Optimization of A Patch-Based Finger Verification with a Convolutional Neural Network

Tugce Arican
Computer Science, University of Twente
Supervised by Raymond Veldhuis and Luuk Spreeuwers

***Abstract*—Finger veins are accepted as unique for each person, and since finger veins are below skin, they are more resistant to forgery. In this paper, a patch-based approach using a convolutional neural network is explored. The patch-based approach increases the number of labeled data, and helps against brightness variations, yet, at the same time, it introduces its own issues such as determining the patch properties, combining the patches, and registration of the image pairs. This research proposes an optimisation to the patch based finger vein verification approach by addressing these issues. The patch-based system has achieved 0.3% of equal error rate and 0.999 area under the curve on UTFVP and 6.6% of equal error rate and 0.969 area under the curve on SDUMLA-HMT after proposed optimisations. Even though the results are far from the state-of-the-art performance, the improvement indicates the potential of the proposed system.**

I. Introduction

Vascular patterns is of interest because they could be more resistant to external factor such as scars or aging, and forgery due to being below skin. Therefore, finger vein patterns could have advantages over other biometrics such as face or finger prints.

An identification system using vascular patterns could perform image verification by comparing a probe image with a registered one. If a matching score computed in this comparison exceeds a threshold, the identity is verified, rejected otherwise. Images belonging to the same subject are called as genuine pairs. If they belong to different subjects, then the pairs are said to be imposter.

Several approaches have been proposed to design a finger vein human identification system. Conventional methods are one of them which mainly utilises manually extracted features and matching distance. Line tracking [1], or cross-sectional areas [2], [3] have been proposed to extract the vein pattern. Gabor filters [4], Local Binary Patterns(LBP) [5] also used as textural features. Beside the whole vein pattern or textural information, end or bifurcation points [28], [29] are also used. These methods generally achieve high recognition accuracy; however, since they are based on manually extracted features, these approaches could be tailored to the dataset or the problem itself. Personalised Best Patch Map(PBPM) [33] and localised sub-regions [34] are also proposed to finger vein verification to achieve a system robust to partial-distortions.

There is a growing body of literature that recognises Machine learning based approaches in finger vein verification. These approaches rely on machine learning methods, such as neural networks, fuzzy logic, for the final decision. Wu and Liu showed machine learning based approaches could achieve high accuracy in finger vein pattern identification by using Support Vector Machines(SVM) [26] and Adaptive Neuro Fuzzy System(ANFIS) [25].

As a machine learning approach, Convolutional Neural Networks(CNNs) are proposed in finger vein human identification. [8], [14], [15], and [16] have achieved promising recognition rates indicating the potential of the CNN based approaches. [17] proposed to use difference image rather than the whole image in order to reduce the complexity of the CNN. Different than the previous ones [35] used a CNN to extract the finger vein structure by assigning labels for pixels as foreground and background.

A patch-based approach using a CNN has been proposed by [11] aiming to achieve a robust finger vein verification system to brightness variations. Rather than scoring the whole image, the network is fed by small square regions, called patches. After scoring, the individual patch scores are combined to an image score for the final decision. The proposed method achieved promising results showing the feasibility of the patch based approach with a CNN. A diagram of the system is shown in Figure1.

However, the patch-based approach in [11] has some issues which could prevent the approach from achieving its optimal. First of all, the use of patch properties such as patch size, shape, and overlapping were not investigated. It is likely that an improper patch could degrade the performance by dividing junctions unintentionally. Secondly, in [11], two fusion methods were explored under uniformity assumption. As not all the patches involve the same information, the uniformity assumption may not reflect the expected results. Finally, the effects of displacements on x-axis on the registration accuracy was not examined. The coarse registration used in [11] uses finger edges for image registration, and it could fail such a case since these kind of displacements do not change the finger edges much. This paper, proposes solutions to these issues by investigating different patch sizes and shapes, also the overlapping patches. Moreover, the fusion method considering the differences among patches and the registration approach which could consider the displacement in the x-axis have been implemented.

Considering the literature and the existing system this work investigates the following cases.

Research Question 1: What is the proper patch size and

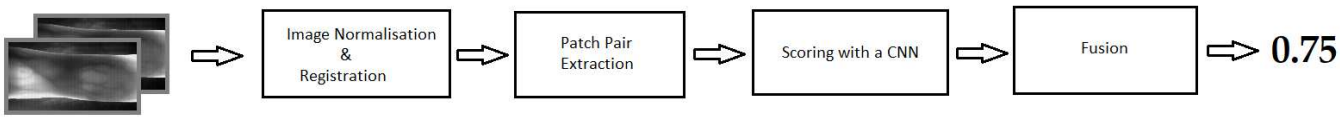


Fig. 1: Block diagram of the system

shape for the existing system?

Research Question 2: How could overlapping patches influence the verification?

Research Question 3: How could the contribution of each patch be determined with a computationally less complex fusion method?

Research Question 4: How could the displacements, which are ignored by the existing registration method, be taken into account?

This paper is organized as follows. Chapter II gives a brief overview about what have been done about finger vein biometrics until now. Chapter III explains the patch-based finger verification system in a detailed way and the methods used in this paper. Chapter IV presents the results achieved. Chapter V discusses the findings. Finally, Chapter VI closes this paper with a conclusion and future work.

II. Related Work

Various studies have indicated the patch-based approaches could benefit in solving several issues. For instance, [20] applied a patch based approach aiming to reduce the computational complexity of cancer image classification with a CNN. Since the input image has very high resolution, e.g. gigapixels, these patches help to reduce the resolution of the input, therefore the complexity of the CNN. Moreover, the patch based approach allowed to select only the relevant patches. [35] applied the patch-based CNN approach to extract finger vein patterns from raw images. The patches are centered on pixels. Then, a CNN assigns a probability of being a foreground pixel to the corresponding patch. The authors were able to achieve significant improvements on two public dataset in terms of finger vein verification accuracy. [11] utilised a patch-based CNN to finger vein verification aiming to achieve more robust verification system to brightness changes. The results achieved indicates the feasibility of the patch-based CNN approach to finger vein verification.

While they provide many opportunities, the patch properties are crucial for the patch-based approaches. [36] proposed a patch-based approach with Collaborative Representation based Classification(CRC) to face recognition in aiming to increase training samples, and different patch sizes have been investigated n this research. The obtained results indicated that the recognition performance is dependent on the selected patch size. [37] argued that not only the patch size but also the shape of the patch affect the performance. The authors proposed to use superpixels instead of fixed shape patches. The results achieved showed the importance of the selected patch shape. Overlapping patches approach used in [36], [22], and [21] achieved promising improvements indicating that overlapping patches could provide improvements on the performance.

