Acta Scientiarum

http://periodicos.uem.br/ojs
ISSN on-line: 1807-8664
Doi: 10.4025/actascitechnol.v45i1.61948

Dorsal hand vein biometrics with a novel deep learning approach for person identification

Felix Olanrewaju Babalola, Yıltan Bitirim and Önsen Toygar*

Computer Engineering Department, Faculty of Engineering, Eastern Mediterranean University, 99628, Famagusta, North Cyprus, via Mersin 10, Turkey.
*Author for correspondence. E-mail: onsen.toygar@emu.edu.tr

ABSTRACT. Hand dorsal biometric recognition system proposed in this study combines the strength of information in regions of dorsal vein biometric trait in a deep learning based Convolutional Neural Networks (CNN) model. The approach divides each dorsal image into five overlapping regions; consequently, five different training and test sets are obtained for each image, modeling a multi-modal biometric system while using only one trait. The test outputs are combined by score-level fusion. Experimental results on FYO, Bosphorus and Badawi datasets indicate the efficiency of the proposed method and its comparability with other recognition systems. The results are also compared with the state-of-the-art dorsal hand vein recognition systems to show the ability of the proposed biometric architecture to perform well in different conditions that may affect dorsal vein pattern acquisition and have con-sequent effect on the efficiency of the recognition system.

Keywords: dorsal vein recognition; convolutional neural network; score-level fusion; overlapping image regions.

Received on November 11, 2021.
Accepted on August 22, 2022.

Introduction

Biometric systems with physical traits such as palmprint, fingerprint, iris, and face are employed for person recognition and secure system authentication. These systems have become widely popular due to enormous success made over the years in biometric research and consequent technological developments for the acquisition of the traits both for research work and real life implementations (Farmanbar & Toygar, 2015). Moreover, hand vein patterns from palm, wrist, fingers and other regions of the hand have also received massive popularity due to inherent attributes of vein pattern, such as stability, uniqueness and spoof-proof properties, as well as the emergence of technologies that can easily capture these traits in a user friendly manner. These properties make vein pattern recognition one of the most reliable way of effectively restricting access to a system and safeguarding data.

This study explores the promising attributes of dorsal vein (back of palm) as bio-metric trait and propose a system which exploits the strength of multi-modal systems for enhancing dorsal vein biometric recognition systems by merging outputs from five over-lapping sections of dorsal vein samples. The system is designed with a Convolutional Neural Network (CNN) model which is trained separately for each overlapping region. The overall decision of the system is determined by merging scores from five predictions which correspond to each region. The experiments were conducted on FYO, Bosphorus and Badawi dorsal databases to demonstrate the efficacy of the proffered system.

It is generally perceived that when human are subjected to different routine activities or environmental conditions, it bears an effect on sensors' ability to capture vein pattern effectively for recognition, or that captured images may be slightly different in different conditions which will therefore affect the efficiency of the system. The robustness of the system against this perception was also tested by experimenting with data captured under different conditions such as very cold weather, using the device after strenuous exercise or under normal conditions, which are available in Bosphorus database.

The contributions of this work can be summed up as follows:

i) This study introduces a new dorsal hand vein biometric authentication system that fuses scores from five overlapping sections of dorsal vein samples using a CNN-based architectures.

ii) The proposed system takes the advantage of multi-modality within a unimodal system.

iii) Texture-based and CNN-based systems are compared with the proposed biometric recognition system.

Biometric recognition systems have been very well researched, Dargan and Kumar (2020) carried out a comprehensive survey of physiological and behavioral modalities that have been used in several research in literature. More importantly, vein patterns have been used in several researches, ranging from vein pattern on fingers (Qiu, Liu, Zhou, Huang, & Nie, 2016; Purushothaman & Mahalakshmi, 2017) to vein pattern on palms (Sudha, Bhanushali, Nikam, & Tripathi, 2015; Li & Yuan, 2018), to wrist vein pattern (Matkowski, Chan, & Kong, 2019), to hand dorsal vein pattern for human recognition (Huang, Zhu, Wang, & Zhang, 2016). Hand shape and geometry have also been employed in similar ways (Sharma, Dubey, Singh, Saxena, & Singh, 2015).

