Rotation Detection in Finger Vein Biometrics using CNNs

Bernhard Prommegger* and Georg Wimmer* and Andreas Uhl
Department of Computer Sciences
University of Salzburg
Salzburg, Austria
Email: {bprommeg, gwimmer, uhl}@cs.sbg.ac.at
*: These authors contributed equally

***Abstract*—Finger vein recognition deals with the identification of subjects based on their venous pattern within the fingers. The recognition accuracy of finger vein recognition systems suffers from different internal and external factors. One of the major problems are misplacements of the finger during acquisition. In particular longitudinal finger rotation poses a severe problem for such recognition systems. The detection and correction of such rotations is a difficult task as typically finger vein scanners acquire only a single image from the vein pattern. Therefore, important information such as the shape of the finger or the depth of the veins within the finger, which are needed for the rotation detection, are not available. This work presents a CNN based rotation detector that is capable of estimating the rotational difference between vein images of the same finger without providing any additional information. The experiments executed not only show that the method delivers highly accurate results, but it also generalizes so that the trained CNN can also be applied on data sets which have not been included during the training of the CNN. Correcting the rotation difference between images using the CNN's rotation prediction leads to EER improvements between 50-260% for a well-established vein-pattern based method (*Maximum Curvature*) on four public finger vein databases.**

I. Introduction

Vascular pattern based biometric systems, commonly denoted as vein biometrics, offer several advantages over other well-established biometric recognition systems. In particular, hand and finger vein systems have become a serious alternative to fingerprint based ones for several applications. Vein based systems use the structure of the blood vessels inside the human body, which becomes visible under near-infrared (NIR) light. As the vein structure is located inside the human body, it is resistant to abrasion and external influences on the skin. Furthermore, due to the blood flow exhibited in NIR finger vein videos, liveness detection techniques can be applied to prevent presentation attacks [2], [25].

The performance of finger vein recognition systems suffers from different internal and external factors. Internal factors include the design and configuration of the sensor itself, especially the NIR light source and the camera module. External factors include environmental conditions (e.g. temperature and humidity) and deformations due to misplacement of the finger, typically including shifts, tilt, bending and longitudinal rotation.

This project was partly funded from the FFG KIRAS project AUTFingerATM under grant No. 864785 and the FWF project "Advanced Methods and Applications for Fingervein Recognition" under grant No. P 32201-NBL.

Performance degradations caused by various types of finger misplacement are not new and have been addressed in several publications. Kumar and Zhou [12] addressed the need for robust finger vein image normalization, including rotational alignment, already in 2012. Chen *et al.* [4] stated that deformations caused by a misplacement of the finger can be corrected either during pre-processing, feature extraction or comparison. Moreover, the design of the finger vein sensor helps to avoid or reduce misplacements of the finger as well. In [20] the authors showed, that longitudinal finger rotation has a severe influence on the recognition performance of finger vein recognition systems. There are several approaches that try to handle these issues during the processing of the vein images. These approaches can be grouped into two different categories: (1) approaches that use classical single perspective capturing devices, e.g. [4], [5], [9], [13], [16], [19], [31] and (2) methods that acquire multiple perspectives either during enrolment [23], [24], or for both, enrolment and recognition, [3], [10], [28]. However, none of theses approaches quantify the extent (i.e. the rotation angle) of the misplacements on which the deformation is based on. Prommegger *et al.* estimated the rotation angles in four publicly available data sets in [22]. For the rotation estimation between two finger vein samples, one sample is rotated 90 times in the range of $\pm 45°$ in steps of $1°$. The other sample is compared to the first sample and its 90 rotated versions, so 91 comparisons in total. The rotation angle is taken from the comparison at which the highest correlation (the highest score) is achieved. However, such a time-intensive empirical approach can only be used to analyse existing data sets, but is not suitable for real world applications. Therefore, a system that is able to determine the rotation angle between two vein images in real time would be desirable. This article proposes a CNN based rotation detector that is capable of doing so.

CNNs have already been used for rotation estimation in several biometric applications. In [26] a siames network based approach was used to estimate the rotation of finger prints. In [6], CNNs were applied to detect hands and estimate their rotation and in [8], CNNs were applied to estimate head pose

angles for face-related applications like face recognition. Up to now, there is no CNN-based prior work that tries to estimate the rotation of finger vein images.

The main contribution of this work is the proposal of a CNN-based rotation detector that estimates the difference in longitudinal rotation between vein images of the same finger. The experimental results show that the CNN is not only capable of estimating the rotation for the data set it is trained on, but can also be used for data sets not included during training. The CNN has been trained using data provided by the *PROTECT Multimodal Dataset* (PMMDB) [27] and evaluated on the *PLUSVein Finger Rotation Dataset* (PLUSVein-FR) [21]. To verify the generalisability of the proposed model, it is also applied on four often used publicly available data sets, namely SDUMLA-HMT [32], FV-USM [1], UTFVP [29] and PLUSVein-FV3 [11]).

