities between each two deep representations $R_W(I_{cl})$ and $R_W(I_{cb})$ from the network can be measured by computing their distance as follows,

$$D_W(I_{cl}, I_{cb}) = ||R_W(I_{cl}) - R_W(I_{cb}))|| \tag{1}$$

Then contrastive loss function can be defined as:

$$L(W) = \sum_{i=1}^{N} l((I_{cl}, I_{cb}, Y)^i, W) \tag{2}$$

$$\begin{aligned} l((I_{cl}, I_{cb}, Y)^i, W) = {} & (1 - Y)D_W(I_{cl}, I_{cb})_P \\ & + Ymax(\alpha - D_W(I_{cl}, I_{cb})_N, 0) \end{aligned} \tag{3}$$

where $(I_{cl}, I_{cb}, Y)^i$ is the $i - th$ training sample pair and P represents the positive pair and N represents the negative pair. The loss function enforces the distance between each two negative pair larger than margin α. During the training phase, this network aims to make the distance between positive pair small and push the negative pairs away. An improved contrastive loss function was developed and has been detailed in [29], which was used in our network.

2.3.3 Match Score using Deep Feature Representation

Several references [2, 26] have successfully incorporated the feature representations that are learned from the fully connected layer of the trained CNNs. In a similar way, we extracted fully connected layer from three sub-networks as the six feature vectors during the test phase. Let $f(\cdot)$ be the feature vector extracted from fully connected layer of proposed network. Then the six feature vectors of contactless

Figure 5: Contactless and corresponding contact-based fingerprint samples from same subject in the dataset [4]

(cl) and contact-based (cb) fingerprint pairs generated from three sub-networks can be represented as $(f(a_{cl}), f(a_{cb}))$, $(f(b_{cl}), f(b_{cb}))$ and $(f(c_{cl}), f(c_{cb}))$, where a represents the input fingerprint ridge map with minutiae, b represents the input ridge map with blurred core point region and c represents the input core point region of ridge map. For each of the test or unknown fingerprint image pair, we generated the corresponding enhanced ridge map with minutiae, core point region of ridge map and ridge map with blurred core point region. In order to generate robust feature representation, we concatenated these three feature vectors to form one fingerprint deep feature representation, which can be represented as $(f((a, b, c)_{cl}), f((a, b, c)_{cb}))$. Then for each test fingerprint image pair, the match score S can be directly computed from the following equation,

$$S = d(f((a, b, c)_{cl}), f((a, b, c)_{cb})) \tag{4}$$

where d represents the Euclidean distance between each two deep feature representations.

3. Experimental Evaluation

In this section, we detail experiments performed to evaluate the proposed approach for contactless to contact-based cross fingerprint matching. We compare the proposed method with several promising CNN-based approach and also with the conventional fingerprint matching method. Two common performance evaluation approaches for biometrics systems are to compute receiving operating characteristic (ROC) curve and equal error rate (EER). They are used in our experiments to ascertain the effectiveness of the proposed approach.

3.1. Dataset and Implementation Details

We used a publicly available database [4] for the performance evaluation. This dataset includes two-session images from 336 different clients (fingers/subjects). 160 clients have *two-session* samples, i.e. 12 contactless and corresponding contact-based fingerprint image samples and the other 176 clients only contains *one-session* 6 fingerprint images samples. The contactless and respective contact-based fingerprint image samples are illustrated in Figure 5. We split the dataset into training, testing and validation set. Each set contains the fingerprints from different fingers/subjects. The training dataset contains 3840 image samples, that is, 24 samples (12 contactless and 12 contact-based image samples) each for 160 clients with *two-session* samples. The other unseen 1920 fingerprint samples, that is, 12 samples each of first 160 clients with only *one-session* samples are selected as test data. The validation dataset contains unseen 192 image samples from the last 16 clients.

The contactless and contact-based fingerprint images from the dataset are cropped and resized into the same resolution for training. In the context of our problem, during contactless and contact-based fingerprint acquisition process, unintentional finger translation and rotation may result in fingerprint image rotation. Although using Siamese based network may avoid overfitting problem to some extent which caused by insufficient training data, we still applied data augmentation operation for fingerprint rotation problem and overfitting problem. Each fingerprint image sample is rotated by four different degrees. After this operation, the size of the dataset is increased five times respectively.

