Deep Learning based Age and Gender Recognition applied to Finger Vein Images*

Georg Wimmer*[0000−0001−5529−0154], Bernhard Prommegger[0000−0003−4002−2602], and Andreas Uhl[0000−0002−5921−8755]

Department of Artificial Inteligence and Human Interfaces, University of Salzburg, Jakob-Haringer-Strasse 2, 5020 Salzburg, Austria
{gwimmer,bprommeg,uhl}@cs.sbg.ac.at

Abstract. Finger vein recognition deals with the identification of subjects based on their venous pattern within the fingers. It was shown in previous work that biometric data can include more than only identity related information like e.g. age, gender and ethnicity. In this work, deep learning based methods are employed to find out if finger vein image data includes information on the age and gender of the subjects. In our experiments we use different CNNs and different loss functions (triplet loss, SoftMax loss and Mean Squared Error loss) to predict gender and age based on finger vein image data. Using three publicly available finger vein image datasets, we show that it is feasible to predict the gender (accuracies of up to 93.1%). By analyzing finger vein data from different genders we found out that the finger thickness and especially the total length over all finger veins are important features to differentiate between images from male and female subjects. On the other hand, estimating the age of the subjects hardly worked at all in our experiments.

Keywords: finger vein recognition, age, gender

1 Introduction

Biometric systems extract identity related information to use it for person recognition or identification. However, biometric data can include more than only identity related information. For example, humans can deduce a lot of information from images of the face of a person like age, gender and ethnicity.

Previous works predicted both the age and gender of subjects based on face images [18, 29], voice recordings [15], gait [21], ear [28] and also using less intuitive biometric traits for age and gender prediction like finger prints [6] and EEG recordings [13]. Furthermore, iris (ocular) images have been used to predict age [19] and gender [23, 1]. Reviews about the demographic bias in biometrics are presented in [22, 5].

* Corresponding author: Georg Wimmer. 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.

In this work we analyze the demographic bias in biometric finger vein image data with respect to the age and gender of the subjects. Research on soft-biometrics showed that privacy-sensitive information can be deduced from biometric templates of an individual. Since these templates are expected to be used for recognition purposes only, this raises major privacy issues. We aim to find out if these privacy issues also apply for finger vein data with regard to age and gender information.

Two medical studies [14, 7], showed that there is no noticeable influence of age on the size of the veins but the size of the veins is in general bigger for men as for woman.

In [4], the authors claim that common finger vein recognition systems are not able to recognize age and gender from finger vein images. However, the considered finger vein recognition systems were purely trained for subject recognition and not to recognize the age or gender of the subjects. There are also previous publications that developed methods specifically for age and gender recognition based on vein images. However, these publications either use a heavily biased experimental setup and/or show serious errors in the experimental setup by speaking of age and gender recognition while actually doing subject recognition:

- [2, 3]: In these two papers from the same authors Local Binary Pattern (LBP) operators are employed together with a nearest neighbor classifier to estimate the age and gender based on finger vein [2] and palm vein [3] images. What the authors actually did is to assign an image the age/gender of its nearest neighbor. The systematical error in the experimental setup is that they did not exclude comparisons between images from the same finger and subject. As the nearest neighbor of nearly each image is an image from the same finger, these two papers do misrepresent a subject recognition as an age/gender recognition.
- [25]: In this publication a 2-layer network combined with a linear SVM classifier is applied for gender recognition of hand dorsal vein images. In the publication there is no splitting in training and evaluation data (same data is used for training and evaluation), which heavily biases the experiments and makes the results unusable.

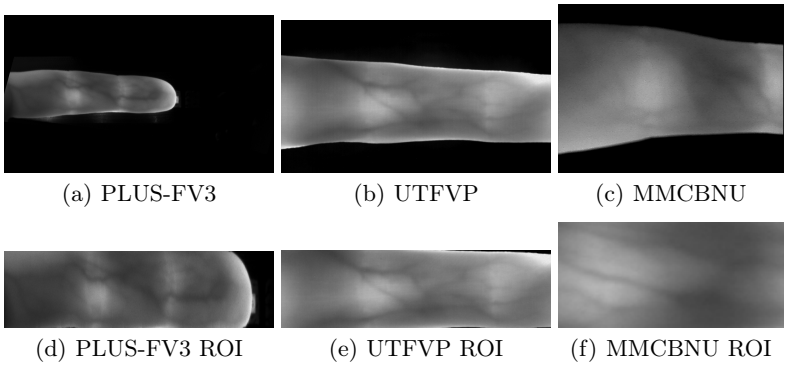
So this is the first paper for age and gender recognition based on finger vein images with methods specifically trained to predict the age and gender of the subjects without a heavily biased or incorrect experimental setup. In the experiments we employ CNN-based methods for gender and age recognition on three publicly available finger vein image datasets. The CNNs are trained using various loss functions that are suited for age respectively gender recognition. Furthermore, this is the first paper that analyses what features are really different between male and female finger vein images.