The reminder of this paper is organized as follows: Longitudinal finger rotation and the problems it causes for finger vein recognition systems are described in more detail in section II. Sections III hold all details on the used CNN model and its training, section IV describes the region of interest detection and section V explains the rotation correction of finger vein images. The experimental setup together with its results are described in section VI. Section VII concludes the paper.

II. Longitudinal Finger Rotation in Finger Vein Recognition

Typically, finger vein scanners are designed to acquire a single finger at a time. Different types of finger misplacement can easily occur with these scanners. The different types of finger misplacement includes planar shifts and rotations, a change of the distance to the camera (scaling), finger bending, finger tilt (finger tip and root are not in the same plane) and longitudinal finger rotation. As described in [20], the influence of some of these problematic misplacements can be reduced or even prevented completely either during acquisition by adding support structures for finger positioning or a correction during pre-processing, feature extraction or comparison. Almost all currently available sensors use such support structures, but most of them still do not prevent longitudinal finger rotation. Thus, longitudinal finger rotation poses a severe problem to finger vein recognition systems.

The vein structure captured in finger vein images is a projection of the blood vessel structure in the 3D space onto a 2D plane. If the finger is rotated along its longitudinal axis, the vein pattern is deformed according to a non-linear transformation. Figure 1 shows the effect of longitudinal finger rotation on the vein pattern. The finger cross section (top row) is rotated from -30° to +30°. As a result of the rotation the projected pattern of the veins (bottom row) changes as well. Depending on the relative position of the veins to each other and the rotation angle, some of the captured veins might even merge into a single one. The vein structures of -30° (left), 0° (middle) and 30° (right) are completely different. Widely used vein recognition schemes can handle such deformations only to a certain extent [20]. If the deformations caused by the longitudinal rotation are corrected, the negative effect can be reduced but not completely prevented [19].

III. Rotation Detection using CNNs

The idea of the CNN-based rotation detector is to have pairs of different rotated but otherwise identical images as inputs for a CNN, so that it can learn to estimate the rotation difference. Typically, CNN inputs in image processing tasks are either 3-channel images (images in RGB or other color spaces) or one channel images (grayscale images). The proposed approach follows a different strategy: It uses a 2-channel input, where both channels contain grayscale finger vein images, with the image of the second channel being a rotated version of the image of the first channel (see figure 2). A somehow similar approach was already applied in [7], where two finger vein images where merged to a 2-channel image which was used as CNN input. However, in [7], the CNN was directly used for identification (using the cross entropy loss) whereas the proposed approach estimates the rotational difference between the two images.

In order to learn the rotation difference $\varphi_{I_i,I_i^{rotated}}$ between a pair of differently rotated images, I_i and $I_i^{rotated}$, the mean squared error (MSE) loss function, which is defined in Eq. (1), is applied.

$$L = \frac{1}{N}\sum_{i=1}^{N}(\varphi_i - \hat{\varphi}_i)^2 \tag{1}$$

φ_i is the actual rotational distance between a pair of training images, $\hat{\varphi}_i$ is the CNN's prediction of the rotation angle of the considered image pair and N is the batch size.

Figure 2 visualizes the CNN training process.

As CNN, the ResNeXt [30] architecture (ResNeXt-101), a highly modularized and deep network architecture for image classification, is used. The CNN weights are initialized from a model that was already trained on the ImageNet database[1]. The pre-trained model was trained on 3-channel RGB images, whereas the proposed model requires two channel input images. The problem is solved by replacing the original 3-channel filter kernels of the first convolutional layer with 2-channel filter kernels, whereas both channels of the new filter kernel are grayscale versions of the original 3-channel filter kernel. Furthermore, instead of the originally 1000-dimensional output from the ImageNet database, a one-dimensional CNN output (the predicted rotation difference) is required. Therefore, the last fully connected layer is resized from 1000*2048 to 1*2048 using randomly initialized weights.

IV. Region of Interest Detection

The *region of interest* (ROI), which serves as input for the CNN as well as the recognition tool-chain in section VI-C2, is extracted as following: First, an edge detection algorithm is used to detect the finger outlines. The area between the two finger lines is accounted as finger region. Next, a straight line is fitted between the two finger lines. This line represents

[1] http://www.image-net.org/

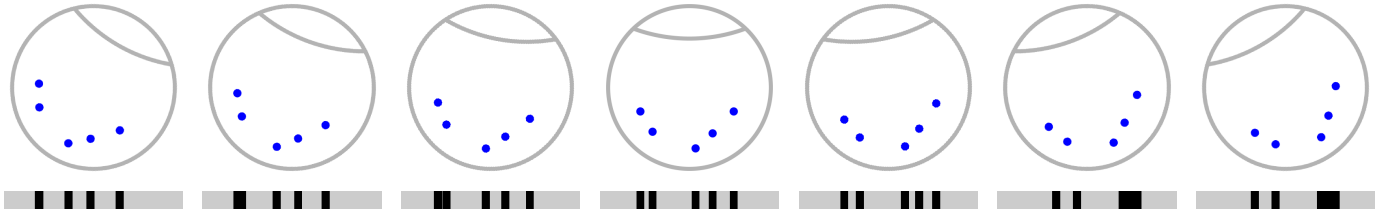

Fig. 1. Longitudinal finger rotation principle: a schematic finger cross section showing five veins (blue dots) rotated from -30° (left) to +30° (right) in 10° steps. The projection of the vein pattern (bottom row) is different depending on the rotation angle according to a non-linear transformation (originally published in [20]).