Available research on hand dorsal vein recognition indicate that vein images are obtained using near infrared or infrared cameras similar to those used to capture vein pattern from the wrist, palm and finger. Subsequently, there has been several proposals in literature on how to use this trait for human recognition, such as in (Hamid, Narang, & Singh, 2017) where a security system was proposed based on dorsal vein patterns where line segments are extracted using Hough transform and modified Hausdorff distance was used for matching.

Our system was motivated by the knowledge of multi-biometrics systems which has proven to improve the performance of biometric systems as shown in several re-searches where more than one hand vein parts or hand vein with other traits were used. For example, palm vein images and vascular patterns of finger vein used with enhanced two dimensional Gabor filter and a methodology based on gradient representation were proposed in (Bharathi & Sudhakar, 2018), Gabor filters on palm and wrist vein preceded by multi-level image improvement has also been studied (Abed, 2017), integration of dorsal and palm vein features was proposed in (Gupta & Gupta, 2015), while concatenation of scores from dorsal, palm and wrist vein features for human authentication was proposed in (Toygar, Babalola, & Bitirim, 2020).

Table 1 shows further research studies that has been undertaken over the years in the field of dorsal vein as a trait for human recognition. It presents available dorsal vein databases and where they have been used in literature, while briefly representing the methods used for recognition in those researches and their major experimental results.

Table 1. Review of dorsal vein research systems.

Reference	Research	Database	Result
Yadav and Rajankar (2012)	Introduced Bosphorus hand vein database and proposed a hand vein recognition system using Independent Component Analysis	Bosphorus	FAR: 0.00% & FRR: 20.00%
Trabelsi, Alima and Sellami (2013)	Recognition using finger vein and hand vein: Monogenic local binary pattern for finger vein and Improved Gaussian Matched Filter (IGMF) for hand vein.	Bosphorus & SDUMLA-HMT[1] (finger vein)	FAR: 0.0% & FRR: 1.0% & EER: 0.004%
Hamid et al. (2017)	Hand vein authentication using line segments extracted using Hough transform and modified Hausdorff distance for matching.	CIR biometrics[2]	FAR: 0.2% & FRR: 0.1%
Al-johania and Elrefaei (2019)	Hand vein authentication using CNN by transfer learning and feature learning approaches.	Badawi & Bosphorus	Acc: 99.25 – 100%
Toygar et al. (2020)	Multi-trait hand vein database comprising of dorsal, palmar and wrist vein datasets. Proposed multimodal recognition system with Binarized Statistical Image Features (BSIF) and a CNN-based technique.	Badawi, Bosphorus & FYO	Acc: 98.13 – 100%
Sayed, Taha and Zayed (2021)	Improvement of smartphone recognition technology with dorsal vein pattern extracted using oriented FAST and rotated BRIEF[3] (ORB[4]) algorithm.	Own database and PUT	EER: 4.33%

Material and methods

The proposed dorsal vein recognition method determines the overlapping vein image regions, trains these regions separately by a CNN architecture and combines the obtained CNN predictions using score-level fusion. Each step of proposed system is illustrated in Figure 1. Firstly, it involves pre-processing stage where input images are converted into required dimension, then the images are enhanced by removing noise after cropping out background areas. The overlapping regions of all images are identified and each region is used

[1]SDUMLA-HMT is a multimodal biometric database, the acronym stands for Shandong University Machine Learning and Applications - Homologous Multi-modal Traits.
[2]CIR: Center for International Research.
[3]BRIEF: Binary Robust Independent Elementary Features.
[4]ORB: Oriented FAST and rotated BRIEF.

in its corresponding CNN training module, followed by classification module which generates a set of five scores for each train image and test image pair. Five scores in each set are added together and results are employed in the determination of the class that the test image belongs to. The system is explained in the following subsections. A pseudo code of the proposed system is given in Figure 2.

