A. Brömme, C. Busch, N. Damer, A. Dantcheva, M. Gomez-Barrero, K. Raja,
C. Rathgeb, A. Sequeira, and A. Uhl (Eds.): BIOSIG 2021,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2021

Vein Enhancement with Deep Auto-Encoders to improve Finger Vein Recognition

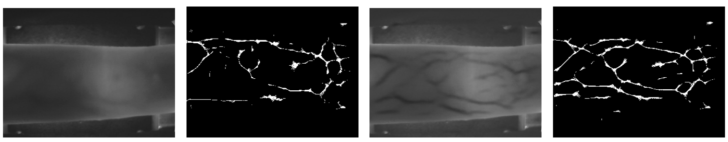
Victor Bros[1], Ketan Kotwal[1], and Sébastien Marcel[1]

Abstract: The field of Vascular Biometric Recognition has drawn a lot of attention recently with the emergence of new computer vision techniques. The different methods using Deep Learning involve a new understanding of deeper features from the vascular network. The specific architecture of the veins needs complex model capable of comprehending the vascular pattern. In this paper, we present an image enhancement method using Deep Convolutional Neural Network. For this task, a residual convolutional auto-encoder architecture has been trained in a supervised way to enhance the vein patterns in near-infrared images. The method has been evaluated on several databases with promising results on the UTFVP database as a main result. In including the model as a preprocessing in the biometric pipelines of recognition for finger vein patterns, the error rate has been reduced from 2.1% to 1.0%.

Keywords: Finger Vein Recognition, Deep Residual Convolutional Auto-encoder, Vein Enhancement

1 Introduction

Automatic biometric recognition has become a reliable technology to perform identification and verification of an individual. This domain has a wide range of applications in everyday life to assess the identity of an individual or to attribute one to them. Since the vascular patterns are believed to be unique from one person to another, they are well-suited for the task of verification. For the convenience of regular use, the veins of the most accessible parts of the body, such as the hands (palm, finger, wrist, etc.), are preferred as a recognition modality.



(a) (b) (c) (d)

Fig. 1: Samples of finger vein images from the SDUMLA dataset. From left to right : (a) shows the original vein image, and (b) depicts its maximum curvature (MC) without enhancement. The resulting image after vein enhancement is shown in (c); while (d) depicts the MC obtained from it.

Usually biometric verification experiments consist of the following stages. The sensor acquires the biometric characteristic. The feature extractor generates the feature descriptor from the input presentation. The feature descriptor is compared against the precomputed templates to obtain the matching score. Despite several finger vein (FV) recognition

[1] Biometrics Security and Privacy Group, Idiap Research Institute, Martigny, Switzerland, {vbros, ketan.kotwal, sebastien.marcel}@idiap.ch

methods being developed [Hu10, KZ11], it may be noted that their performance is often strongly correlated to the quality of the input presentations captured in the NIR spectra. Due to the nature of the FV structure (under the skin with many flesh artefacts), the quality of the captured presentations may often be poor, especially with lack of contrast. For instance, Figure 1 (a) and (b) show the input FV presentation and its corresponding FV pattern extracted using Maximum Curvature. Since the input presentation exhibits low contrast, it is indeed a challenging task to identify relatively darker vein patterns. As we may observe from Figure 1 (b), the FV extraction algorithm may miss several smaller veins, which could significantly enhance the discriminative capabilities of the FV recognition system. An efficient mechanism for preprocessing of FV presentations may result in the extraction of subtle vein patterns, and thus, obtaining a robust feature for the subsequent recognition pipeline.

In this work, we propose a deep learning-based preprocessing method that, in particular, enhances the vein patterns acquired in the NIR spectra. The objective of our method is to improve the separation between the background and the vascular networks in FV images. We use a Deep Convolutional Neural Network (DCNN) Auto-Encoder, more particularly a residual convolutional auto-encoder (RCAE), that can function as a preprocessor in the recognition pipeline before the feature extractor stage. Our model has been trained to reconstruct the enhanced versions of the input FV presentations with darker vein patterns. Using these enhanced presentations, the feature extractor has been able to identify even subtle FV patterns, that the standalone Maximum Curvature was unable to extract from unprocessed presentations. Figure 1 (c)-(d) depict the output of the proposed method and the extracted FV pattern, respectively. When compared against Figures 1 (a-b), that represent the equivalent results without any preprocessing; a simple visual inspection can demonstrate the improvement in the quality of presentations in terms of higher contrast and better separation between vein patterns from the rest of the content. In the context of FV, an objective evaluation of quality of the presentation is an open problem; and there is no universally accepted measure for the same. Therefore, in this work, we measure efficacy of the proposed FV enhancement method in an indirect manner, *i.e.,* through the gain observed in overall recognition accuracy after incorporating the proposed preprocessing mechanism.
Although the present work has been conducted in the context of finger vein recognition, it is possible to imagine other applications of such enhancement across a variety of domains. Some examples could be improvement in the contouring of a vascular network, a correction of low quality captured images which are quite expensive tasks in medical imaging.