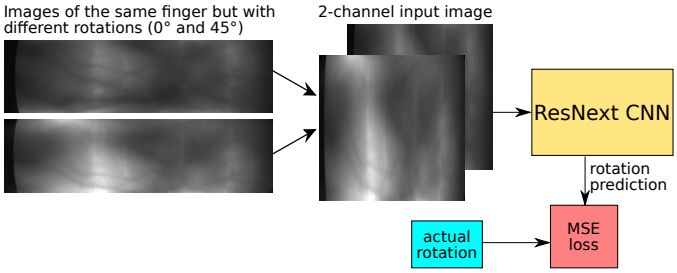

Fig. 2. Scheme of CNN training for rotation estimation

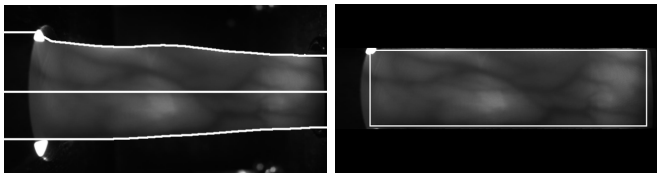

Fig. 3. ROI extraction - left: finger line detection. The straight line in the middle represents the center of the finger at which it is horizontally aligned. The top and bottom lines are the detected finger outlines which separate the finger from the background. The region between the lines is regarded as finger region. Right: the finger region is transformed to a fixed height. Afterwards the ROI, visualized as white square, of a fixed size is cut out.

the center of the finger. Based on this center line, the finger is aligned (rotated and vertically shifted) such that it is in a horizontal position and the center line of the finger is in the middle of the image. The area outside of the finger lines is masked out (pixels set to black). Afterwards, the finger region is transformed to a fixed height. In the last step, the finger ROI is cut out. Fig. 3 visualizes this process. The left image shows the finger with the center and finger outlines, the right image shows the final ROI after its transformation to a fixed height.

V. Rotation Correction of Finger Vein Images

For the generation of rotated vein images, not only the position of the veins in the 2D image, but also the shape of the finger and the depth of the veins within the finger have to be known. As the last two informations are not available in general, both need to be estimated. In this work, it is assumed that the cross section of a finger approximately resembles a circle (as e.g. Matsuda *et al.* did in [16]) and that the captured veins are close to the finger surface (cf. Huang *et al.* [9]).

The rotation of the veins by an angle of φ_{rotate} is calculated by applying a rotation matrix given in Eq. (2).

$$\begin{bmatrix} x_r \\ y_r \end{bmatrix} = \begin{bmatrix} cos(-\varphi_{rotate}) & -sin(-\varphi_{rotate}) \\ sin(-\varphi_{rotate}) & cos(-\varphi_{rotate}) \end{bmatrix} * \begin{bmatrix} x \\ y \end{bmatrix} \quad (2)$$

x and y are the coordinates of the vein pixel in the acquired image, x_r and y_r the ones in the rotated image. x is the position of the pixel in the vein pattern, y is calculated by

$$y = \sqrt{r^2 - x^2} \quad (3)$$

where r is the approximated radius of the finger, which corresponds to half of the height of the vein image. The part of the rotated ROI image that contains no information (due to the transform) is filled with the average gray level of the image. For more details, the interested reader is referred to [22].

VI. Experiments

The aim of the experiments is to show that a CNN can be trained to estimate the longitudinal rotation between two finger vein samples of the same finger using the approach presented ins section III. The rotational range, for which the CNN should be capable of estimating the longitudinal finger rotation, was determined based on the results of [19]. There, Prommegger *et al.* showed that a rotation correction gives very good results in the range of ±30°. However, a correction for rotation angles of more than 45° no longer makes sense as the recognition rates drop rapidly. Therefore, this work analyses the rotation estimation in the range of ±45°. Although, the range of particular interest is that between ±30°.

The training of the CNN model is done using data from the PMMDB data set. The range Θ, from which the training samples are taken is varied from ±45° to ±60° in steps of 5°. The evaluations are carried out on the PLUSVein-FR. To prove the generalizability of the model, it is also applied on four publicly available data sets, SDUMLA-HMT, FV-USM, UTFVP and PLUSVein-FV3, which have not been included during training. The rotation angles of these four data sets have also been evaluated in [22] and therefore, a direct comparison of the results from this work and [22] is possible.