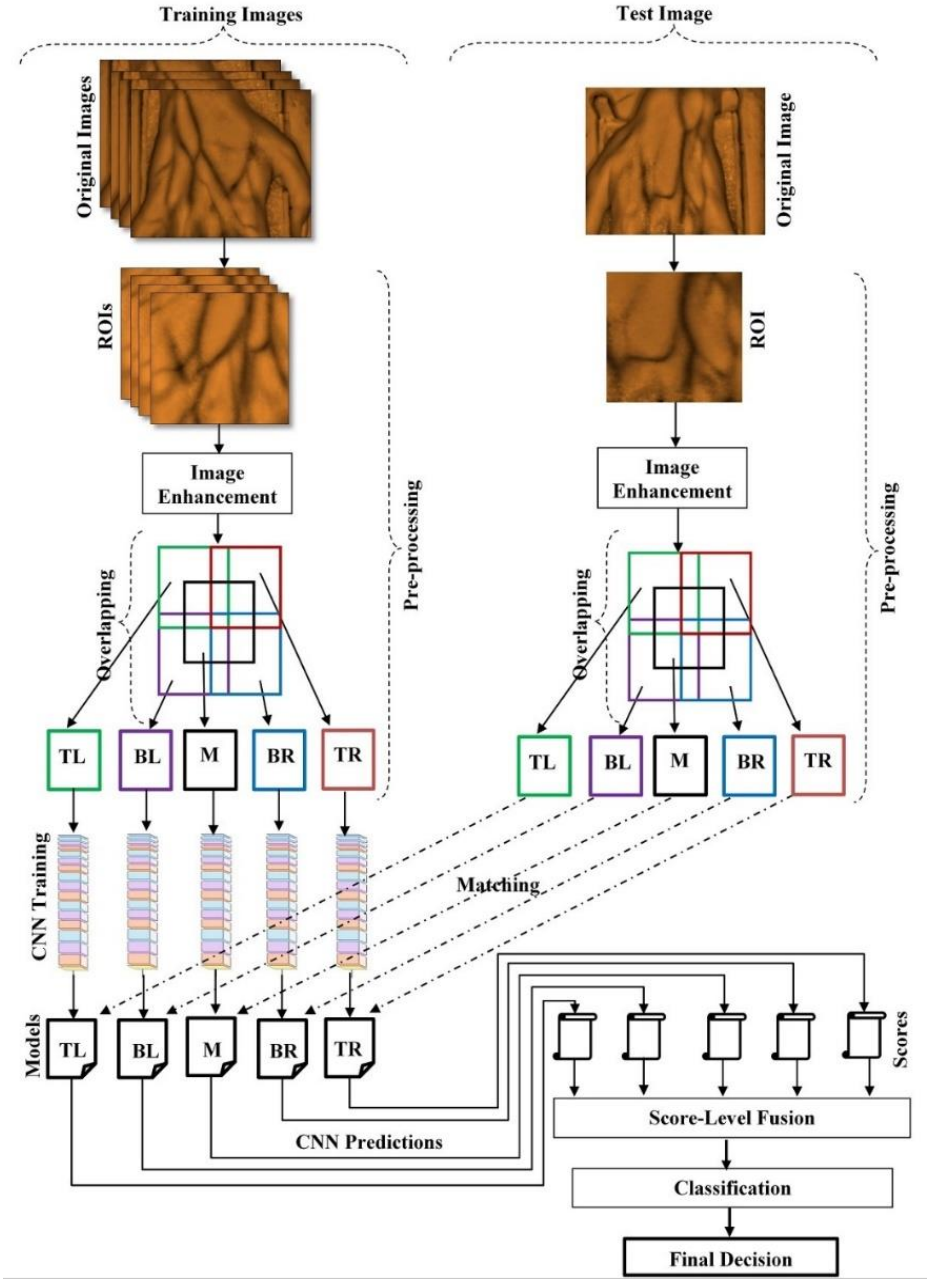

Figure 1. Proposed method with separate training for each region.