2 Databases

In the experiments, three publicly available finger vein datasets are employed. All of them provide - next to the vein images - also information on age and gender

Database	Subjects	Instances	Samples
PLUS-FV3	76	456	2268
UTFVP	60	360	1440
MMCBNU	100	600	6000

Table 1. The number of subjects, instances and samples of each of the three employed finger vein image datasets.



(a) PLUS-FV3 (b) UTFVP (c) MMCBNU

(d) PLUS-FV3 ROI (e) UTFVP ROI (f) MMCBNU ROI

Fig. 1. Exemplar finger vein images of the three datasets. In the top row we show the original finger vein image and in the bottom row the ROI extracted versions of the images above.

of the acquired subjects. The first dataset is a combination of the PLUSVein-FV3 [12] and the PROTECT Multimodal DB [8] further donated as PLUS-FV3. Due to missing age and gender information, two out of the 78 subjects have been omitted. The other datasets are the UTFVP [24] and the MMCBNU [16]. The number of subjects, instances and image samples are listed in Table 1. The experiments are applied to the original finger vein images as well as to images that are reduced to the region of interest (ROI) using a ROI extraction technique based on [17]. In Figure 1, example images (original ones as well as ROI extracted version of the images) are presented for each dataset.

3 Methods

Since age and gender recognition are quite different tasks to handle we employ different CNN-based approaches with different loss functions for the two tasks which also results in different performance measures for different approaches. Since the datasets are mostly quite inbalanced with respect to the number of images per gender (see Table 2) and the age distribution among the subjects (see Figure 2, black bars), the CNN training is applied using balanced batch sampling (each batch contains the same number of images per class) to mitigate the bias in the data. For all experiments, the CNNs are trained using 2-fold cross validation. The CNNs are trained on one fold and evaluation is applied on the other one and then vice versa. The images of half of the subjects are in one fold and the

images of the remaining subjects are in the other fold. All CNNs are trained for 400 epochs using nets that were pretrained on the ImageNet database. Data augmentation is applied by first resizing the input images to a size of 229×229 and then extracting a patch of size 224×224 at a random position of the resized image (± 5 pixels in each direction). As performance measures we report the mean accuracy over the two folds and partly also the mean Equal Error Rate (EER) over the two folds.

3.1 Methods for gender recognition

To train a CNN to recognize the gender based on finger vein images, each image of the training fold is assigned to the class (male or female) of the respective subject. In that way the CNN can learn similarities between images of the same class (gender) and differences between images of different classes. For more meaningful results and to see if different methods produce different results, two quite different CNN approaches are applied with very different loss fuctions for CNN training and also different CNN architectures:

1. Triplet loss: Same as in a previous publication on finger vein recognition [26], we employ the triplet loss function for CNN training. Per training step, the triplet loss requires three input images at once (a so called triplet), where two images belong to the same class (the anchor image and a sample from the same class, further denoted as positive) and the third belongs to a different class (further denoted as negative). The triplet loss learns the network to minimize the distance between the anchor and the positive and maximize the distance between the anchor and the negative. The triplet loss using the squared Euclidean distance is defined as follows:

$$L(A, P, N) = \max(||f(A) - f(P)||^2 - ||f(A) - f(N)||^2 + \alpha, 0), \quad (1)$$

   where A is the anchor, P the positive and N the negative. α is a margin that is enforced between positive and negative pairs and is set to $\alpha = 1$. $f(x)$ is an embedding (the CNN output of an input image x). So the CNN is trained so that the squared distances between all embeddings of finger vein images from the same class (gender) is small, whereas the squared distance between embeddings of any pairs of finger vein images from different classes is large. As CNN architecture we employ the SqueezeNet [10], a neural networks with low memory requirements. The size of the last layers convolutional filter is adapted so that a 256-dimensional output (embedding) is produced. As first performance measure we compute the EER, where the similarity score between two finger vein images is defined as the inversed euclidean distances between the CNN outputs of two images. As second performance measure we compute the (Rank-1) accuracy. Since CNNs that were trained with the triplet loss produce an feature vector (embedding) for each input image instead of a class prediction (like CNNs trained with the the SoftMax loss), a classifier is required to obtain gender predictions. Ftor this we employ

a linear SVM to predict the gender based on the CNN outputs. Same as the CNNs, the SVM is trained on the training fold while evaluation is applied on the images of the evaluation fold.
2. SoftMax loss: Another obvious choice for this 2-class gender classification problem is to employ the widely known SoftMax loss function to train a CNN. As net architecture we employ the DenseNet-161 [9]. Since the CNN predicts gender directly without generating any feature vector output per image, we only report the accuracy and not the EER, for which distances between images would be required.

As can be observed in Table 2, only the PLUS-FV3 dataset is somehow balanced with respect to the gender distribution of the subjects. The other two datasets consist of distinctly more male subjects than female ones. So for a better overview on the CNN classification outcomes, we not only report the accuracy over the whole dataset but also the percentage of images from female subjects that were correctly classified as female (ACC women) and the percentage of images from male subjects that were correctly classified as male (ACC men).