A. Data Sets

The PLUSVein-FR provides vein images of perspectives all around the finger (360°) in steps of 1°. It was acquired for 63 different subjects, 4 fingers per subject, which sums up to

a total of 252 unique fingers. For each finger, five samples were acquired. Each one of the five samples consists of 361 images (one per perspective, 0° and 360° have been acquired separately). This results in $252 \cdot 5 = 1.260$ finger vein images per perspective. This work uses the publicly available subset *±45° Around the Palmar View* [19], which contains all images acquired for the 92 perspectives perspectives between -45° and +45°, resulting in a total of $92 \times 1.260 = 115.920$ vein images.

The PMMDB finger vein database was acquired in two data acquisition events with one year between the two sessions. In this work only data acquired during the second session using the same capturing device as for the PLUSVein-FR is used. From the 33 acquired subjects, only 29 are used in this work (4 subjects are part of both, PMMDB and PLUSVein-FR, and therefore were removed from PMMDB). This sums up to a total of 116 unique fingers (4 fingers per subject). As the two data sets were acquired using the same capturing device and the same acquisition protocol, they are very similar.

To show the generalizability of the presented rotation detector, it was used to evaluate the rotation angles in four publicly available finger vein data sets: SDUMLA-HMT, FV-USM, UTFVP and PLUSVein-FV3 (only the dorsal images acquired by the laser version of the sensor). The data sets itself do not provide any information on the longitudinal rotation of the samples. In [22], their rotation angles have been estimated. According to these estimates, PLUSVein-FV3 shows the lowest degree of longitudinal finger rotation, followed by UTFVP and FV-USM, while SDUMLA-HMT exhibits the highest amount.

B. Rotation Detection using CNNs

This Section describes the experimental setup for CNN training and evaluation and presents the CNN results.

1) Experimental Setup for CNN Training: CNN training is performed on pairs of images from the PMMDB database. For each image of the PMMDB database in the relevant range of $\pm\Theta$, a randomly chosen image of the same subject and sample but from a different perspective is selected as the second image of the input image pair (remember, for the PMMDB vein images are acquired all around the finger in steps of 1° $\Rightarrow$ one sample consists of 361 vein images from 361 different perspectives, where 0° and 360° are acquired separately). For these pairs of images, the exact rotational difference is known. Theoretically, 0° and 360° should be the same but practically there can occur small differences because of accumulated errors across the 360° rotation or small pose changes of the finger during data acquisition. Hence, to avoid any training errors, the pairs of images for CNN training are always selected either from the positive $[0°, \Theta]$ or negative range $[-\Theta, 360°]$, but no combinations across both rotational ranges.

The CNNs are trained for 60 epochs using a batch size of 8. Training starts with a learning rate of 0.0001 and is subsequently reduced by multiplying it with factor 0.3 after 20, 30, 40 and 50 epochs of CNN training. In every epoch, each of the 77.704 images of the PMMDB database is used once as first image of an image pair, the second image is randomly chosen from the same range as the first one. First, the 2-channel input image (the image pair) is resized to size 224×254. Data augmentation is applied by randomly cropping image patches of size 224×224 (the required input size for the CNN) from the resized image independently for each of the two channels. In that way the CNN's robustness to horizontal shifts (resulting from finger misplacements) is increased. This is important since there are no shifts between finger images of the same sample, whereas for evaluation, image pairs are built of images from different samples and so shifts do occur (also in practical application, the acquired images are subject to such misplacements). For evaluation, data augmentation is skipped and patches of size 224×224 are directly cropped from the center of the resized 224×254 images.

2) Evaluation Protocol: For evaluation, the trained CNN is applied on the subset *±45° around the palmar view* of the PLUSVein-FR data set. As already mentioned, PLUSVein-FR and PMMDB have been acquired using the same sensor and acquisition protocol. The only difference are the acquired subjects.

The rotation angles are always evaluated with respect to the palmar perspective (0° or 360°) of the first sample of each finger (denoted as reference image). Just as for the training setup, depending on the rotation angle α of the probe image, the rotation detection is made against the reference image at 0° ($\alpha >= 0°$) or 360° ($\alpha < 0°$), respectively. To achieve a more robust result, similar to [22], the rotation angle $\Phi_{i(\alpha)}$ is calculated as the average of $\hat{\varphi}_{i(\alpha),\text{ref}}$ (the predicted angle of the i^{th} sample at α against the reference image) and $\hat{\varphi}_{\text{ref},i(\alpha)}$ (the predicted angle of the reference image against the i^{th} sample at α):

$$\Phi_{i(\alpha)} = \text{avg}\left(\hat{\varphi}_{i(\alpha),\text{ref}}, -\hat{\varphi}_{\text{ref},i(\alpha)}\right) \tag{4}$$

As the CNN is trained on the range of $\pm\Theta$, it is not capable of estimating rotation angles outside of this range. Prediction results that exceed Θ are rejected and the resulting rotation angle $\Phi_{i(\alpha)}$ is taken from the remaining prediction. If both estimates, $\hat{\varphi}_{i(\alpha),\text{ref}}$ and $\hat{\varphi}_{\text{ref},i(\alpha)}$, are rejected, the rotation angle is set to 0°.