In patch-based systems, the patches are scored individually, therefore, these scores must be fused to an image score. Since each patch could carry different information, a fusion operation should consider these differences. However, such a fusion is computationally expensive. [10] showed such a fusion is possible with less computational complexity. The authors applied a patch-based approach to face recognition aiming to achieve a robust system to brightness variations and facial expressions. They proposed a fusion method which determines a threshold for each patch by using only one parameter, called False Acceptance Rate (FAR) value. The FAR value is set at the beginning, and it is assumed that patches having poor scores could not contribute much to the final score since the FAR value will set a high threshold for those patches. The promising results they achieved in face recognition indicate the feasibility of the proposed approach.

In a patch-based approach, patch pairs are extracted according to their relative locations. Therefore, image registration has an influence on the overall performance. Registration is generally done based on the physical properties, e.g. edges, or reference points, e.g. landmarks. [30] and [13] showed that a better registration accuracy is possible with a matching score based approach. In [30], a performance metric computed from face recognition similarity scores were attempted to be maximized among a set of alignment candidate. One of the alignment candidates reaching the maximum performance metric has been selected as the aligned image. Similarly, [13] used an iterative method for image registration utilising a matching score based approach. The authors searched a set of geometric translations. One of the geometric translations minimising the matching score has been accepted as the registration parameters.

III. The Existing System and Methodology

A. Patch-based Finger Vein Verification

The finger vein verification system proposed by [11] consists of 5 steps, namely image registration, patch extraction, scoring, fusion, and decision. Figure 1 indicates the block diagram of the existing system.

Image pairs are registered based on utilising Iterative Closest Point (ICP) [12] algorithm. ICP uses the finger edges in order to align two finger vein images. Later, the center line of the finger images is used to correct the orientation of the fingers. After registration, 31 pixels square patch pairs are extracted only from the finger region. A CNN is fed by these patch pairs, and outputs a matching score for each. These

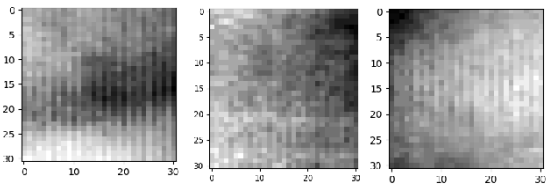


Fig. 2: 31-pixel patch samples from UTFVP

patch pair scores are fused to an image score. [11] compared two fusion methods namely decision and score level fusions. Finally, the fused scores are compared against a threshold for the final decision.

B. Methodology

Following sections describes the proposed solutions to the issues found in the existing systems and explains how to apply them.

1) Patch Size: In [11], it is stated that the maximum width of finger veins are approximately 20 pixels. Therefore, a vein might occupy a large area in a 31 pixels patch such that the network could not learn much from it. A visual inspection of the patches revealed that veins are ambiguous or not visible in 31 x 31 patches (Figure 2). In this research, the larger patch sizes than 31-pixel were investigated, The patch sizes were selected as 49, 57, 63, 69, 75, 82, and 88 pixels.

2) Patch Shape: Finger veins lay horizontally. Thus, square patch shape might not be the optimal choice for the proposed system. Rectangular patches could capture the horizontal vein structure better than square ones. Height of the patch was fixed to the best performing patch size. Table **??** presents the patch widths used in experiments.

3) Overlapping: Overlapping patches are another aspect of patch-based systems. Overlapping patches could help to catch some vein structures which cannot be seen with non-overlapping patches. Moreover, the overlapping increases the number of labelled data by providing more variations about the veins, which help the network to learn the vein structures better.

The overlapping patches were extracted in the same way described in [11], except a smaller stride than the patch size has been used in both height and width. The smaller stride leads more overlap, hence more similarity among adjacent patches. The strides used are presented in Experiment 3.

4) Fixed-Far Voting Fusion: A non-uniform voting fusion method is formulated by 1 and 2. LR_i defines the score of the system *i*, while T_i denotes the individual threshold for the system *i*. V_i votes are collected by comparing the score of each system against its threshold.

$$V_i = \begin{cases} 0 & , LR_i < T_i \\ 1 & , LR_i \geq T_i \end{cases} \tag{1}$$

After collecting all V_i votes, a score *S* is computed as a sum of these votes, as shown in 2. This *S* score is compared to a threshold *T* to form a final decision. However, determining the T_i values for each system individually is computationally expensive.

$$D = \begin{cases} \text{reject}, S = \sum_i V_i < T \\ \text{accept}, S = \sum_i V_i \geq T \end{cases} \tag{2}$$

Fixed-Far Voting Fusion(FFVF) proposed by Spreeuwers [10] aimed to simplify the determination of the individual threshold T_i. The authors proposed to set a False Acceptance Rate(FAR) value for each system in order to determine the individual thresholds. It was assumed that poor systems would not be able to cast too many votes since their FAR value determines their threshold as high.

In this research, this approach implemented in the patches. Each patch was considered as an individual system. By setting a FAR value, an individual threshold was computed for each patch.

5) Fine Registration: A fine registration step has been implemented based on the idea of minimising/maximising an objective proposed by Spreeuwers et.al. [13]. The objective in this research was determined as maximizing the output of the network, aka. matching score. Different alignment candidates have been generated by applying a shift operation within a range of values on the object pair. Then, each pair has been scored, and the candidate pair having the maximum matching score has been selected as the registered pair. Figure 3 shows the steps involved in fine registration.

Two implementations have been done by utilising the proposed matching score based approach.

a) Local Fine Registration: The shift operation has been applied on individual patch pairs. The object patch was shifted up to 4 pixels in 8 directions, namely up, down, left, right, and their combinations. Since the displacements were in the patch level, the small range of shift values has been selected.

b) Global Fine Registration: The shift operation has been applied on the object image. The whole image has been shifted in 8 directions, namely up,down, left, right, an their combinations. In the global level, larger displacements were considered. Therefore, shift range has been determined as 40 pixels.

IV. Experiments and Results

In this section, five experiments were conducted to demonstrate the performance of the proposed solutions. The first three experiments have been done to find optimal patch properties namely patch size, patch shape, and overlapping stride. The fourth experiment was executed to investigate the feasibility of Fixed-FAR voting fusion on the patch based finger vein verification. In the last experiment, the performance of the proposed fine registration method was examined.

In all experiments, except Experiment 3, the datasets more or less have the same size, approximately 190000 patch pairs. Since the changes in patch size and shape affect the number of patches extracted, some overlap was applied to keep the dataset sizes stable.

The network was trained with the appropriate settings for each experiment. For instance, if en evaluation is done with

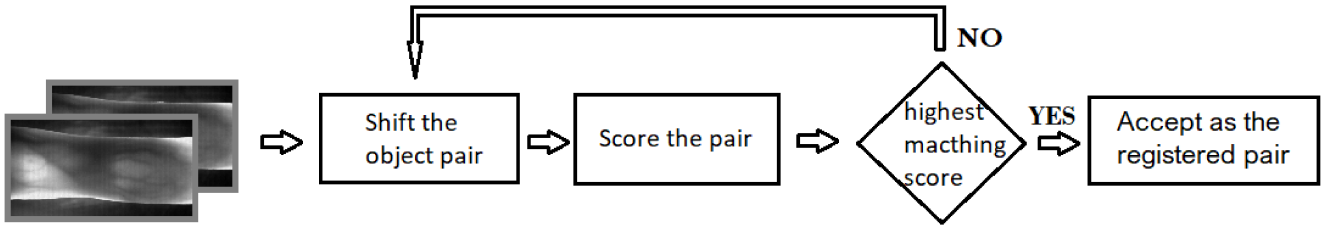


Fig. 3: Steps involved in fine registration

82-pixel patches, the training also done with the 82-pixel patch size. Only UTFVP dataset was used to train the network due to its higher quality.