Pre-processing

This is the stage where the dataset is prepared for use in the system. It was carried out in four phases: Region of Interest (ROI) determination and cropping, overlapping regions determination, image enhancement and training sets collation.

Captured hand-dorsal vein images usually have other parts of the hand such as finger and wrist, as well as the background. These are unwanted parts that must be removed before processing such images further. In this study, determining and cropping of ROI is done semi-automatically in two phases. In the first phase, a suitable dimension for the ROI is determined that best fits all sample images and all images are automatically cropped using the defined dimension. However, due to hand rotation and variation in size of hands, dimension is not the same for all images in the dataset. Consequently, in the second phase,

badly cropped images are automatically re-cropped with different dimensions. An image is considered to be badly cropped if it contains finger crest, finger, wrist or background. The two phases are repeated until there are no badly cropped images (Babalola, Bitirim, & Toygar, 2021).

Score-Level Fusion of five CNN models based on overlapping regions of hand vein images.

Overlapping region definition:

- Images in the dataset are resized to same size: 336 x 336 pixels.
- Histogram equalization is carried out to enhance the image.
- Divide five overlapping regions (s1, s2, s3, s4 and s5) from each image, where
  - X1 is 224 x 224 pixels from top left corner of image
  - X2 is 224 x 224 pixels from top right corner of image
  - X3 is 224 x 224 pixels from bottom left corner of image
  - X4 is 224 x 224 pixels from bottom right corner of image
  - X5 is 224 x 224 pixels centered at the middle of image

Class label is assigned to each image and the dataset is divided into training and test sets. Then both training and test sets are randomly reshuffled in order to have an unbiased training process.

Training:

Five separate models (model1, model2… model5) are trained using corresponding regions defined earlier; model1 is trained using X1s… model5 is trained using X5s.

Model Evaluation:

- Scores are obtained using the test set samples. This is done by assigning them as input to their corresponding model for prediction; X1s to model1… X5s to model5.
- Five scores are obtained per test image; these are fused using the chosen score level fusion technique.
- Classification is carried out using simple k-NN algorithm.

Figure 2. Pseudo code of overlapping regions system for hand vein recognition.

The regions are formed in the next stage, where five overlapping regions for each of the dorsal ROI image are determined. These regions were chosen based on preliminary experiment in (Babalola et al., 2021) which showed that the middle region represent a stable part of the images and is the best representation of the whole image. Additionally, the overlapping quartet were found to have comparable strength with the middle region, and can also account for vertical and horizontal shifts in samples during acquisition and ROI cropping. Since the ROI will have different dimensions after cropping, re-sizing is carried out to ensure all images of the same dimension (334 x 334 pixels), while a region is defined as 224 x 224 pixels. Therefore, the overlapping regions are 224 x 224 kernels shifted to the corners of the image, that is, top-left, top-right, bottom-left and bottom-right corners of the ROI images. However, the middle overlapping region is 224 x 224 pixels centered at the center of the ROI image.

Afterwards, image enhancement using histogram equalization is the third pre-processing stage. Histogram equalization is a well-known image enhancement technique that can be used to improve contrast in images. This can potentially remove noise from images and enable vein pattern more distinctive.

In the last phase of pre-processing stage, ROI images are categorized as training and test sets. Then from training images, for each overlapping region type, namely; Top-Left (TL), Top-Right (TR), Bottom-Left (BL), Bottom-Right (BR) and Middle (M), a training set is formed. Hence, there are five training sets which are Training-set TL (for TL region), Training-set TR (for TR region), Training-set BL (for BL region), Training-set BR (for BR region), and Training-set M (for M region). Class labels are assigned to every region in the training set.