In order to correct any rotational misalignments between two samples of the same finger, the rotational distance of the i^{th} sample at the palmar view to the palmar view of the first sample ($\Phi_{i(palmar)}$) is subtracted from $\Phi_{i(\alpha)}$. The predicted rotation angle $\hat{\alpha}$ is thus determined as follows:

$$\hat{\alpha} = \Phi_{i(\alpha)} - \Phi_{i(palmar)} \tag{5}$$

3) Results: Figure 4 shows the result for all four training ranges Θ. The red solid line represents the median, the blue dashed lines mark the limits of the 90% quantile. The thinner dash-dotted lines serve as ledger lines for the ideal rotation prediction (prediction error = 0°) and ±15° (this is the range in which, according to [19], commonly used recognition deliver good recognition rates even without any rotation correction or compensation). For all four training ranges Θ, up to a rotation angle of ±30°, the median of the predictions is quite close to the ideal prediction. Outside of this range, the prediction

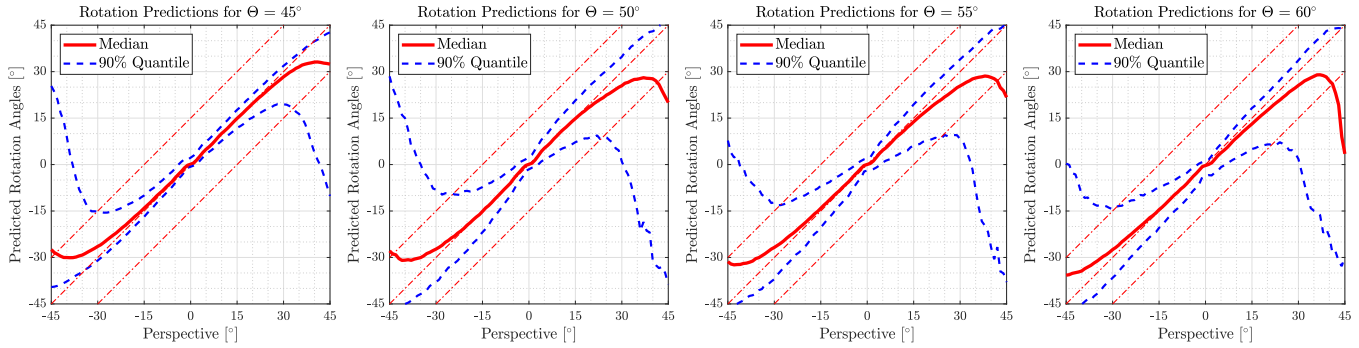

Fig. 4. Results of rotation detection on PLUSVein-FR (median and 90% quantile) for different training ranges Θ. From left to right: ±45°, ±50°, ±55°and ±60°

error increases. In general, the proposed CNN-model tends to underestimate the rotation angle. This can be seen in the plots by the fact, that the median is below the ideal estimate for rotation angles >0° and above for angles <0°, respectively. For $\Theta = 45°$, the estimation error stays below 15° on the examined range. With increasing training range, the prediction error for larger rotation angles increase. This is especially obvious for rotations $> +35°$ for $\Theta = 60°$.

The 90% quantile visualizes the accuracy of the determined rotation angles. For $\Theta = 45°$, up to ±30°, the limits of the quantile are close to the median. In this area it also remains within the ±15° range. For rotation angles that exceed 30°, the deviations of the prediction increase. This widens the 90% quantile. Increasing the angular range Θ, from which the training data is taken, does not improve the predictions. In contrary, the estimates differ more from their actual value.

C. Application on Public Finger Vein Data Sets

Both data sets used in section VI-B, PMMDB and PLUSVein-FR, have been acquired using the same capturing device following the same acquisition protocol. Therefore, the data sets (with exception of the acquired subjects) are very similar. To show that the trained model is not limited to a specific data set or sensor, it is also applied on four publicly available data sets, namely SDUMLA-HMT, FV-USM, UTFVP and PLUSVein-FV3, that were acquired using different capturing devices and acquisition protocols. Experiments are conducted to show the positive effect of rotation correction on the EERs for the four public datasets. The performance results are evaluated using the best performing model of section VI-B ($\Theta = 45°$).

1) Evaluation Protocol: As in [22], the data sets are corrected on the basis of the estimated rotation angles. As a result, all images of a finger should be aligned with each other with respect to their longitudinal rotation. The results are compared to those of the original data set (ORI) and the results achieved in [22]. For the experiments two corrected data sets are generated: In the first version (ROT), all samples are corrected with respect to the first sample of each finger, in the second one (ROT Mean), all samples of a finger are corrected with respect to the calculated mean rotation angle of this finger. In real world applications, the rotation correction differs from the two approaches mentioned above (ROT, ROT Mean). There, it is not clear if the enrolment and probe sample are actually from the same finger/subject. As a result of this, the rotation estimation and correction needs to be executed prior to every comparison (regardless of whether it is a genuine or impostor comparison).