The performance of the system was evaluated by ROC curves which shows the relationship between True Positive Rate(TPR) and False Positive Rate(FPR) of a system. In addition to ROC curves, Area Under the Curve(AUC) and Equal Error Rate(EER) were also used to measure the systems performance. Fusion experiments was evaluated by using FRR@FAR=0.1% performance metric.

A. Databases

In this work, 2 different databases from two universities, which are explained below, were used to evaluate the proposed optimisation solutions.

1) University of Twente: UTFVP [9] consists of 1440 finger vein images from 60 subjects. In total, 24 finger vein images were captured from each subject, which consists of total 6 fingers from 2 hands e.g., index, middle, and ring fingers. The fingers were illuminated above by 8 near infrared LEDs. The LEDs had been adaptively controlled for more uniform intensity. Some samples are seen in Figure 4.

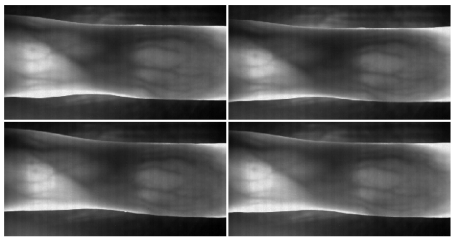


Fig. 4: Sample finger vein images from UTFVP

The finger vein images were saved in 8-bit gray scale png files with a resolution of 672x380 pixels. No guide bar was used during capturing process. Image capturing was done in two sessions.

2) Shandung Univesity: SDUMLA-HMT [27] was developed by Shandung University. It is a muti-modal biometrics database which consists of face, iris, finger vein, and gait images. The finger vein database composes 3,816 finger vein images from 106 persons. 6 images were captured for each of the 6 fingers from 2 hands e.g. index, middle and ring fingers. Figure 5 shows some samples from the database.

The finger vein images were saved as 8 bit rgb format in bmp files. Resolution of the images are 320x240 pixels.

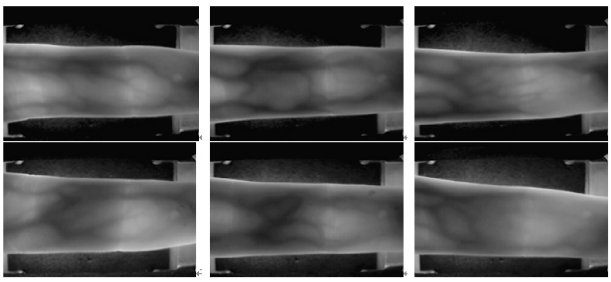
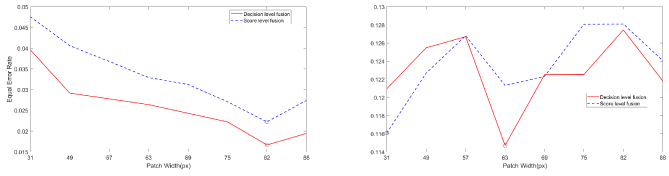


Fig. 5: Sample finger vein images from SDUMLA-HMT

No guide bar was used during capturing the images, and no session information was given.

B. Experiment 1 - Patch Size

The performance of different patch sizes was investigated in this experiment. The patch sizes selected as 49, 57, 63, 69, 75, 82, and 88 pixels. A 10-subject subset of UTFVP was used. This dataset involves 2120 finger vein image pairs (720 genuine, 1400 imposter). All 106 subjects in SDUMLA-HMT database were used in the experiments. This dataset involves 57240 image pairs in total (19080 genuine, 38160 imposter).

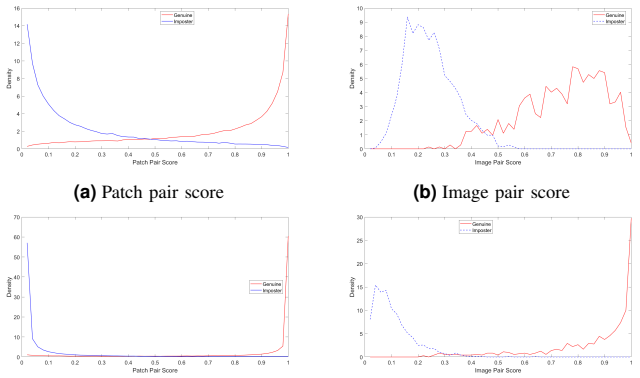


(a) (b)

Fig. 6: EER for varying values of patch sizes (a) UTFVP, (b) SDUMLA-HMT

Figure 6a shows a decrease in EER on UTFVP with the larger patch sizes. Figure 8 compares the ROC curves of 31 and 82-pixel patches. 82-pixel patches performed significantly better than 31-pixels. Moreover, the gap between decision and score level fusion performance was closed. The larger patch involves more vein structure, therefore, the network could find better matches. Figure 7 supports this claim. The larger patch leaded a better score distribution compared to 31-pixel patches.

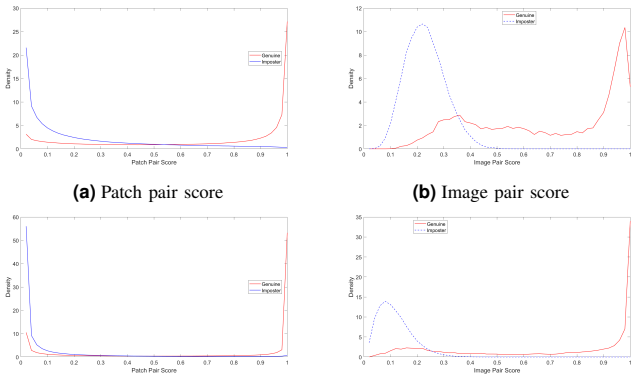
On the other hand, Figure 6b does not show any significant trend on SDUMLA-HMT compared to UTFVP. However, Figure 9 indicates a better separation between genuine and imposter scores with a larger patch size. Different from



(a) Patch pair score **(b)** Image pair score

(c) Patch pair score **(d)** Image pair score

Fig. 7: Pair score distributions on UTFVP (a) and (b) 31-pixel, (c) and (d) 82-pixel patches compared

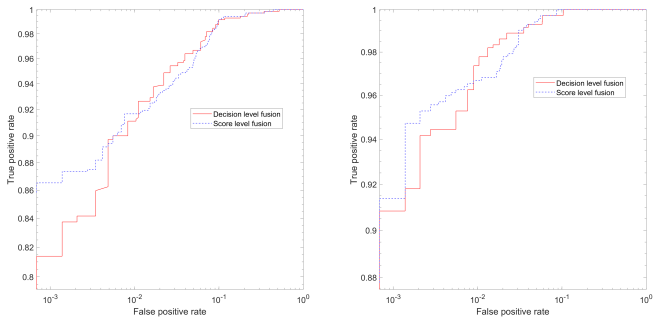


(a) Patch pair score **(b)** Image pair score

(c) Patch pair score **(d)** Image pair score

Fig. 9: Pair score distributions on SDUMLA-HMT (a) and (b) 31-pixel, (c) and (d) 69-pixel patches compared

UTFVP, some low genuine scores were persistent to change in the patch size while high score genuine were moving to the right edge of the plot. These genuine pairs generally had an extreme translation on the object pair. Therefore, the patch size did not help on these pairs alone. On the other hand, Figure 10 shows a slight improvement in performance. The change in score distribution also improved the verification performance, yet this improvement was not as remarkable as seen in UTFVP.