The gender recognition experiments are applied to the original images as well as to the ROI extracted images. The original images have the advantage that the images contain information on the length and thickness of the fingers, while the ROI extracted images have the advantage of a higher image resolution since the images need to be resized to the required CNN input size (224×224) and by downsizing only the part of the image containing the finger the resolution is higher than by downsizing the whole image.

3.2 Methods for age recognition

For all age recognition experiments, we only employ the ROI extracted images. This is done because the length and thickness of the fingers does not matter for age recognition. Age recognition will probably be a more difficult task than gender recognition, since there are no known differences in the vein structure depending on the age of the subjects (see [14]). Because of that, we do not directly start with a direct estimation of the age but firstly conduct an experiment to find out if CNNs are at least able to discriminate between finger vein images of young and old subjects. For this, we divide the datasets in two classes, with one class consisting of all images of subjects under 25 years and the other class of all images of subjects over 45 years. The images of all subjects with an age between 25 and 45 years are removed for this experiment. If the CNNs are not even capable to differentiate between these two age groups, then all further experiments to estimate the age of subjects based on finger vein images would probably be futile. For this experiments we employ the same methods (SqueezeNet trained with triplet loss and DenseNet trained with SoftMax loss) as in the gender recognition experiment. Similar to the gender recognition experiment, the datasets are quite unbalanced with respect to the number of images per class and hence we once again not only report the overall accuracy but also the percentage of correctly classified images of the class containing all subjects under 25 years age and of the class containing all subjects above 45 years age.

In a second and more difficult experiment, we aim to directly estimate the age of the subjects using a CNN trained with the mean squared error (MSE) loss function. For this experiment we use all images of the datasets contrary to the previous one. As net architecture we employ ResNeXt101 [27]. For training, each finger vein image of the training fold is assigned to the age of the according subject so that the CNN can learn to estimate the age based on finger vein images. For the images of the evaluation fold, an age estimation is made using the trained CNN. Then the folds for training and validation are exchanged. In this experiments we do not classify images like in the previous experiments but directly estimate the age of the subjects from the images. Thus, we need to apply different performance measures in this experiment and report the mean absolute error (MAE) over all images:

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |y_i - \hat{y}_i|, \tag{2}$$

where N is the number of images of the dataset, $\hat{y}_i$ the CNN's age prediction of image i and y_i the actual age of the image's subject. Furthermore, we divide the images into several age groups with an age group having a range of 10 years. Then we compare the actual number of images per age group with the number of images that are correctly and incorrectly predicted to the considered age group. Additionally, we present the mean over the CNN age predictions over all images per age group.

4 Results

4.1 Gender recognition

In Table 2, we present the gender recognition results using the SoftMax loss trained CNN (along with the number of images per gender) and in Table 3 the results using the triplet loss trained CNN. For the triplet loss trained net we report the EER additional to the accuracy. We can observe that on all three datasets accuracies between about 79 and 93 % are achieved for both kinds of CNNs. In general, the accuracies for the two kinds of CNNs are similar. For the PLUS-FV3 and UTFVP dataset, the classification rates are between about 76% and 87% for images from female subjects and between 84% and 97% for images from male subjects. In case of the MMCBNU dataset, the inbalance between the number of male and female subjects is huge (nearly 5 times as much men as women) which probably leads to the fact that most images are predicted to be male by the CNNs, even the images from women. For the ROI images of the MMCBNU dataset, about 50% of the images from female subjects are classified correctly but for the original images only 10-15% of the images from female subjects are classified correctly. In the discussion we will analyze if the inbalance between the number of images from male and female subjects is actually the reason that most of the images of the MMCBNU dataset are predicted to be male.

Databases	Images per gender		ACC		
	women	men	overall	women	men
PLUS-FV3 (ROI)	951	1317	81.7	76.4	86.0
MMCBNU (ROI)	1020	4980	87.5	49.3	93.4
UTFVP (ROI)	384	1056	90.5	78.8	96.3
PLUS-FV3 (Orig.)	951	1317	86.1	84.0	88.2
MMCBNU (Orig.)	1020	4980	79.3	10.1	93.5
UTFVP (Orig.)	384	1056	92.2	82.1	97.1

Table 2. Gender classification results (accuracy (ACC) in %) using SoftMax loss trained CNNs

Databases	EER	ACC		
		all	women	men
PLUS-FV3 (ROI)	34.2	80.3	73.3	84.0
MMCBNU (ROI)	27.1	88.6	51.4	96.1
UTFVP (ROI)	18.3	92.1	83.6	96.4
PLUS-FV3 (Original)	29.4	84.4	78.3	89.5
MMCBNU (Original)	46.9	80.3	14.6	93.9
UTFVP (Original)	13.8	93.1	86.5	96.7