The remainder of the paper is structured as follows: In the Section 2, we briefly describe the relevant work and the related work on enhancement of finger vein patterns. The deep convolutional auto-encoder model that reconstructs the image with enhanced vein patterns is described in Section 3. We discuss the details of the experiments based on the proposed method in Section 4. Section 5 summarizes the conclusions.

2 Related Work

To the best of our knowledge, the proposed work is the first attempt to use a DCNN Auto-Encoder to enhance FV patterns aimed towards improving the subsequent feature extraction. Therefore, in this section, we present a brief overview of commonly used building blocks of the FV recognition pipeline. We also discuss a few similar image enhancement methodologies- that have been developed for different end applications. Inspired by the idea of preprocessing images for their specific use, the method aims to enhance veins in NIR images of fingers in the context of biometric recognition. This method has been incorporated into verification pipelines, whose algorithms have been previously chosen for their reliability and performance. These are canonical algorithms that will be the baseline for the comparison with and without enhancement. The extraction of vein patterns is done by Maximum Curvature [MNM07]. The principle is to compute a binary map of the vascular patterns by calculating the centers of the veins by their intensity profiles and connecting them via a filter operation. Then the comparator, designed by the same team, is the Miura Matching [MNM04]. This method calculates a similarity score between two binary patterns by retrieving their maximum superposition score with possible displacements represented by a sliding window.

Within the pipelines, other preprocessing have been tested to improve the image quality. The first idea was to center the image on the finger, either by performing a crop of the region of interest (ROI) [YS12], or by including as much as possible the minutiae of the veins [LLP09]. Also since the patterns have a third dimension component, the second preprocessing reduces the distortion of the capture in adding a normalization for the position [Hu10]. Hence for the baseline, the whole sequence has been considered [Pe13].

With the idea of learning vein patterns by deep neural networks, other works have shown the use of convolutional networks for FV experiments. A study was able to use a convolution neural network as a comparator between two patterns with great success [Li17]. It was then questioned if an auto-encoder can compress the patterns in a reduced bottleneck embedding which is then compared with a Support Vector Machine classifier for the verification task [HY20]. From these experiments, it seems possible that network models have the ability to retrieve latent information from vascular patterns, in particular those mentioned above. The combination of an auto-encoder with convolutional layers has been tested for vascular patterns in the fundus of the eye. In [Li20], Li *et al* have proposed a neural network, NuI-Go, that aims to reduce the non-uniform illumination of the images of the eye. Here, they generated a dataset of retina vessel images with a synthetic degradation of the illumination. Their method is built on a deep residual convolutional auto-encoder to perform the reconstruction from the degraded image towards the original high quality image of the fundus.

3 Proposed Method

For the enhancement of FV images, we design an RCAE consisting of the encoder and decoder blocks that are linked through a residual connection. The encoder accepts an input FV image, and the decoder attempts to reconstruct its enhanced version. In the next subsection, we first outline the architecture of our RCAE network, and then provide details of the training process.

3.1 RCAE Architecture

For FV enhancement, we need to design a network structure that selectively identifies specific patterns in the input, and enhances their representation, in terms of good contrast and sharpness, during the reconstruction phase. Recently, Li *et al* have demonstrated the use of convolutional auto-encoder with a non-local unit for enhancement of the fundus of the eye image [Li20]. Although their use-case is quite different from ours, the quality of their results indeed suggests that a convolutional AE can be a highly effective method for enhancement of finer structures in the image.