(a) 31-pixel **(b)** 82-pixel

Fig. 8: Image pair decision and score level fusion ROCs on UTFVP. 31-pixel and 82-pixel patches are compared.

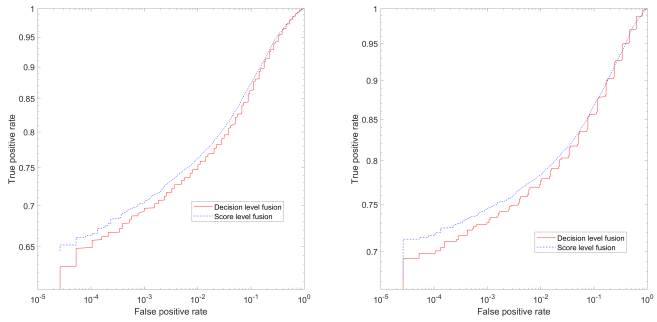
C. Experiment 2 - Patch Shape

The purpose of Experiment 2 was to question the square shape patches. Since the veins lay horizontally, rectangular patch shape with different widths were investigated. The height of the patch was determined in the first experiment. The patch widths used in the experiments can be seen in Table I.

Dataset	Patch Width (px)			
UTFVP	123	164	205	126
SDUMLA-HMT	104	134	174	209

TABLE I: Patch widths used in experiments

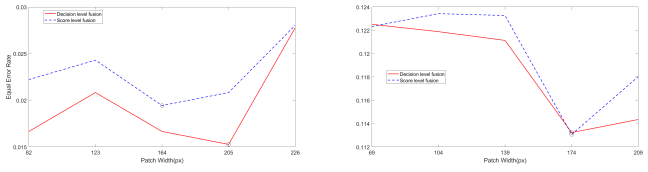
Figure 11 shows a decrease in EER on both datasets with a rectangular patch shape. The improvement was not as remarkable as seen with the patch size on UTFVP. The UTFVP



(a) 31-pixel **(b)** 69-pixel

Fig. 10: Image pair decision and score level fusion ROCs on SDUMLA-HMT. 31-pixel and 82-pixel patches are compared.

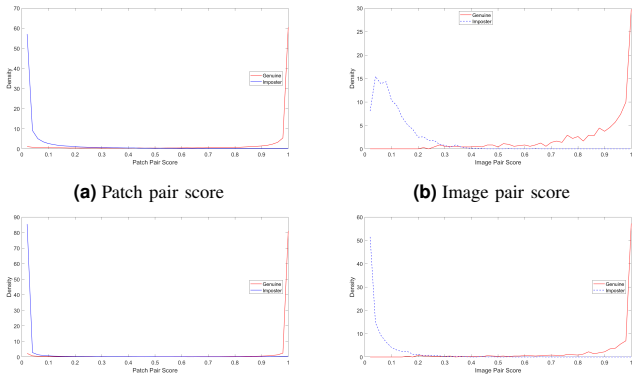
can be stated as a higher quality dataset. Therefore, a large enough square patch performs as good as a rectangular patch. Figure 12 point outs that the rectangular shape provided better separation than the square one. More horizontal information about the veins helped to solve some ambiguity between genuine and imposter pair.



(a) **(b)**

Fig. 11: EER for varying patch widths (a) UTFVP, (b) SDUMLA-HMT

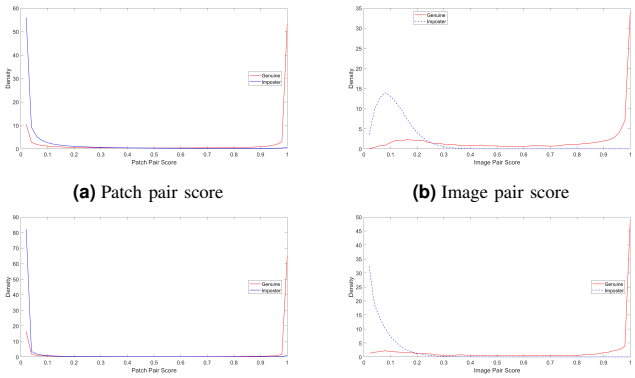
On the other hand, the improvement was more remarkable on SDUMLA-HMT dataset. Figure 13 indicates that the low genuine score density decreased significantly with the rectangular patch. Horizontal information helped more on SDUMLA-HMT. Since the translations seen on the dataset deforms the vein structure on vertical axis more than horizontal one, more horizontal information leaded a better score separation than vertical information.



(a) Patch pair score **(b)** Image pair score

(c) Patch pair score **(d)** Image pair score

Fig. 12: Pair score distributions on UTFVP (a) and (b) 82-pixel, (c) and (d) 205-pixel patch widths compared



(a) Patch pair score **(b)** Image pair score

(c) Patch pair score **(d)** Image pair score

Fig. 13: Pair score distributions on SDUMLA (a) and (b) 69-pixel, (c) and (d) 174-pixel patch widths compared

A small search around the best patch shape was conducted on both datasets. Table II shows the selected patch shapes for these experiments. Even though the improvement on UTFVP was not significant, 87 x 210-pixel patches were selected for the further experiments. For SDUMLA-HMT, the selected patch sizes was kept the same.

Dataset	Heights (px)	Widths (px)
UTFVP	77, 82, 87	200, 205, 210, 215, 220
SDUMLA-HMT	64, 69, 74	169, 174, 179

TABLE II: Search space around the best performing patch shape

D. Experiment 3 - Overlapping Patches

In this section, two sub-experiments were conducted to investigate the influence of the overlapping patches on training and evaluation stages. In both experiments, the best patch shapes were used. The sizes of the data-sets used in these experiments were proportional to the stride used in each dataset. The stride values for each dataset are given in Table III. In both experiments, the network was trained with only UTFVP due to its better image quality.