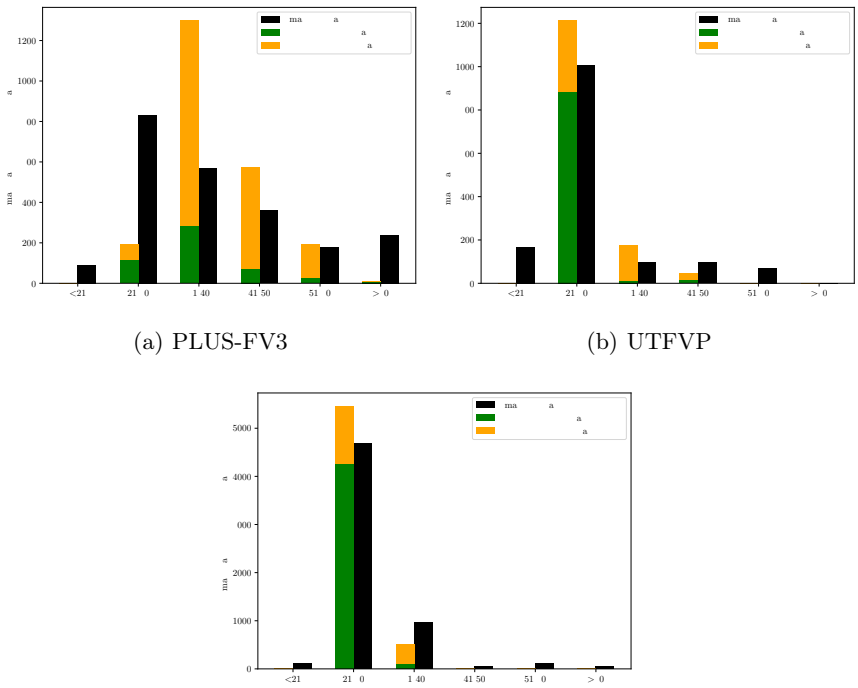
Table 3. Gender recognition (EER in %) and classification results (ACC in %) using triplet loss trained CNNs

In general, it does not matter whether we use the original finger vein images or the ROI extracted images with respect to the results, except for the MMCBNU dataset.

4.2 Age recognition

2-class age recognition experiment: In Table 4 we present the results for the 2-class age recognition experiment, where all images of subjects under 25 years of age are assigned to one class and all images of subjects over 45 years are assigned to the other one. On all three datasets, the images from the class with fewer images are misclassified clearly more often than the images from the class with more images. The differences in the results of the two classes are even more pronounced than in the gender recognition experiments. For the MMCBNU dataset, nearly all images are predicted to belong to the class with distinctly more samples (< 25). We can observe that at least for the PLUS-FV3 and the UTFVP dataset, the prediction of the age class is clearly better than random assignment with classification rates of up to nearly 90% for the UTFVP dataset and up to 76% for the PLUS-FV3 dataset.

Age estimation experiment: In Figure 2 we present the number of images per age group for all three datasets as well as outcomes of the age recognition experiment for images at different age groups. The age of the subjects is estimated using the CNN trained with the MSE loss. Each age group except the first



(a) PLUS-FV3

(b) UTFVP

(c) MMCBNU

Fig. 2. Actual number of images per age group (black bars) compared to the number of CNN age predictions that are correctly (green bar) and incorrectly (orange bar) assigned to an age group

(till 20) and last one (over 60) has a range of 10 years. The black bar shows the number of images per age group, the green bar the number of images that are correctly predicted to the considered age group and the orange bar the number of images that are incorrectly predicted to the considered age group. So if we add up the numbers of the orange and green bar for an age group, then we get the number of images whose CNN age estimate is within the range of the age group under consideration. In Figure 2 we can observe that only the PLUS-FV3 dataset has a at least somehow balanced distribution of subjects across the different age groups. The other two datasets mainly consist of subjects at an age between 21 and 30. We can observe that most of the images on all datasets are predicted to the age group that contains the mean age over a dataset. In case of the two datasets UTFVP and MMCBNU, this is also the age group that contains the majority of image samples. In case of the PLUS-FV3 dataset, where the images are distributed over several different age groups, most images are predicted to the wrong age group (31-40 years).

In Table 5, we present the CNN results on all three datasets using the Mean Absolute Error (MAE) in years as performance measure. To get an understand-

Database	Nr. of images		CNN results with Triplet loss				CNN results with SoftMax		
	<25	>45	EER	ACC overall	ACC<25	ACC>45	ACC overall	ACC<25	ACC>45
PLUS-FV3	297	657	40.3	76.0	63.3	83.9	74.0	58.9	85.1
MMCBNU	1440	240	51.0	83.6	97.2	6.7	85.1	99.4	0.0
UTFVP	672	167	26.4	89.7	99.4	59.7	87.4	98.8	50.7

Table 4. CNN results for grouping the images in two classes, where one class comprises all images below an age of 25 years and the other one all images above 45 years. For the triplet loss trained CNNs we present the accuracy and the EER in %, for the SoftMax loss trained CNNs we only present the accuracy. Furthermore, we present the number of images per class.