TABLE I
NUMBER OF COMPARISONS FOR EACH DATA SET

Name	Genuine	Impostor	Total
SDUMLA-HMT	9540	200340	209880
UTFVP	2160	63720	65880
FV-USM	32472	120048	152520
PLUSVein-FV3	3600	63720	67320

To quantify the performance of the data sets, the EER is used. The experiments follow the test protocol of the FVC2004 [15]: For calculating the genuine scores, all possible genuine comparisons are performed. For calculating the impostor scores, only the first sample of a finger is compared against the first sample of all other fingers. The resulting numbers of comparisons for all data sets are listed in table I. To quantify the change of the performance, the relative performance increase (RPI) as stated in Eq. (6) is used.

$$RPI = \frac{EER_{ref} - EER_x}{EER_x} \tag{6}$$

EER_{ref} is the EER of the reference data set and EER_x the EER of the evaluated data set.

2) Recognition Tool-Chain: The finger vein recognition tool-chain consists of the following components:

1) The ROIs are extracted as described in section IV
2) The rotation angle between two vein images is estimated using the CNN rotation estimator in section VI-B (training range $\Theta = 45°$)
3) Rotated versions of the input images are generated as described in section V
4) To improve the visibility of the vein pattern *High Frequency Emphasis Filtering* (HFE) [34], *Circular Gabor Filter* (CGF) [33] and simple *CLAHE* (local histogram equalisation) [35] are used as *pre-processing* steps.

TABLE II
RECOGNITION PERFORMANCE OF THE EVALUATED DATA SETS AND ITS CORRECTED VERSIONS: ORI = ORIGINAL DATA SET, ROT = ROTATION CORRECTED TO 1ST IMAGE, ROT MEAN = ROTATION CORRECTED TO MEAN OF FINGER. BEST ACHIEVED EER AND RPI VALUES ARE HIGHLIGHTED IN BOLD.

Data Set	Correction	Method	Performance Indicators EER [%]	RPI [%]
SDUMLA-HMT	ORI	-	4.73	-
	ROT	proposed	1.30	263.40
		ICB'19 [22]	**1.07**	**341.59**
	ROT Mean	proposed	1.37	244.77
		ICB'19 [22]	1.14	315.85
FV-USM	ORI	-	1.23	-
	ROT	proposed	**0.52**	**137.03**
		ICB'19 [22]	0.56	120.06
	ROT Mean	proposed	0.76	61.89
		ICB'19 [22]	0.77	59.38
UTFVP	ORI	-	0.42	-
	ROT	proposed	0.18	125.47
		ICB'19 [22]	0.19	124.53
	ROT Mean	proposed	0.19	115.54
		ICB'19 [22]	**0.09**	**349.06**
PLUSVein-FV3	ORI	-	0.08	-
	ROT	proposed	**0.05**	**61.23**
		ICB'19 [22]	0.06	50.00
	ROT Mean	proposed	0.06	52.12
		ICB'19 [22]	0.08	0.94

5) As *feature extraction* method, the well-established vein-pattern based *Maximum Curvature* method (MC) [18] is employed.
6) The *comparison* of the binary feature images is done using a correlation measure, calculated between the input images and in x- and y-direction shifted versions of the reference image as described in [17].

3) Results: Table II holds the performance results of the proposed method as well as for the unmodified data set (ORI) and the results achieved in [22] for all four data sets. Both corrected data sets outperform the original data set in all four cases. For all data sets, the correction with respect to the first sample achieves slightly better results. The highest performance increase is achieved for SDUMLA-HMT. There the EER dropped from 4.73% (ORI) to 1.3% (ROT), which corresponds to a RPI of 263%. For the three other data sets the performance increased as well, but not to the same extent. For FV-USM the performance increased by 137%, for UTFVP by 125% and for PLUSVein-FV3 by 61%, respectively. These results essentially correspond with those achieved in [22]. SDUMLA-HMT shows the biggest performance differences. Such a result was to be expected since (according to [22]) this data set contains the largest rotation angles, including rotations up to 45°. According to the results of section VI-B3, the accuracy of the predicted rotations decreases noticeable for rotations above $\pm 30°$.

TABLE III
MEAN AND STANDARD DEVIATION OVER THE ABSOLUTE VALUED DIFFERENCES OF THE PREDICTED ROTATION ANGLES OF THE PROPOSED SYSTEM TO THE RESULTS IN [22]

Name	Mean	Standard Deviation
SDUMLA-HMT	3.07°	3.87°
UTFVP	2.18°	2.04°
FV-USM	1.78°	2.00°
PLUSVein-FV3	1.32°	1.35°

The main advantage of the proposed rotation detector is its applicability in real world finger vein recognition systems. The detection of one rotation angle requires only a single forward pass of the proposed CNN which takes in average 15ms (GPU: GeForce GTX Titan X). The approach in [22] needs 91 comparisons (2.621ms per comparison or 238ms in total) if all 91 rotated versions of each image are already available in storage.