Slide(h-w) / Dataset	s0	s1	s2	s3	s4
UTFVP	no-overlap	53 - 172	50 - 103	19 - 24	16 - 13
SDUMLA-HMT	no-overlap	42 - 143	39 - 85	26 - 22	12 - 10

TABLE III: Slide values used in overlap experiments for each dataset

1) Without overlapping Training: The network was trained by a non-overlapping data-set. This set consisted of 23,930 patch pairs in total. The experiments were conducted on the 5 overlapping settings. The dataset settings are presented in Table III

Figure 14a shows a continuous decrease in EER on UTFVP with smaller strides. More overlap on classifiers increases the number of votes; therefore, the performance increased with more overlap. Between *s3* and *s4* evaluation sets, almost no change was observed in EER. After a stride value, the score distribution did not change. Thus, the performance remained the same.

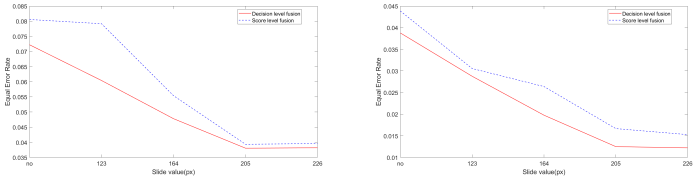
SDUMLA-HMT showed a different trend than UTFVP with smaller strides. Figure 15a indicates an increase in EER after *s1* evaluation set. Larger overlap leaded a better separation between genuine and imposter scores. However, the low score genuine pairs mentioned on the previous experiments were treated as imposter pairs. Their score also decreased with smaller strides, so the EER increased. Similarly, after a stride, the patches started to become too similar such that the distribution did not change.

2) Overlapping Training: The same experiments were executed using a network trained with an overlapping dataset. *s3* dataset was used for training. This data-set was the same with the dataset used in patch shape experiments. It consisted of 952,970 patch pairs, in total.

As seen in the Figure 14b and Figure 15b, the EER was lower with overlapping training. Overlapping applied on the training data helped the network to learn the veins better by providing more variation. Therefore, the network was able to find a better matching. Other than the decrease in EER, the performance of the overlapping classifiers had almost the same trend.

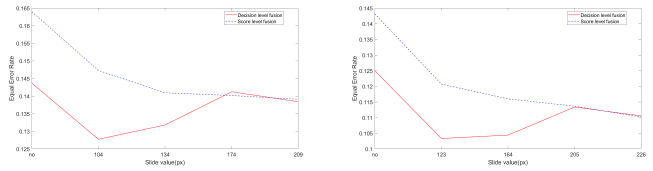
E. Experiment 4 - Fusion

The aim of this experiment is to demonstrate the feasibility of a non-uniform fusion approach on patch based system. This approach consists two stages, namely training and evaluation. Therefore, the datasets were divided into two partitions training and evaluation. Training part was used to determine



(a) Training with without-overlap **(b)** Training with overlap

Fig. 14: EER of UTFVP for varying overlap strides. The network trained (a) without-overlap, (b) with overlap



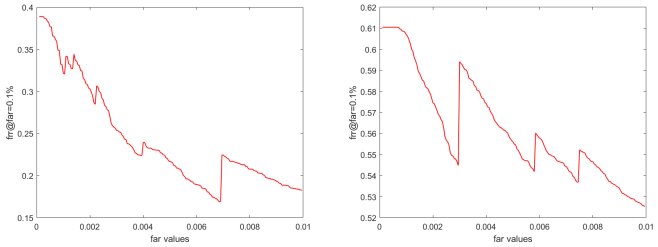
(a) Training with without-overlap **(b)** Training with overlap

Fig. 15: EER of SDUMLA-HMT for varying overlap strides. The network trained (a) without overlap, (b) with overlap.

individual thresholds using a set of FAR values. Evaluation set was used to fine tune the individual thresholds at FAR=0.1%. The number of image pairs used in training and evaluation are given in Table IV.

	Train	Evaluation
UTFVP	7776	3044
SDUMLA-HMT	34560	20925

TABLE IV: Number of image pairs involved by train and evaluation sets in Fixed-FAR voting fusion experiments for UTFVP and SDUMLA-HMT



(a) UTFVP **(b)** SDUMLA-HMT

Fig. 16: FRR@FAR=0.1% plots (a) UTFVP, (b) SDUMLA-HMT

The performance of the fused patch pairs on the evaluation set in terms of FRR@FAR=0.1% as a function of fixed FAR values is seen in Figure 16 for both datasets. Fixed FAR values were selected as 0.7% and 0.99% for UTFVP and SDUMLA-HMT, respectively, since the minimum FRR was reached.

	UTFVP		SDUMLA-HMT	
	EER%	FRR@FAR=0.1%	EER%	FRR@FAR=0.1%
FFVF	0.0614	0.254	0.213	0.603
Uniform Voting Fusion	0.0395	0.186	0.147	0.448

TABLE V: Performance comaprison of FFVT and non-uniform voting fusion methods on both UTFVP and SDUMLA-HMT

Table V compares the performance of the proposed FFVF and the current uniform vote fusion methods. A remarkable decrease in performance was observed with the proposed FFVF. At training stage, the individual thresholds were computed so high so that few genuine patch pairs were able to vote. Moreover, no significant difference between individual patch thresholds was observed.

As the uniform vote fusion outperformed the proposed FFVF approach, further experiments were terminated.

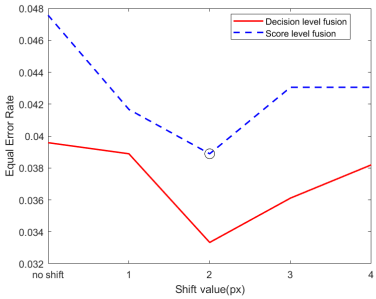
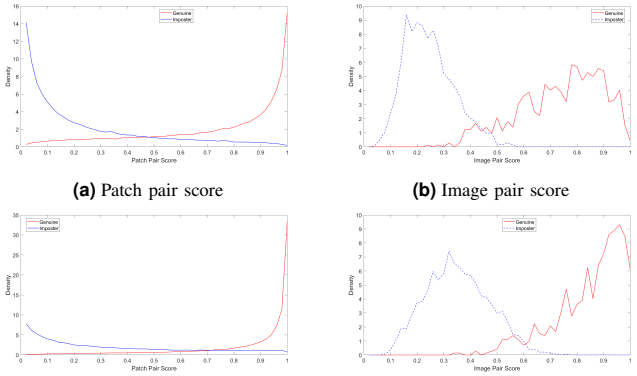


Fig. 17: EER of Local fine registration applied on UTFVP up to 4-pixel shift.

F. Experiment 5 - Fine Registration

Two sub-experiments were conducted to investigate the feasibility of a fine registration approach.

1) Local Fine Registration: Local fine registration assumes that finger vein images are captured with small displacements on the x-axis. Therefore, a small set of shift values up to 4 pixels was experimented.