Databases	Mean age	MAE		Mean CNN prediction per age group					
		CNN	MP	0-20	21-30	31-40	41-50	51-60	>60
PLUS-FV3	38.5	11.4	11.5	33.4	36.2	38.8	37.2	41.6	44.7
MMCBNU	27.7	4.5	4.5	24.9	27.0	26.8	28.8	25.9	26.6
UTFVP	28.4	5.9	6.2	25.1	26.4	25.9	32.1	31.9	-

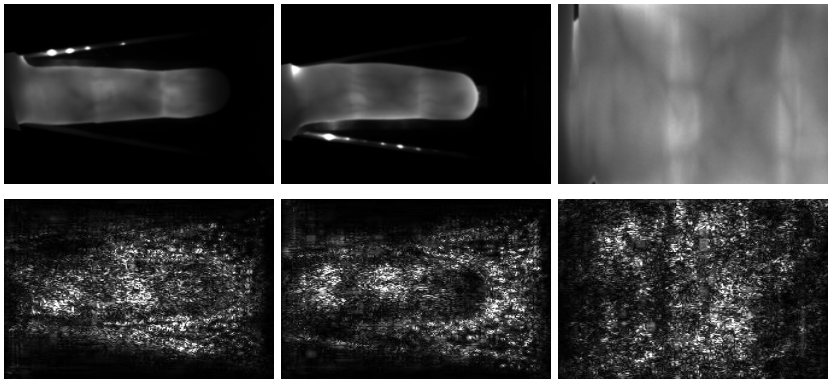
Table 5. CNN age prediction results using the MSE-loss. The left column shows the mean age per dataset and the middle columns the Mean Absolute Error (MAE) of the CNN and of the method that simply predicts the mean age for all samples (MP). The right side columns present the mean over the CNN age predictions over all images per age group

ing if the MAE results of the proposed CNN age prediction method are good or rather not, we compare the MAE of the CNN with the MAE of a method denoted as 'mean prediction' (MP). This method simply assigns the mean age (average age over the ages of all subjects from a dataset) to each image of the dataset. In addition, we present the mean over the CNN age estimates separately for the images of each age group. For example, if we consider the age group '0-20', it means that we present the average over all CNN age estimates from images of subjects between the ages of 0 and 20 years. We can observe in Table 5 that the MAE of the CNN is only slightly better or equal as the MAE for predicting the mean age for each image of the dataset (MP). So the age estimation using our CNN does not work well.

When we observe the mean CNN predictions per age group on the PLUS-FV3 dataset (the two other datasets mainly consist of images from subject with an age between 21 and 30, which limits the information we can extract from the results of these two datasets), then we can see that the CNN age predictions on finger vein images from subjects of higher age are indeed higher as the age predictions on images of subjects from lower age. However, the CNN predictions are clearly too close to the mean age over the dataset (38.5) and hence too high for images of younger subjects and too low for images of older subjects.

5 Discussion

In this section we aim to find out which features enable the CNNs to discriminate between genders using a gradient based CNN visualization technique. Further-



(a) Female Original (b) Male Original (c) Female ROI

Fig. 3. CNN visualizations using Vanilla backpropagation of male and female finger vein images using original as well as ROI images. Top row: original image, bottom row: saliency map

more, we want to find which differences actually do exist between male and female finger vein images by analyzing segmentation masks of the fingers and manual extracted segmentations of the finger vein structure. In addition, we want to find the cause for the poor gender recognition result on the MMCBNU dataset. Age recognition clearly did not work and hence there is no need to discuss the results any further.

The idea behind gradient based CNN visualization techniques is to compute the gradient of the network's prediction with respect to the input, holding the weights fixed. This determines which pixels of an input image need to be changed the least to affect the prediction the most. In this work we employ gradient visualization using Vanilla Backpropagation [20] to get the saliency maps of images. With these saliency maps we can measure the relative importance of each pixel to the ultimate prediction by the model. In Figure 3 we present the saliency maps of three different finger vein images from the PLUS-FV3 dataset (two original images and one ROI image) using a CNN that was trained to predict the gender with the SoftMax loss function. From the saliency maps of the two original images (one of a man's finger and one of a woman's finger) we can observe that image regions from the background that are surrounding the finger have an impact on the CNN predictions as high as image regions within the finger. This indicates that the finger thickness and length are features used by the CNN to determine the gender. By segmenting the fingers on the UTFVP dataset (using the segmentation method for the ROI extraction) and computing the finger thickness, we found that the fingers of male subjects are in average 15% thicker than those of female subjects. For the ROI image, the saliency map is fairly uniform across the regions of the finger without any clear visible correspondence to the vein pattern, so no conclusion can be drawn about what features are important for the CNN to predict the gender.