CNN Training modules

The proposed scheme was carried out on three CNN models; two based on AlexNet architecture (referred to as AlexNet-M1 and AlexNet-M2), one based on VGG16 architecture (referred to as VGG16-M1). AlexNet-M1 and AlexNet-M2 are composed of five convolution layers just like AlexNet (Krizhevsky, Sutskever, & Hinton, 2017) but with lesser number of filters per layer to decrease computation time while also maintaining efficiency (Babalola et al., 2021). Similarly, VGG16-M1 has sixteen convolution layers group into 5 layers as in VGG16 (Simonya & Zisserman, 2014). The differences in the number of filters are presented in Table 2.

Table 2. Number of filters per convolution layer.

Layer	AlexNet	AlexNet-M1	AlexNet-M2	VGG16	VGG16-M1
1st layer	96	32	32	64	32
2nd layer	256	64	64	128	64
3rd layer	384	96	128	256	96
4th layer	384	96	128	512	128
5th layer	256	128	64	512	128

In AlexNet-M1 and AlexNet-M2, each convolution layer convolution layer is followed by an activation function, Rectified Linear Unit (ReLU) as in AlexNet. This converts all negative values to zero and leaves all positive values untouched. A Max pooling and Batch normalization layers follows. However in VGG16-M1, Max pooling and Batch normalization follows each layer as in VGG16, not each convolution layer. The fully-connected layers to end the architecture maintains the original structures from AlexNet and VGG16, respectively.

As shown in Figure 1, five models are trained, one for each overlapping region, using the corresponding training sets formed at the pre-processing stage. Each trained model is used for forecasting the class label of test sample from the same region. It should be noted that though the training process is five times longer, the testing phase is not affected by multiple training; therefore computation time is not a detrimental factor in real world usage. Additionally, the models used are fast due to reduced number of filters in the convolution layers.

Classification module

This session is illustrated under testing phase in Figure 1, where sample test images are introduced to the system for prediction. Sample test images go through the same initial pre-processing stages as the training set as shown in Figure 1; after carrying out enhancement on the ROI image and determining overlapping regions TL, TR, BL, BR, and M. Each region is introduced to the matching stage of its corresponding trained model for prediction. This generates probabilities of a test sample to belong to each of the classes in the system which represents the scores. Consequently, there will be five scores for each class, which are added together and entered into the classification phase to assign class to test samples, where k-NN classifier is employed to identify the closest training sample to the test sample. k classes correspond to the number of unique subjects in the datasets used. The class with the highest fused score is considered as the predicted class. This is a relatively fast and less computation intensive method of classification.

Databases

Three dorsal vein datasets obtained from FYO, Bosphorus and Badawi databases were used in this study. Samples of images from the three datasets are shown in Figure 3. FYO database from (Toygar et al., 2020), is a multi-trait hand vein database. The traits include dorsal, palmar and wrist vein patterns. The dataset was acquired from 160 volunteers in two separate sessions, where hand vein patterns were captured from both hands, comprising of left and right dorsal vein, left and right palmar vein and left and right wrist vein pattern images. These sum up to 1920 sample in all, 640 images per trait (left and right included).