(a) Patch pair score **(b)** Image pair score

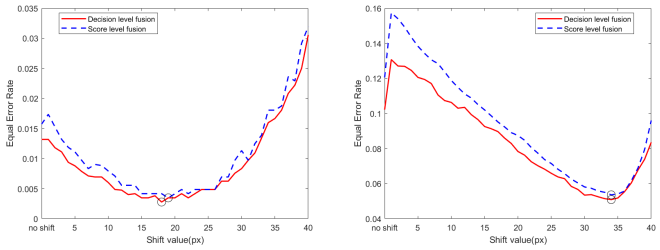
(c) Patch pair score **(d)** Image pair score

Fig. 18: Pair score distribution comparison between (a), (b) no shift and (c), (d) 4-pixel local shift on UTFVP

Figure 17 indicates that local fine registration did not provide much improvement on the EER of the system. Even though small improvements were observed, no significant change on the EER was measured. The score distribution graphs presented in Figure 18 compare the difference between no-shift and 4-pixel local shift. After local shift operation, both imposter patch pair and image pair score distribution moved to right significantly. Due to simultaneous increase in both genuine and imposter pair scores, the overall performance of the system did not show an improvement as expected. Therefore, further experiments has been terminated.

2) Global Fine Registration: Fine registration was applied on the whole image. Thus, a larger range for shift values was able to be investigated. The shift operation was applied maximum 40 pixels on both UTFVP and SDUMLA-HMT databases in 8 directions such as up, down, right, left, and the combinations of these directions.

As seen in the Figure 19a, lower EER values were observed with larger shift values on UTFVP. However, there is a limit



(a) UTFVP **(b)** SDUMLA-HMT

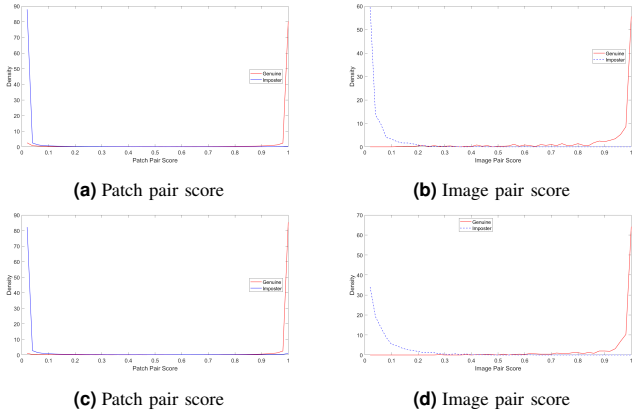
Fig. 19: EER of Global fine registration applied on (a) UTFVP, (b) SDUMLA-HMT, up to 40-pixel shift.

for the shift. After approximately 30 pixels, the EER values started to increase gradually. This was an expected increase. Because of larger shift values, the object pair image had started to go out of the window, and the orientation between the image pairs was distorted. At the end, lower matching scores were obtained.

Figure 20 compares the score distributions of the best performing patch shape with and without the global fine registration. It is seen in Figure 21a and Figure 21b indicate that global fine registration leaded an increase in the performance. This could be interpreted as the global fine registration leaded a better registration accuracy; therefore the performance improved. Figure 20b and Figure 20d indicate a significant increase in imposter scores; however, different form the local approach, genuine scores increased more than imposter scores. Since all the patch pairs were involved in the best matching score, even some imposter patch pair scores increases too much, the rest were able to keep the balance.

Figure 19b shows that SDUMLA needs larger shift values for a better registration. This difference between UTFVP and SDUMLA-HMT could be expressed in the translations seen on SDUMLA-HMT. Larger shifts would help finding a better match for these pairs.

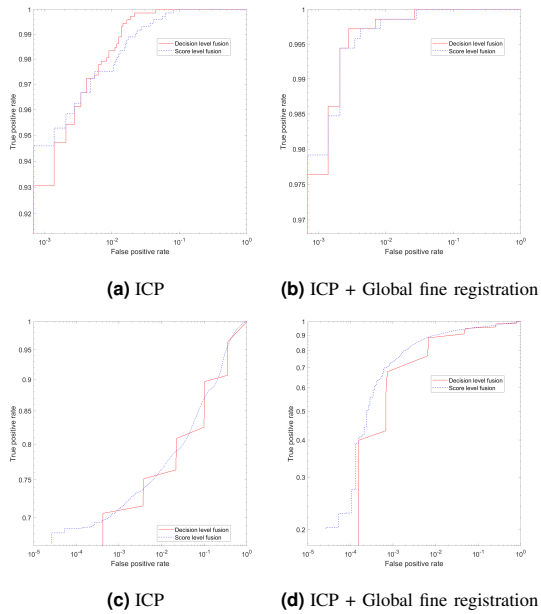
However, Figure 22d shows a significant distortion in image pair score distribution compared to Figure 22b. Low score



(a) Patch pair score **(b)** Image pair score

(c) Patch pair score **(d)** Image pair score

Fig. 20: Comparison of score distributions (a),(b) without and (c),(d)with global fine registration on UTFVP. Global fine registration distributions plotted where the EER is lowest(18-pixel shift)

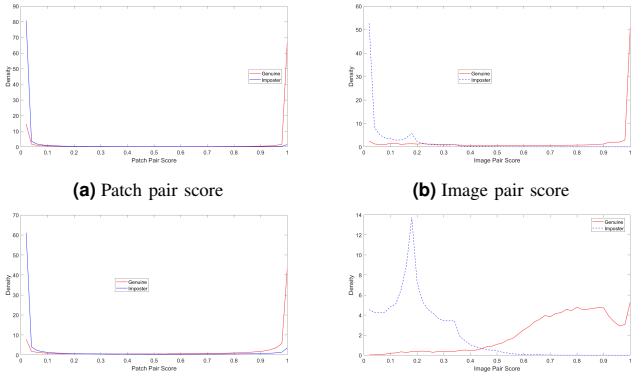


(a) ICP **(b)** ICP + Global fine registration

(c) ICP **(d)** ICP + Global fine registration

Fig. 21: Comparison of ROCs with and without global fine registration (a) and (b) on UTFVP, (c) and (d) SDUMLA-HMT. ROCs with global fine registration was plotted at 18-pixel and 35-pixel shift, respectively(the lowest EER)

genuine image pair density decreased as expected, yet high scored genuine image pair distribution moved to left. This distortion in genuine pair scores, together with the significant improvement in imposter scores, caused a performance degradation at lower FARs, seen in Figure 21d.



(a) Patch pair score **(b)** Image pair score

(c) Patch pair score **(d)** Image pair score

Fig. 22: Comparison of score distributions (a),(b) without and (c),(d)with global fine registration on SDUMLA-HMT. Global fine registration distributions plotted where the EER is lowest(35-pixel shift)

On the other hand, Figure 23a indicates that a modest shift around 25-pixel was able to keep the imposter and genuine pair score distribution separated while providing a performance improvement.

V. Discussion and Future Work

The purpose of this research is to provide an optimisation to the patch-based finger vein verification system. Patch properties, fusion method, and registration step has been tried to be optimised.