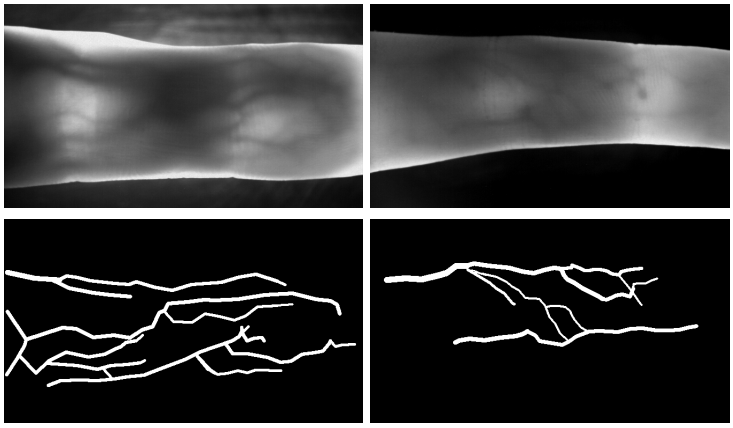


Fig. 4. Two example images of the UTFVP dataset (top row) and their manually extracted segmentation masks (bottom row) from the $UTFVP_{seg}$ dataset.

As already mentioned in the introduction, it was shown in a medical study that the cross section area of veins is in general bigger for men as for women. We now aim to verify if that is also the case for the finger veins using a dataset containing manually segmented veins from finger vein images. This dataset [11], further denotes as UTFVP$_{seg}$, contains 388 segmentation masks from images of the UTFVP dataset. 292 segmentation masks are from male subjects and 92 from female subjects. UTFVP$_{seg}$ includes at least one segmentation mask per finger of the UTFVP dataset, for some fingers it includes two. In Figure 4 we show two finger vein images of the UTFVP dataset and their segmentation masks from the UTFVP$_{seg}$ dataset. We can observe that the segmentation masks only cover the clearly visible finger veins but not the very fine ones that are hardly visible. In addition, the segmentation masks do not perfectly match the finger vein thickness, but the masks are still much better than for finger vein pattern extraction techniques like Maximum Curvature (MC) and Principle Curvature (PC), which do absolutely not reflect the actual vein thickness.

By analyzing the average vein thickness using the segmentation masks, we can find out if there are differences in the vein thickness depending on the gender of the subjects. This is done by first computing the size of the area of the finger vein structure by counting the number of pixels in the segmentation masks that are indicating a vein. Secondly, we compute the summed up length over all finger veins by applying skeletonization to the segmentation mask and summing up the number of pixels of the skeletonized finger vein segmentation mask. By dividing the area of the finger vein structure by its lengths we get the average finger vein thickness/diameter of a finger vein image. Now by averaging the average finger vein thickness over all images per gender we can find out if there are truly differences between men and women with respect to the finger vein thickness. It turned out that the mean vein diameter for finger vein images of men is 8.6

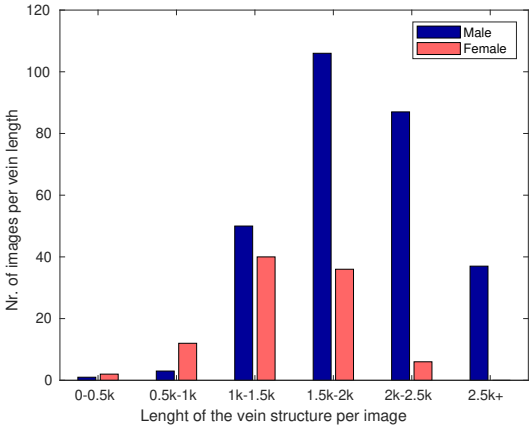


Fig. 5. Bar plot showing the distribution of the summed up length of the finger vein structure on the images of the UTFVP$_{seg}$ database for the two genders

pixels whereas the mean vein diameter for finger vein images of women is 8.5 pixels. So there is hardly any difference between the genders.

However, we found out that the summed up length over all finger veins of an image is for male subjects about 1.4 times higher in average than for female subjects. That means that there are clearly more well visible finger veins for men as for women. Figure 5 presents the distribution of the vein length on the UTFVP$_{seg}$ database of male and female subjects, where the length of the vein structure is defined as the sum of the pixels of the skeletonized segmentation mask. The higher length of the vein structure in finger vein images of male subjects could be one of the reasons that the CNNs were able to determine the gender based on finger vein images.

Finally, we want to find out if the inbalance of the MMCBNU dataset data with respect to the gender distribution (only 17% of the images are from female subjects) is the reason that so much images of female subjects were predicted as male. For this we apply a gender recognition experiment to a subset of the MMCBNU dataset consisting of all 17 female subjects but only 17 of the 83 male subjects (we simply chose the first 17 male subjects of the dataset). So, in total the sub dataset consists of 1020 images of female subjects and 1020 images from male subjects. The gender recognition experiments are employed exactly the same as for the original dataset, except using the smaller but balanced subset of the MMCBNU dataset. In Table 6 we present the results of this experiment.