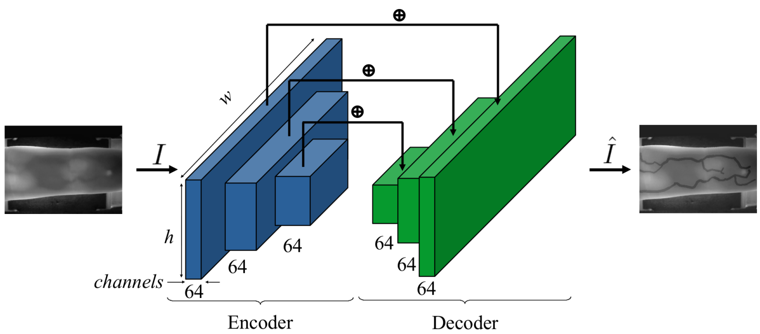


Fig. 2: The proposed architecture, a Residual Convolutional Auto-Encoder (RCAE), for the vein enhancement.

Following an extensive study, we design a RCAE network for the enhancement of FV aimed at improved detection of vein patterns. Figure 2 shows the schematic of the proposed RCAE. The encoder of our RCAE consists of 3 blocks: each of which includes a convolution, activation, and normalization. These blocks are connected to each other in sequence through pooling operation- thereby reducing the spatial dimensions of the effective input at every stage. On the decoder side, we have a succession of 3 decoding blocks: each consisting of a convolution and transposed convolution filters, along with activation and normalization layers. The network includes a residual transmission across every pair of blocks in encoder-decoder. This results into a differential component that learns the difference between the input and the target across layers. In the proposed RCAE, we have performed max-pooling, across the encoder, over the window of 2×2, and retained the stride of 2 for transposed convolutions in the decoder. The ReLU (Rectified Linear Unit) has been chosen as the activation operator; and each block is interspersed with a batch normalization layer to reduce the dependence on the training dataset and help the generalization [IS15].

The size of convolutional kernel is an important factor that determines the (effective) receptive field of the input and later layers of the deep network. Since the width or thickness of the FV patterns varies within a given presentation, the optimal size of convolutional kernel may not be easily decided. While the filters with relatively larger size are capable of learning the spatial relationship across somewhat distant pixels in the image; the smaller filters focus on encoding the features in local patches of the input. To explore the effects of kernel size of convolutional filters, we design two variants of the RCAE: (a) *Model 1*: with a constant kernel shape of 3×3 for each convolutional layer; and (b) *Model 2*: where the dimensions of convolutional kernel have been gradually decreased from 9×9 to 3×3 in the encoder, whereas the decoder layers observe a gradual increase in the kernel sizes. Also the number of channels inside the network have been drastically increased to 64 to introduce the complexity needed for the task.

3.2 Training Procedure

Generation of targets: The first step in the training of the RCAE is the generation of the targets, *i.e.,* synthetically generated enhanced FV images that act as the reference outputs during training. Our training datasets have been manually annotated for vein patterns in the form of binary masks. We consider the enhanced image (target) as the linear combination of actual input image and the vein-annotated binary mask. If $\mathbf{x}$ is the input FV presentation with $\mathbf{h}$ as the binary mask depicting vein structure, then the target image, $\mathbf{y}$, is obtained as: $\mathbf{y} = \alpha\mathbf{x} + (1-\alpha)\mathbf{h}$, where α refers to the fixed weight parameter. It should be noted that $\mathbf{x}$, $\mathbf{h}$, and $\mathbf{y}$ have the same dimensions. Figure 3 shows the example of input presentation and its mask, along with the generated target.

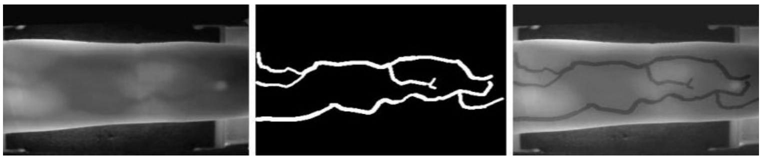


Fig. 3: Generation of target presentations: On the left, the original FV image; in the middle the binary annotation; and on the right the target image.

Data Augmentation: To deal with a possible lack of training data and generalization, we have incorporated data augmentation strategies during the training process of the RCAE model. We have considered the following four strategies for data augmentation: (a) random horizontal flip with probability $p = 50\%$, (b) random rotation with maximum angle $\theta = 2.5^\circ$, (c) random translation with maximum distance $d = 5\%$, and (d) random shear with maximum degrees $x = 5^\circ$.