Table III shows the mean and the standard deviation for the difference between the predictions of the proposed approach and the estimated rotation angles of [22]. The mean deviation is below 3.1° for all four databases, the standard deviation is never higher than 4°. The low differences between the rotation angles of both independent approaches imply that the predicted angles should be fairly accurate estimates. Therefore, also the majority of the errors in the prediction will be below $\pm 15°$, which, according to [19], can be compensated by commonly used finger vein recognition systems.

VII. CONCLUSIONS

In this article, a CNN-based rotation detector for finger vein biometrics was presented. The detector accurately estimates the longitudinal rotation between two finger vein images of the same finger and is not limited on the data sets it was trained on. The prediction of one rotation angle is very fast (one estimation takes approximately 15ms on a GPU system). This makes the proposed detector the first system that can be practically applied in finger vein recognition systems.

The first part of the experiments analysed the accuracy of the estimated rotation angles. The results showed, that the rotation detector delivers accurate results in the range of particular interest ($\pm 30°$). For rotation angles $>30°$, the estimation error rises noticeable.

To show that the system is not limited to the data set it was trained on, it was applied on four publicly available finger vein data sets, that differ from the training data (different capturing devices and acquisition protocols). A rotation correction using the estimated rotation angles leads to distinct improvements in the EER on all four data sets between 50 and 260% compared to the performance without rotation correction. The only prerequisite to apply the proposed system is that the ROIs of the finger vein images are extracted in the same way as during the training of the CNN.

REFERENCES

[1] M. S. M. Asaari, S. A. Suandi, and B. A. Rosdi. Fusion of band limited phase only correlation and width centroid contour distance for finger

based biometrics. *Expert Systems with Applications*, 41(7):3367–3382, 2014.

[2] J. Y. Bok, K. H. Suh, and E. C. Lee. Detecting fake finger-vein data using remote photoplethysmography. *Electronics*, 8(9):1016, 2019.

[3] S. Bunda. 3D point cloud reconstruction based on the finger vascular pattern. B.S. thesis, University of Twente, 2018.

[4] Q. Chen, L. Yang, G. Yang, and Y. Yin. Geometric shape analysis based finger vein deformation detection and correction. *Neurocomputing*, 2018.

[5] Q. Chen, L. Yang, G. Yang, Y. Yin, and X. Meng. DFVR: Deformable finger vein recognition. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1278–1282, March 2017.

[6] X. Deng, Y. Yuan, Y. Zhang, P. Tan, L. Chang, S. Yang, and H. Wang. Joint hand detection and rotation estimation by using CNN. *CoRR*, abs/1612.02742, 2016.

[7] Houjun Huang, S. Liu, H. Zheng, L. Ni, Yi Zhang, and W. Li. Deepvein: Novel finger vein verification methods based on deep convolutional neural networks. In *2017 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA)*, pages 1–8, Feb 2017.

[8] H. Hsu, T. Wu, S. Wan, W. H. Wong, and C. Lee. Quatnet: Quaternion-based head pose estimation with multiregression loss. *IEEE Transactions on Multimedia*, 21(4):1035–1046, 2019.

[9] B. Huang, Y. Dai, R. Li, D. Tang, and W. Li. Finger-vein authentication based on wide line detector and pattern normalization. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 1269–1272. IEEE, 2010.

[10] W. Kang, H. Liu, W. Luo, and F. Deng. Study of a full-view 3D finger vein verification technique. *IEEE Transactions on Information Forensics and Security*, pages 1–1, 2019.

[11] C. Kauba, B. Prommegger, and A. Uhl. Focussing the beam - a new laser illumination based data set providing insights to finger-vein recognition. In *Proceedings of the IEEE 9th International Conference on Biometrics: Theory, Applications, and Systems (BTAS2018)*, pages 1–9, Los Angeles, California, USA, 2018.

[12] A. Kumar and Y. Zhou. Human identification using finger images. *Image Processing, IEEE Transactions on*, 21(4):2228–2244, 2012.

[13] E. C. Lee, H. C. Lee, and K. R. Park. 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.

[14] Y. Lu, S. J. Xie, S. Yoon, J. Yang, and D. S. Park. Robust finger vein roi localization based on flexible segmentation. *Sensors*, 13(11):14339–14366, 2013.

[15] D. Maio, D. Maltoni, R. Cappelli, J. L. Wayman, and A. K. Jain. FVC2004: Third Fingerprint Verification Competition. In *ICBA*, volume 3072 of *LNCS*, pages 1–7. Springer Verlag, 2004.