As we can observe in Table 6, female images are predicted with about the same accuracy as male images on the balanced MMCBNU sub dataset. So the huge inbalance of the MMCBNU dataset with respect to the gender distribution was actually the reason that so many female images were classified incorrectly on the MMCBNU dataset.

MMCBNU	ACC triplet			ACC SoftMax		
	all	women	men	all	women	men
ROI	78.5	82.0	74.2	75.6	75.9	74.2
Original	53.4	50.4	57.1	51.5	49.8	54.3

Table 6. Gender classification results (ACC in %) on a balanced subset of the MMCBNU dataset (ROI as well as original images) using CNNs trained with the triplet loss (SVM results) and CNNs trained with the SoftMax loss

6 Conclusion

In our experiments on three public finger vein image datasets we showed that it is indeed feasible to predict the gender based on finger vein images. For the UTFVP dataset we achieved accuracies of up to 93.1%, for the PLUS dataset accuracies of up to 86.1% and for the MMCBNU dataset accuracies of up to 88.6%. However, it should be noted that the data sets used are rather small and unbalanced in terms of the number of images from male and female subjects and therefore predictions are more accurate for the larger group (male). In general, it did not matter which CNN architecture or loss function was used. As the results for original and ROI images (which contains no shape or background information) are similar, one can conclude that the finger region itself contains enough information to predict gender. However, we showed that the thickness and length of the fingers are important features for the CNNs to discriminate between the genders. Furthermore, there is a big difference between the genders with respect to the length of the finger vein structure. In average, the summed up length of the finger vein structure for male subjects is about 1.4 times higher than for female subjects.

The experiments to estimate the age based on finger vein images did not perform well. For two of the three datasets, it did work at least to some extent to differentiate between young (> 25) and old subjects (> 45), but on the MMCBNU dataset nearly all images were predicted as male. The direct estimation of the age using a MSE loss trained CNN did not work at all and only performed slightly better than simply predicting the mean age over all subjects for all images.

References

1. Bobeldyk, D., Ross, A.: Analyzing covariate influence on gender and race prediction from near-infrared ocular images. IEEE Access **7**, 7905–7919 (2019). https://doi.org/10.1109/ACCESS.2018.2886275
2. Damak, W., Boukhris Trabelsi, R., Damak Masmoudi, A., Sellami, D., Nait-Ali, A.: Age and gender classification from finger vein patterns. In: Madureira, A.M., Abraham, A., Gamboa, D., Novais, P. (eds.) Intelligent Systems Design and Applications. p. 811–820. Springer International Publishing, Cham (2017)
3. Damak, W., Trabelsi, R.B., Masmoudi, A.D., Sellami, D.: Palm vein age and gender estimation using center symmetric-local binary pattern. In: International Joint

Conference: 12th International Conference on Computational Intelligence in Security for Information Systems (CISIS 2019) and 10th International Conference on EUropean Transnational Education (ICEUTE 2019). pp. 114–123. Springer International Publishing, Cham (2020)
4. Drozdowski, P., Prommegger, B., Wimmer, G., Schraml, R., Rathgeb, C., Uhl, A., Busch, C.: Demographic bias: A challenge for fingervein recognition systems? In: 2020 28th European Signal Processing Conference (EUSIPCO). pp. 825–829 (2021). https://doi.org/10.23919/Eusipco47968.2020.9287722
5. Drozdowski, P., Rathgeb, C., Dantcheva, A., Damer, N., Busch, C.: Demographic bias in biometrics: A survey on an emerging challenge. CoRR **abs/2003.02488** (2020), https://arxiv.org/abs/2003.02488
6. Falohun, A., Deborah, F., Ajala, F.: A fingerprint-based age and gender detector system using fingerprint pattern analysis. International Journal of Computer Applications **136**, 975–8887 (03 2016). https://doi.org/10.5120/ijca2016908474
7. Gagne, P., Sharma, K.: Relationship of common vascular anatomy to cannulated catheters. International Journal of Vascular Medicine **2017** (2017)
8. Galdi, C., Boyle, J., Chen, L., Chiesa, V., Debiasi, L., Dugelay, J.L., Ferryman, J., Grudzień, A., Kauba, C., Kirchgasser, S., Kowalski, M., Linortner, M., Maik, P., Michoń, K., Patino, L., Prommegger, B., Sequeira, A.F., Łukasz Szklarski, Uhl, A.: Protect: Pervasive and user focused biometrics border project – a case study. IET Biometrics **9**(6), 297–308 (2020). https://doi.org/10.1049/iet-bmt.2020.0033
9. Gao, H., Liu, Z., Weinberger, K.Q.: Densely connected convolutional networks. CoRR **abs/1608.06993** (2016), http://arxiv.org/abs/1608.06993
10. Iandola, F.N., Moskewicz, M.W., Ashraf, K., Han, S., Dally, W.J., Keutzer, K.: Squeezenet: Alexnet-level accuracy with 50x fewer parameters and <1mb model size. CoRR **abs/1602.07360** (2016), http://arxiv.org/abs/1602.07360
11. Jalilian, E., Uhl, A.: Enhanced segmentation-cnn based finger-vein recognition by joint training with automatically generated and manual labels. In: Proceedings of the IEEE 5th International Conference on Identity, Security and Behavior Analysis (ISBA 2019). pp. 1–8. IDRBT (2019)
12. Kauba, C., Prommegger, B., Uhl, A.: 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). pp. 1–9. Los Angeles, California, USA (2018), accepted
13. Kaushik, P., Gupta, A., Roy, P.P., Dogra, D.P.: Eeg-based age and gender prediction using deep blstm-lstm network model. IEEE Sensors Journal **19**(7), 2634–2641 (2019). https://doi.org/10.1109/JSEN.2018.2885582
14. Kröger, K., Ose, C., Rudofsky, G., Roesener, J., Weiland, D., Hirche, H.: Peripheral veins: influence of gender, body mass index, age and varicose veins on cross-sectional area. Vascular Medicine **8**(4), 249–255 (2003). https://doi.org/10.1191/1358863x03vm508oa
15. Kwasny, D., Hemmerling, D.: Gender and age estimation methods based on speech using deep neural networks. Sensors **21**(14) (2021). https://doi.org/10.3390/s21144785, https://www.mdpi.com/1424-8220/21/14/4785
16. Lu, Y., Xie, S.J., Yoon, S., Wang, Z., Park, D.S.: An available database for the research of finger vein recognition. In: 2013 6th International Congress on Image and Signal Processing (CISP). vol. 01, pp. 410–415 (2013). https://doi.org/10.1109/CISP.2013.6744030
17. Lu, Y., Xie, S.J., Yoon, S., Yang, J., Park, D.S.: Robust finger vein roi localization based on flexible segmentation. Sensors **13**(11), 14339–14366 (2013)