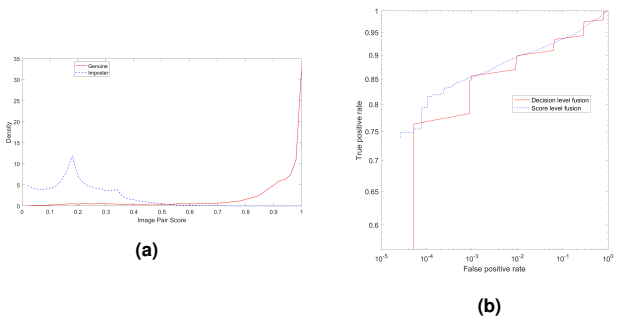


Fig. 23: (a) image pair score distribution and (b) ROC where shift is 25-pixel on SDUMLA-HMT.

Table VI presents some examples of the experiment results. Since couple of experiments improvement in AUC became too small only the EER results are presented. Overall, the optimisation achieved promising results with 0.3% of EER on UTFVP. The patch-based finger vein verification system outperformed some CNN based methods such as [16] (0.42%) and [17] (0.4%). However, the performance is still below the conventional methods. For comparison, conventional approaches in [1], [3] have achieved 0.145% and 0.25%, respectively. However, the small difference between the obtained and state-of-the-art results indicates the potential of the proposed patch-based system.

The patch size method performed better on UTFVP, while the patch shape method was more successful on SDUMLA-HMT. The datasets had some major differences. The image qualities could be considered as high in UTFVP. However, in SDUMLA-HMT, images without visible veins were more common. Moreover, SDUMLA-HMT provided many finger samples with extreme translations causing a deformation in the vein structure. Rectangular patches might avoid these deformations on the vertical axis, while involving less deformed information on the horizontal axis. As in UTFVP, these translations are not common, a large enough square patch could involve as much information as a rectangular patch.

Overlapping patches improved the performance on both datasets. It has been found that overlapping applied on training data leaded the network to learn the veins better by adding more variation on training stage. Moreover, overlapping applied on classifiers also improved the performance. As it increases the number of votes for a finger, the more vote generally leaded a better performance.

Proposed FFVF showed a different behavior than stated in [10] on evaluation stage. Unlike [10], FRR@FAR=0.1% plots had a continuous declining trend. The difference might be expressed with the different experimental settings. The number of imposter patch pairs extracted from relative locations were higher than expected; therefore, individual thresholds were set too high so that genuine votes decreased drastically. Fixed FAR criteria may not be suitable for this research for the individual patch pair setting. Additionally, computed thresholds did not differ much among patch pairs. Thus, differences between individual patches could be ignored. Fine tuning on one uniform threshold for all patches performed better.

Fine registration approach was successful in global level. By shifting, generally a better matching has been found for an imposter pair; therefore, imposter pair scores generally increased after shifting. However, this increase has been controlled by the mean score used in global fine registration. Even some imposter patch pairs scored high after shifting, the lower scored patches were able to keep this increase a balance since the mean matching score was selected as the objective for the global level. On the other hand, the increase in imposter scores could not be controlled in local level approach. Therefore, the increase in imposter pair scores extinguished the improvement on genuine pairs.

In this research, only a few aspects of the patch-based system have been selected for optimisation. In addition to the existing ones, the research arose new research topics.

First of all, some patches could have a little influence on the final score. [20] showed that filtering out the less relevant patches could lead more accurate verification results. Therefore, an algorithm selecting the relevant patches could lead an improvement on the performance.

The implemented fusion revealed that the importance of a patch pair over the other pairs could be ignored. However, some regions could still be more important than the others. For example, joint regions are generally dark and does not involve much visible vein. The implemented fusion method could be adapted to work with regions rather than individual patches. Moreover, such a fusion system could be implemented by defining different weight to different locations on the existing fusion methods.

The network was out of the scope of this research; however, changes in the CNN could also lead an improvement. The input size of the network did not change during the patch size and shape experiments. Rather, the extracted larger patch was re-scaled to the input size of the network. This re-scaling operation might deform the vein structure. By changing the input size and the network organisation, better results might be achieved.

Moreover, rather than using the same network, a new network structure could be investigated. Siamese network structures contain two or more identical sub-networks. They are popular among the tasks involving finding similarity or a relationship between two comparable things. Since the weights will be shared among the sub-networks, they tend to have less complexity, therefore less data is needed. In this respect, a Siamese structure could provide more improvement on the patch-based approach.

VI. Conclusion

In this research, the feasibility of an optimisation on the patch-based finger vein verification system has been investigated. The optimisation has been applied on the patch properties, fusion strategy, and registration approach.

The proposed solutions achieved promising results with 0.3% of EER and 0.999 AUC on UTFVP, and 6.59% of EER anf 0.969 AUC on SDUMLA-HMT.

		University of Twente				SDUMLA-HMT			
		Before		After		Before		After	
		Decision	Score	Decision	Score	Decison	Score	Decison	Score
Patch Properties	Size	0.0395	0.0476	0.0166	0.0222	0.121	0.116	0.122	0.123
	Size+Shape	0.0166	0.0222	0.0153	0.0208	0.122	0.123	0.113	0.113
	Size+Shape+Overall	0.0153	0.0208	0.0122	0.0153	0.113	0.113	0.103	0.121
Fine Registration	Local	0.0395	0.0476	0.0333	0.0388	-	-	-	-
	Patch Properties+ Global	0.0122	0.0153	0.003	0.004	0.103	0.121	0.066	0.0715
Overall		0.0395	0.0476	0.003	0.004	0.121	0.116	0.066	0.0715

TABLE VI: Overall comparison of the experiemnts

Optimal patch properties are dependent on the dataset characteristics. On a lower quality dataset with extreme translations, rectangle patches could lead better results than square ones.

Overlapping helped in both cases. When it is applied on the training data, overlapping leads better learning. Overlapping on the classifiers improves the performance by increasing the number of votes per image pair.

Contrary to the expected, differences between individual patches could be ignored because the difference between computed thresholds were negligible. Setting and fine tuning on one threshold for all patches have performed better than individual thresholds on both datasets.

The matching score based fine registration approach leaded a better registration accuracy on global level. Local level did not perform as expected due to the uncontrolled increase in imposter pair scores.

Overall, the proposed optimisations achieved promising results and reinforced the potential of the patch-based finger vein verification approach. Even though the obtained results are less satisfied than of the state-of-the-art, any improvements made on these approaches proposed in this research may provide more satisfied results.

REFERENCES

[1] N. Miura, A. Nagasaka, T. Miyatake, Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification,T. Machine Vision and Applications, October 2004, Vol. 15, pp. 194-203.

[2] N. Miura, A. Nagasaka, T. Miyatake,Extraction of Finger-Vein Patterns Using Maximum Curvature Points in Image Profiles, IEICE TRANSACTIONS on Information and Systems, August 2007, Vol. E90-D, pp. 1185-119.