Since the dimensions of the FV presentations across datasets may differ, it is necessary to resize the input to a fixed size (as determined during the training process). In our experiments, we have fixed the input size to 320×240 pixels. If the input FV presentation has different size, then it is re-scaled, in anti-aliased manner, after the enhancements to yield its original dimensions.

Loss Function: Since the FV images are quite distant from usual images in most common datasets, we have chosen a Mean Squared Error (MSE) loss ($\mathscr{L}_{\text{mse}}$) for the training RCAE. For the reconstructed image $\hat{\mathbf{y}}$, and the target (reference) image $\mathbf{y}$, the MSE loss is computed as $\mathscr{L}_{\text{mse}} = \frac{1}{N}(\mathbf{y} - \hat{\mathbf{y}})^\intercal.(\mathbf{y} - \hat{\mathbf{y}})$; where N is the number of pixels in the image $\mathbf{y}$.

For training the RCAE, we have chosen the Adam optimizer [KB17] with a learning rate of 0.001. In the beginning, the weights were initialized by random values normalized on a Gaussian centered around 0 with a bias of 0.05.

4 Experiments

We have implemented the experiments using PyTorch and Bob[2] frameworks with a focus on reproducible research. The python code and protocols to reproduce the experimental results are available publicly[3].

4.1 Datasets and Protocols

We demonstrate the efficacy of the proposed vein-enhancement RCAE on two publicly available FV datasets: UTFVP [Ro18] and SDUMLA [Lu13].[4]
The SDUMLA dataset consists of 634 masks, of medium quality FV presentations (320×240), from 636 identities (106 persons) collected in 6 sessions. The UTFVP dataset has 389 masks for high quality FV images with dimensions of (672×380).
As part of the verification experiments, it was necessary to define protocols for the use of the data in order to limit the correlation between the different sets for each database. In our experimental protocols, we ensure that the identities from training, validation, and testing of the RCAE do not overlap. We omit the details of protocols due to brevity of the space, but they can be obtained from the code repository.

4.2 Metrics for evaluation

Since the efficacy of the RCAE is measured through the FV recognition experiment, we have considered the False Match Rate (FMR) and False Non-Match Rate (FMNR) as the performance metrics. For performance comparison as well as for selection of score threshold, we have chosen the Half Total Error Rate (HTER) criteria which is computed as $HTER = 0.5 \times (FMR + FNMR)$. To evaluate the performance through the distributions of scores, the genuine (same identity presented) and the imposter (different identity), we have employed statistical tests as well. A *Cohen's d* test was used to measure the distance between the two empirical distributions to observe the impact of the preprocessing. This distance has been used as a comparison of the differentiation ability of the FV verification system. A higher value of the distance indicates a better separability of both distributions.

4.3 Experimental Results

Table 1 provides the results of verification experiments on the two datasets. For SDUMLA, a significant impact of the preprocessing can be seen through the performance of the verification system. For both models, the HTER has reduced by nearly 20%, and the *Cohen's distance* is higher by nearly 10% as compared to the corresponding values without the RCAE preprocessing. This second experiment was conducted to show the generalisation

[2] `https://www.idiap.ch/software/bob`
[3] `https://gitlab.idiap.ch/bob/bob.paper.biosig2021_deep_vein_enhancement`
[4] The vein annotations for both datasets were provided by the University of Salzburg.

		SDUMLA		UTFVP	
Model	Statistical test	development set	evaluation set	development set	evaluation set
w/o	HTER	15.4%	14.2%	0.2%	2.1%
	Cohen's d	5.0	6.8	15.5	15.2
Model 1	HTER	12.0%	9.8%	0.2%	1.3%
	Cohen's d	5.7	7.6	15.7	16.7
Model 2	HTER	12.5%	10.1%	0.2%	1.0%
	Cohen's d	5.6	7.6	15.7	16.2

Tab. 1: Results of the verification experiment on SDUMLA dataset and UTFVP dataset.

of the model to other databases, which means other sensors and other image quality. Similarly, the HTER on the evaluation set is lower for both RCAE models for the UTFVP dataset. The Cohen's d test also shows that the impact of the matching algorithm is higher with a preprocessing since the distance is higher for both models than without preprocessing. Figure 4 also highlights the overall improvement brought by the models for the verification on UTFVP, with lower *FNMR* at all *FMR* on the evaluation set. It may be inferred that the preprocessing has been able to generalize to other databases with success.