Our network is implemented based on the open-source framework Caffe [10]. The initial learning rate is set to 0.0001, the batch size is 100 and weight decay is 0.005. Mini-batch stochastic gradient descent method is used to minimize the loss functions. The processing time is measured on the PC with GPU NVIDIA 980Ti. Although we applied data augmentation operation, the overfitting problem is still observed during training process. Therefore, we used early stopping criteria. The training process of each sub-network is stopped after around 50k iterations when the loss for the validation data starts to increase. Overall training takes around 5 hours for each sub-network.

3.2. Experimental Results

The first set of experiments was performed to validate the effectiveness of the deep feature fusion approach, i.e. incorporating fingerprint ridge image with minutiae and specific region features. The verification experiments generated 5760 $(160 \times 6 \times 6)$ positive match scores and 915840 $(160 \times 6 \times 159 \times 6)$ negative match scores using dataset [4]. We computed the Euclidean distance between each two test fingerprints feature representations as the match score.

The experimental results for cross fingerprint matching using fingerprint ridge map with or without related minutiae map were illustrated in Figure 6. As can be observed from the ROC curve, incorporating fingerprint ridge image with minutiae features improved the performance for contactless to contact-based fingerprint cross matching. The EER de-

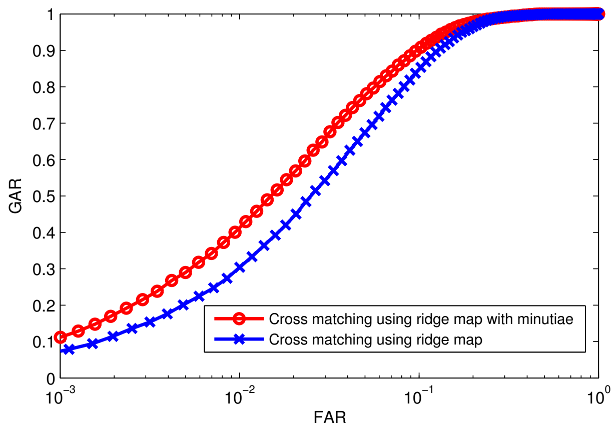

Figure 6: Comparison ROC curves for cross matching using fingerprint ridge map with minutiae and fingerprint ridge map only

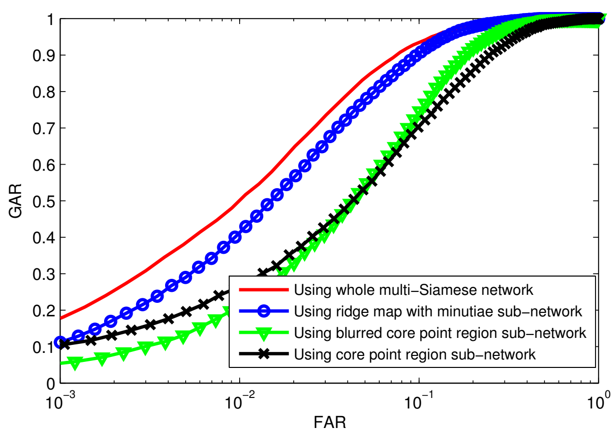

Figure 7: Comparison ROC curves for cross matching using whole multi-Siamese CNN and three sub-networks

creases from 11.65% to 9.89% when enhanced fingerprint ridge map with respective minutiae features are employed. The experimental results for cross-fingerprint matching by separately using three sub-networks and the whole network are illustrated in Figure 7 using respective ROC curves. Our results indicate that the deep fingerprint representations learned from the proposed multi-Siamese network are more robust/effective. The performance for the cross-fingerprint matching using the proposed multi-Siamese network framework is significantly better than those from individual sub-network. The EER is further decreased to 8.39% using the proposed approach.

3.3. Comparison with Other Competing Methods

We also conducted experiments to compare proposed method with several popular CNN-based methods including fine-tuning VGG network [25], 2-channel-based architecture [34] and triplet CNN [24] using fingerprint ridge map images with related minutiae feature. The comparative results were illustrated in Figure 8. To fine-tune the VGG network which uses conventional CNN structure, we combined contactless and contact-based fingerprint samples from same subject together and considered it as a classification problem. During the train phase, we found that it's difficult to achieve accurate classification due to the high intra-class variation between contactless and contact-based fingerprints. It also resulted in poor cross matching performance by fine-tuning VGG network. In [34], the authors trained the 2-channel networks with hinge loss function. However, in our case, we found that the 2-channel network couldn't converge with hinge loss function. Hence we replaced it with softmax loss function. In [13], the authors indicated that for patch matching problem, using triplet CNN can achieve better performance than Siamese CNN. However, for our cross fingerprint matching problem, it's difficult to randomly select the proper anchor for triplet network. This is plausible reason that the results using triplet CNN are worse than using the Siamese CNN. In summary, the experimental results from the ROC curves in Figure 8, indicate outperforming results over other popular CNN-based methods in the literature.