[16] Y. Matsuda, N. Miura, A. Nagasaka, H. Kiyomiu, and T. Miyatake. Finger-vein authentication based on deformation-tolerant feature-point matching. *Machine Vision and Applications*, 27(2):237–250, 2016.

[17] N. Miura, A. Nagasaka, and T. Miyatake. Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification. *Machine Vision and Applications*, 15(4):194–203, 2004.

[18] N. Miura, A. Nagasaka, and T. Miyatake. Extraction of finger-vein patterns using maximum curvature points in image profiles. *IEICE transactions on information and systems*, 90(8):1185–1194, 2007.

[19] B. Prommegger, C. Kauba, M. Linortner, and A. Uhl. Longitudinal finger rotation - deformation detection and correction. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 1(2):123–138, 2019.

[20] B. Prommegger, C. Kauba, and A. Uhl. Longitudinal finger rotation - problems and effects in finger-vein recognition. In *Proceedings of the International Conference of the Biometrics Special Interest Group (BIOSIG'18)*, Darmstadt, Germany, 2018.

[21] B. Prommegger, C. Kauba, and A. Uhl. Multi-perspective finger-vein biometrics. In *Proceedings of the IEEE 9th International Conference on Biometrics: Theory, Applications, and Systems (BTAS2018)*, Los Angeles, California, USA, 2018.

[22] B. Prommegger, C. Kauba, and A. Uhl. On the extent of longitudinal finger rotation in publicly available finger vein data sets. In *Proceedings of the 12th IAPR/IEEE International Conference on Biometrics (ICB'19)*, pages 1–8, Crete, Greece, 2019.

[23] B. Prommegger and A. Uhl. Perspective multiplication for multi-perspective enrolment in finger vein recognition. In *Proceedings of the 18th International Conference of the Biometrics Special Interest Group (BIOSIG'19)*, Darmstadt, Germany, 2019.

[24] B. Prommegger and A. Uhl. Rotation invariant finger vein recognition. In *Proceedings of the IEEE 10th International Conference on Biometrics: Theory, Applications, and Systems (BTAS2019)*, Tampa, Florida, USA, 2019.

[25] R. Raghavendra and C. Busch. Exploring dorsal finger vein pattern for robust person recognition. In *2015 International Conference on Biometrics (ICB)*, pages 341–348, May 2015.

[26] P. Schuch, J. M. May, and C. Busch. Unsupervised learning of fingerprint rotations. In *2018 International Conference of the Biometrics Special Interest Group (BIOSIG)*, pages 1–6, 2018.

[27] A. F. Sequeira, J. Ferryman, L. Chen, C. Galdi, J.-L. Dugelay, V. Chiesa, A. Uhl, B. Prommegger, C. Kauba, S. Kirchgasser, A. Grudzien, M. Kowalski, L. Szklarski, P. Maik, and P. Gmitrowicz. Protect multimodal db: a multimodal biometrics dataset envisaging border control. In *Proceedings of the International Conference of the Biometrics Special Interest Group (BIOSIG'18)*, pages 1–8, Darmstadt, Germany, 2018.

[28] L. Sonna Momo, L. Cerqueira Torres, S. Marcel, A. Anjos, M. Liebling, A. Shajkofci, S. Amoos, A. Woeffray, A. Sierro, P. Roduit, P. Ferez, and L. Bonvin. Method and Device for Biometric Vascular Recognition and/or Identification, WIPO (PCT) Patent WO/2019/150254, 08 2019.

[29] B. Ton and R. Veldhuis. A high quality finger vascular pattern dataset collected using a custom designed capturing device. In *International Conference on Biometrics, ICB 2013*. IEEE, 2013.

[30] S. Xie, R. B. Girshick, P. Dollár, Z. Tu, and K. He. Aggregated residual transformations for deep neural networks. *CoRR*, abs/1611.05431, 2016.

[31] L. Yang, G. Yang, Y. Yin, and X. Xi. Finger vein recognition with anatomy structure analysis. *IEEE Transactions on Circuits and Systems for Video Technology*, pages 1–1, 2017.

[32] Y. Yin, L. Liu, and X. Sun. Sdumla-hmt: a multimodal biometric database. *Biometric Recognition*, pages 260–268, 2011.

[33] J. Zhang and J. Yang. Finger-vein image enhancement based on combination of gray-level grouping and circular gabor filter. In *Information Engineering and Computer Science, 2009. ICIECS 2009. International Conference on*, pages 1–4. IEEE, 2009.

[34] J. Zhao, H. Tian, W. Xu, and X. Li. A new approach to hand vein image enhancement. In *Intelligent Computation Technology and Automation, 2009. ICICTA'09. Second International Conference on*, volume 1, pages 499–501. IEEE, 2009.

[35] K. Zuiderveld. Contrast limited adaptive histogram equalization. In P. S. Heckbert, editor, *Graphics Gems IV*, pages 474–485. Morgan Kaufmann, 1994.