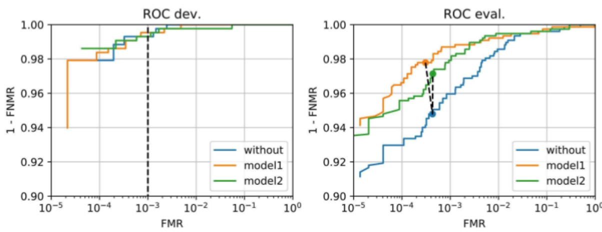


Fig. 4: ROC curves of the experiments on UTFVP: without RCAE preprocessing, model 1, and model 2.

5 Conclusion

In this work, we have proposed an enhancement method for FV images captured in NIR spectra. We have developed an RCAE model that can be integrated as the preprocessor into a biometric recognition pipeline. The purpose of the RCAE is to learn the prominent as well as subtle vein patterns in the image, and improve the quality of presentation, in terms of a better contrast. We have demonstrated that with the proposed preprocessing, the overall accuracy of the FV recognition has increased, as well as the separation between the distributions of recognition scores of genuine and imposter identities has also improved. For two publicly available FV datasets, our method has resulted in nearly 20% reduction in the average error in recognition.

The proposed RCAE enhances major FV structures, and also occasionally identifies the subtle vein patterns that might have been missed by human annotators. However, in some examples, it may lead to generate spurious vein patterns. We are presently working on improving the accuracy of enhanced vein patterns. We are also working on extending the application of RCAE beyond preprocessing tasks.

6 Acknowledgement

This work was supported by the Swiss Center for Biometrics Research & Testing, and the InnoSuisse project CANDY. Also this work was made possible thanks to Pr. Andreas Uhl from Salzburg University and his team for providing the vein annotations required for our experiments.

References

[Hu10] Huang, Beining; Dai, Yanggang; Li, Rongfeng; Tang, Darun; Li, Wenxin: Finger-vein authentication based on wide line detector and pattern normalization. pp. 1269–1272, Aug. 2010.

[HY20] Hou, B.; Yan, R.: Convolutional Autoencoder Model for Finger-Vein Verification. IEEE Transactions on Instrumentation and Measurement, 69(5):2067–2074, 2020.

[IS15] Ioffe, Sergey; Szegedy, Christian: Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. pp. 448–456, 2015.

[KB17] Kingma, Diederik P.; Ba, Jimmy: Adam: A Method for Stochastic Optimization. 2017.

[KZ11] Kumar, Ajay; Zhou, Yingbo: Human identification using finger images. IEEE Transactions on Image Processing, 21(4):2228–2244, 2011.

[Li17] Liu, W.; Li, W.; Sun, L.; Zhang, L.; Chen, P.: Finger vein recognition based on deep learning. pp. 205–210, 2017.

[Li20] Li, Chongyi; Fu, Huazhu; Cong, Runmin; Li, Zechao; Xu, Qianqian: NuI-Go: Recursive Non-Local Encoder-Decoder Network for Retinal Image Non-Uniform Illumination Removal. 2020.

[LLP09] Lee, Eui Chul; Lee, Hyeon Chang; Park, Kang Ryoung: Finger vein recognition using minutia-based alignment and local binary pattern-based feature extraction. International Journal of Imaging Systems and Technology, 19(3):179–186, 2009.

[Lu13] Lu, Y.; Xie, S. J.; Yoon, S.; Wang, Z.; Park, D. S.: An available database for the research of finger vein recognition. 01:410–415, 2013.

[MNM04] Miura, Naoto; Nagasaka, Akio; Miyatake, Takafumi: Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification. Machine Vision and Applications, 15:194–203, 10 2004.

[MNM07] Miura, Naoto; Nagasaka, Akio; Miyatake, Takafumi: Extraction of Finger-Vein Patterns Using Maximum Curvature Points in Image Profiles. IEICE - Trans. Inf. Syst., E90-D(8):1185–1194, August 2007.

[Pe13] Peng, Jialiang; Li, Qiong; Wang, Ning; Abd El-Latif, Ahmed; Niu, Xiamu: An Effective Preprocessing Method for Finger Vein Recognition. Proc SPIE, 8878:08–, 07 2013.

[Ro18] Rozendal, S.P.: Redesign of a finger vein scanner. February 2018.

[YS12] Yang, Jinfeng; Shi, Yihua: Finger–vein ROI localization and vein ridge enhancement. Pattern Recognition Letters, 33(12):1569–1579, 2012.