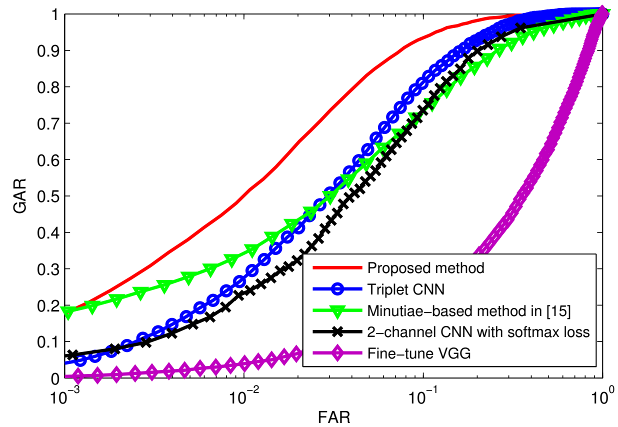

Figure 8: Comparison ROC curves for cross matching using proposed method and other methods

In addition to comparing our approach with other CNN-based methods, traditional minutiae-based fingerprint methods were also implemented to ascertain the effectiveness of our approach. A more recent approach in [15] provides comparative results with many popular minutiae-based matching methods and was also considered as candidate approach for the performance comparison. The comparative results using this approach are also presented in Figure 8. The EERs from the different methods are illustrated in Table 2. These comparative results also indicate that the proposed approach for the contactless to contact-

TABLE 2: Summary of performance using EER for different architectures/methods

	Proposed method	Sub-net1	Sub-net2	Sub-net3	Triplet CNN	2-channel CNN	Minutiae-based method [15]
Equal Error Rate	**8.39%**	9.89%	14.63%	16.95%	13.26%	15.42%	16.17%

based cross fingerprint matching can achieve outperforming results and validate the effectiveness of the proposed method.

4. Conclusions and Further Work

This paper has investigated a new approach for matching contactless to contact-based fingerprint images. Our approach is based on a multi-Siamese convolutional neural network whose architecture emphasis on generating deep feature representations for the robust matching. Fingerprint minutiae feature are extracted and combined with respective ridge map. The specific regions generated based on the fingerprint core point features are integrated into proposed network. More robust fingerprint deep feature representation learned from the proposed multi-Siamese network can effectively measure the similarities between contactless and respective contact-based fingerprint. Our approach can be regarded as feature level fusion of deep representations from multiple networks. The effectiveness of proposed approach is demonstrated from the comparative experimental results on a publicly available dataset. The comparative experimental results presented in this paper suggest that our approach outperforms several CNN-based methods and traditional minutiae-based methods for the cross fingerprint matching.

Despite the promising results there are several limitations. Firstly, the complexity of proposed framework is high due to training multiple CNNs. Secondly, although significant performance improvement is achieved by proposed approach the results are still below the expectations for the deployments. Thirdly, only one database is utilized for evaluation due to the lack of publicly available database. In addition, the size of the used database is not large enough for training deep CNN without overfitting problem. Therefore generating/using large size fingerprint databases is required to ensure the effectiveness of proposed approach. Future extension of this work should focus on developing more efficient framework to further enhance accuracy for the contactless to contact-based fingerprint matching.

Acknowledgment

This work is supported by General Research Fund from Research Grant Council of Hong Kong, project number PolyU 516913.

References

[1] S. Chopra, R. Hadsell, and Y. LeCun. Learning a similarity metric discriminatively, with application to face verification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 539–546. IEEE, 2005.

[2] Y. Cui, F. Zhou, Y. Lin, and S. Belongie. Fine-grained categorization and dataset bootstrapping using deep metric learning with humans in the loop. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1153–1162, 2016.

[3] L. Ericson, S. Shine, I. ManTech Advanced Systems International. Evaluation of contactless versus contact fingerprint data, phase 2 (version 1.1). *Report No. 249552, I. ManTech Advanced Systems International, DOJ Office of Justice Programs*, 2015.

[4] Weblink for downloading *PolyU Contactless to Contact-based Fingerprint Database*. `http://www.comp.polyu.edu.hk/~csajaykr/fingerprint.htm/`.

[5] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.