18. Rwigema, J., Mfitumukiza, J., Tae-Yong, K.: A hybrid approach of neural networks for age and gender classification through decision fusion. Biomedical Signal Processing and Control **66**, 102459 (2021). https://doi.org/https://doi.org/10.1016/j.bspc.2021.102459, https://www.sciencedirect.com/science/article/pii/S1746809421000562
19. Sgroi, A., Bowyer, K.W., Flynn, P.J.: The prediction of old and young subjects from iris texture. In: 2013 International Conference on Biometrics (ICB). pp. 1–5 (2013). https://doi.org/10.1109/ICB.2013.6613010
20. Simonyan, K., Vedaldi, A., Zisserman, A.: Deep inside convolutional networks: Visualising image classification models and saliency maps. In: Workshop at International Conference on Learning Representations (2014)
21. Sun, Y., Lo, F.P.W., Lo, B.: A deep learning approach on gender and age recognition using a single inertial sensor. In: 2019 IEEE 16th International Conference on Wearable and Implantable Body Sensor Networks (BSN). pp. 1–4 (2019). https://doi.org/10.1109/BSN.2019.8771075
22. Sun, Y., Zhang, M., Sun, Z., Tan, T.: Demographic analysis from biometric data: Achievements, challenges, and new frontiers. IEEE Transactions on Pattern Analysis and Machine Intelligence **40**(2), 332–351 (2018). https://doi.org/10.1109/TPAMI.2017.2669035
23. Tapia, J.E., Perez, C.A., Bowyer, K.W.: Gender classification from the same iris code used for recognition. IEEE Transactions on Information Forensics and Security **11**(8), 1760–1770 (2016). https://doi.org/10.1109/TIFS.2016.2550418
24. Ton, B.T., Veldhuis, R.N.J.: A high quality finger vascular pattern dataset collected using a custom designed capturing device. In: 2013 International Conference on Biometrics (ICB). pp. 1–5 (2013). https://doi.org/10.1109/ICB.2013.6612966
25. Wang, J., Wang, G., Pan, Z.: Gender attribute mining with hand-dorsa vein image based on unsupervised sparse feature learning. IEICE Transactions on Information and Systems **E101.D**(1), 257–260 (2018). https://doi.org/10.1587/transinf.2017EDL8098
26. Wimmer, G., Prommegger, B., Uhl, A.: Finger vein recognition and intra-subject similarity evaluation of finger veins using the cnn triplet loss. In: Proceedings of the 25th International Conference on Pattern Recognition (ICPR). pp. 400–406 (2020)
27. Xie, S., Girshick, R., Dollár, P., Tu, Z., He, K.: Aggregated residual transformations for deep neural networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 5987–5995 (2017). https://doi.org/10.1109/CVPR.2017.634
28. Yaman, D., Irem Eyiokur, F., Kemal Ekenel, H.: Multimodal age and gender classification using ear and profile face images. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. pp. 0–0 (2019)
29. Zhang, K., Gao, C., Guo, L., Sun, M., Yuan, X., Han, T.X., Zhao, Z., Li, B.: Age group and gender estimation in the wild with deep ror architecture. IEEE Access **5**, 22492–22503 (2017). https://doi.org/10.1109/ACCESS.2017.2761849