[3] W. Song, T. Kim, H. C. Kim, J. H, Choi, H. J. Kong, S. R. Lee, A finger-vein verification system using mean curvature, Pattern Recognition Letters, August 2011, vol. 32, pp. 1541-1547.

[4] J. Peng, N. Wang, A. A. El-Latif, Q. Li, X. Niu, Finger-vein Verification Using Gabor Filter and SIFT Feature Matching, International Conference on Intelligent Information Hiding and Multimedia Signal Processing, July 2012.

[5] B. A. Rosdi, C. W. Shing, S. A. Suandi, Finger Vein Recognition Using Local Line Binary Pattern, 2011, Sensors, vol. 11, pp. 11357-11371.

[6] X. Meng, X. Xi, G. Yang, Y. Yin, Finger vein recognition based on deformation information, Science China Information Sciences, May 2018.

[7] H. Qin, L Qin, L. Xue, C. Yu, X. Liang, Finger-Vein Verification Based on Multi-Features Fusion, Sensors, 2013, vol. 13, pp. 15048-15067.

[8] A. Radzi, M. K. Hani, R. Bakhteri, Finger-vein biometric identification using convolutional neural network, Turkish Journal of Electrical Engineering and Computer Sciences, 2016, vol. 24, pp. 1863 1878.

[9] B. T. Ton and R. N. J. Veldhuis, A high quality finger vascular pattern dataset collected using a custom designed capturing device, International Conference on Biometrics (ICB), 2013, pp. 1 - 5.

[10] L. Spreeuwers, and R. N. J. Veldhuis, S. Sultanali,J. Diephuis, Fixed FAR vote fusion of regional facial classifiers, International Conference of the Biometrics Special Interest Group (BIOSIG), 2014, pp. 1 - 4.

[11] S. de Boer, Patch-based Finger Vein Verification with Convolutional Neural Network, 2017.

[12] Paul J Besl, Neil D McKay, et al. A method for registration of 3-d shapes. IEEE Transactions on pattern analysis and machine intelligence, 1992, vol. 14(2), pp. 239 - 256.

[13] L. Spreeuwers, B. Boom, Better than best:Matching score based face registration, 2007.

[14] G. Meng, P. Fang, B. Znag, Finger Vein Recognition Based on Convolutional Neural Network, MATEC Web Conference, 2017, vol. 128.

[15] W. Liu, W. Li, L. Sun, L. Zhang, P. Cheng, Finger vein recognition based on deep learning, IEEE Conference on Industrial Electronics and Applications (ICIEA), 2017.

[16] H. Huang, S. Liu, H. Zeng, L. Ni, Y. Zang, W. Li, DeepVein: Novel finger vein verification methods based on Deep Convolutional Neural Networks, 2017, Security and Behavior Analysis (ISBA).

[17] H. G. Hong, M. B. Lee, K. R. Park, Convolutional Neural Network-Based Finger-Vein Recognition Using NIR Image Sensors, Sensors, 2017.

[18] A. Krizhevsky, N. Vinod, H. Geoffrey Hinton, The CIFAR-10 dataset, 2014, online: http://www.cs.toronto.edu/kriz/cifar

[19] J. Deng, W. Dong, R. Socher, L.J. Li, K. Li, L. Fei-Fei, Imagenet: A large-scale hierarchical image database, Computer Vision and Pattern Recognition(CVPR), June 2009, pp. 248-255.

[20] L. Hou, D. Samaras, T. M. Kurc, Y. Gao, Patch Based Convolutional Neural Network for Whole Slide Tissue Image Clssification, IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2016, 2424 - 2433.

[21] G. Keren, J. Deng, J. Pohjalinan, B.W. Schuller, Convolutional Neural Network with Data Augmentation for Classifying Speakers' Native Language , Interspeech, 2016.

[22] S. Hoffman, R. Sharma, A. Ross, Convolutional Neural Network for Iris Presentation Attack Detection: Toward Cross-Dataset and Cross-Sensor Generalisation, 2018.

[23] B. Huang, Y. Dai, R. Li, D. tang, W. Li, Finger Vein Authentication based on Wide Line Detector and Pattern Normalisation, International Conference on Pattern Recognition, 2010.

[24] E.C. Lee, H. Jung, D. Kim, New Finger Vein Biometrics using Near Infrared Imaging, 2011, Sensors, vol. 11.

[25] JD. Wu, LT. Lui, Finger Vein Identification Using Principal Component Analysis and the Neural Network Technique, 2011, Expert Systems with Applications, Vol. 38, pp. 5423 - 5427.

[26] JD. Wu, LT. Liu, Finger Vein Identification using SVM and Neural Network Technique, 2011, Expert Systems with Applications, Vol. 38, pp. 14284 - 14289

[27] Y. Yin, L. Liu, X. Sun, SDUMLA-HMT: A Multimodal Biometric Database, Chinese Conference on Biometric Recognition (CCBR), 2011, LNCS 7098, pp. 260-268.

[28] C.B. Yu, H.F. Qin, Y.Z. Cui, X.Q. Hu, Finger-vein image recognition combining modified hausdorff distance with minutiae feature matching, 2009, Interdisciplinary Sciences: Computational Life Sciences, Vol. 1, pp. 280 - 289.

[29] F.Liu, G. Yang, Y. Yin, S. Wang, Singular value decomposition based minutiae matching method for finger vein recognition, Neurocomputing, 2014, Vol. 145, pp. 75 - 89.

[30] P. Wang, L.C. Tran, Q.Ji, Improving Face Recognition by Online Image Alignment, International Conference on Pattern Recognition (ICPR'06), 2006, pp. 311 - 314.

[31] A. Kortylewski, A. Schneider,T. Gerig, B. Egger, A. Morel-Forster, T. Vetter, Thomas,Training Deep Face Recognition Systems with Synthetic Data, 2018.
[32] L. Perez, J. Wang, Jason, The effectiveness of data augmentation in image classification using deep learning, 2017.
[33] L. Dong, G. Yang, Y. Yin, F Liu, X Xi, Finger vein verification based on a personalized best patches map, Biometrics (IJCB), 2014, pp. 1 - 8.
[34] H. Qin, L. Qin, L. Xue, X. He, C. Yu, X. Liang, Finger-vein verification based on multi-features fusion, Sensors, 2013, vol. 13, pp. 15048 - 15067.
[35] H. Qin, M. El-Yacoubi, Deep representation-based feature extraction and recovering for finger-vein verification, IEEE Transactions on Information Forensics and Security, 2017, vol. 12, pp. 1816 - 1829.
[36] P. Zhu, L. Zhang, Q. Hu, S.C.K. Shiu, Multi-scale patch based collaborative representation for face recognition with margin distribution optimization, European Conference on Computer Vision, 2012, 822 - 835.
[37] F. Liu, Y. Yin, G. Yang, L. Dong, X. Xi, Finger vein recognition with superpixel-based features,Biometrics (IJCB), 2014 IEEE International Joint Conference on, 2014, 1 - 8.