[6] Azimuth Inc. Evaluation of contact versus contactless fingerprint data (final report v2). *Report No.245146, I. ManTech Advanced Systems International, DOJ Office of Justice Programs*, 2014.

[7] A. Jain, L. Hong, and R. Bolle. On-line fingerprint verification. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, pages 302–314, 1997.

[8] A. K. Jain, S. Prabhakar, L. Hong, and S. Pankanti. Filterbank-based fingerprint matching. *Image Processing, IEEE Transactions on*, pages 846–859, 2000.

[9] H. Zhang, T. Xu, M. Elhoseiny, X. Huang, S. Zhang, A. Elgammal, and D. Metaxas. Spda-cnn: Unifying semantic part detection and abstraction for fine-grained recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1143–1152, 2016.

[10] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell. Caffe: Convolutional architecture for fast feature embedding. In *Proceedings of the 22nd ACM international conference on Multimedia*, pages 675–678. ACM, 2014.

[11] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1106–1114, 2012.

[12] A. Kumar and C. Kwong. Towards contactless, low-cost and accurate 3d fingerprint identification. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, pages 681–696, 2015.

[13] B. Kumar, G. Carneiro, and I. Reid. Learning local image descriptors with deep siamese and triplet convolutional networks by minimising global loss functions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5385–5394, 2016.

[14] Y. Chen, G. Parziale, E. Diaz-Santana, and A. K. Jain. *3d touchless fingerprints: compatibility with legacy rolled images*. In *Biometric Consortium Conference, 2006 Biometrics Symposium: Special Session on Research*, pages 1–6, 2006.

[15] C. Lin and A. Kumar. Improving cross sensor interoperability for fingerprint identification. In *International Conference on Pattern Recognition (ICPR)*, pages 943–948, 2016.

[16] T.-Y. Lin, A. RoyChowdhury, and S. Maji. Bilinear cnn models for fine-grained visual recognition. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1449–1457, 2015.

[17] I. Masi, S. Rawls, G. Medioni, and P. Natarajan. Pose-aware face recognition in the wild. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4838–4846, 2016.

[18] MorphoWave. `http://www.morpho.com/`. Accessed Jan, 2017.

[19] R. F. Nogueira, R. de Alencar Lotufo, and R. C. Machado. Fingerprint liveness detection using convolutional neural networks. *IEEE Transactions on Information Forensics and Security*, pages 1206–1213, 2016.

[20] G. Parziale and Y. Chen. Advanced technologies for touchless fingerprint recognition. In *Handbook of Remote Biometrics*, pages 83–109. Springer, 2009.

[21] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld. Adaptive histogram equalization and its variations. *Computer vision, graphics, and image processing*, pages 355–368, 1987.

[22] Y. Qi, Y.-Z. Song, H. Zhang, and J. Liu. Sketch-based image retrieval via siamese convolutional neural network. In *Image Processing (ICIP), 2016 IEEE International Conference on*, pages 2460–2464. IEEE, 2016.

[23] A. Ross and A. Jain. *Biometric sensor interoperability: A case study in fingerprints*. Springer, 2004.

[24] F. Schroff, D. Kalenichenko, and J. Philbin. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 815–823, 2015.

[25] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[26] Y. Sun, Y. Chen, X. Wang, and X. Tang. Deep learning face representation by joint identification-verification. In *Advances in neural information processing systems*, pages 1988–1996, 2014.

[27] Y. Sun, X. Wang, and X. Tang. Deeply learned face representations are sparse, selective, and robust. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2892–2900, 2015.

[28] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–9, 2015.

[29] C. Lin and A. Kumar, "A preliminary study on matching contactless to contact-based fingerprint images using deep learning," *Technical Report No. COMP-K-24*, Jul. 2017.

[30] TBS. `http://www.tbs-biometrics.com/`. Accessed Feb, 2017.

[31] F. Wang, L. Kang, and Y. Li. Sketch-based 3d shape retrieval using convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1875–1883, 2015.

[32] J. Wang, Y. Song, T. Leung, C. Rosenberg, J. Wang, J. Philbin, B. Chen, and Y. Wu. Learning fine-grained image similarity with deep ranking. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1386–1393, 2014.

[33] W. Yani, W. Zhendong, Z. Jianwu, and C. Hongli. A robust damaged fingerprint identification algorithm based on deep learning. In *Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), 2016 IEEE*, pages 1048–1052. IEEE, 2016.

[34] S. Zagoruyko and N. Komodakis. Learning to compare image patches via convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4353–4361, 2015.