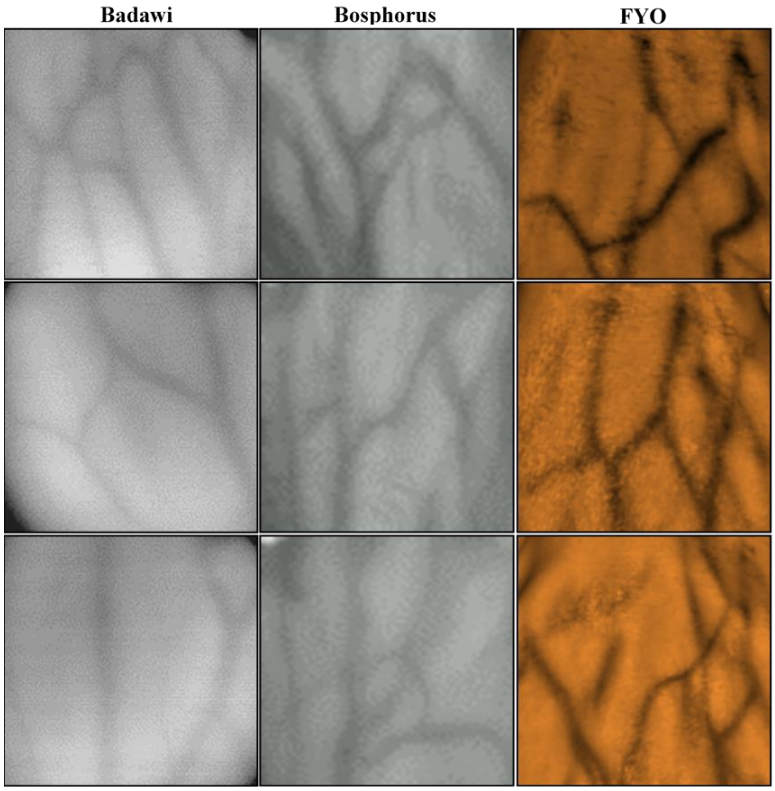

Figure 3. Sample images from the datasets.

A dataset of generated images based on the ROI of their datasets is also provided. The images were generated using Keras data generator. It contains 6,400 images for each of the three traits (10 images were generated per sample ROI image).

Bosphorus hand vein database contains 1,200 left hand images from 100 individuals taken in four sessions simulating physical conditions that could take place in real-life before the hand is presented to the sensor for sample acquisition, such as, carrying a 3 kg sack for about 60 seconds (Bag), squeezing an elastic ball by opening and closing repeatedly for about 60 seconds (Activity), lowering the temperature of the hand by placing a sack of ice on the back of the hand (Ice) and lastly, under normal condition (Normal). Three images were taken per subject in each session (Yuksel, Akarun, & Sankur, 2011).

Badawi database from (Shahin, Badawi, & Rasmy, 2011), has 100 subjects taken from 50 volunteers of age 16 and above, both male and female. 10 images were acquired per person at varying intervals (5 images each from left and right hands). This gives a total of 500 samples, the complete dataset was used in our experiments.

On the other hand, the sample size of each of the above named datasets are too small for training in deep learning. Training samples generally perform better when they are about 5,000 samples or more. Therefore data augmentation was performed using Keras data generator which creates new images using parameters such as varying zoom, height and width shift, brightness adjustment and applying slight rotation on the original images. Consequently, the datasets used for testing the CNN models were organized as follows:

i) Badawi: 11 new images were generated for each sample in the original dataset (500 images, 5 samples per person), summing up to an aggregate of 5,500 images, 5,000 of which were for training (about 91 percent) and 500 were used as test samples.

ii) Bosphorus: 5 new images were generated for each sample in the original dataset (1,200 images, 12 samples per person), summing up to a total of 6,000 images with 60 samples per subject. 90 percent of the augmented dataset was used for training, while 10 percent was used as test samples.

FYO: The generated dorsal vein dataset was used from this database, comprising of 6,400 images (20 samples per subject). Similarly, 90 percent of the dataset was for training and the remaining 10 percent for testing.

Results and discussion

Experiments were performed to ascertain the performance of the proposed biometric architecture and to present comparison with texture-based methods that has been used in similar experiments. A texture based

algorithm called Binarized Statistical Image Features (BSIF) (Toygar et al., 2020; Babalola et al., 2021) and its variant; Multiple-Binarized Statistical Image Features (M-BSIF) based on Bank of Binarized Statistical Image Features (B-BSIF) (Attia, Chaa, Akhtar, & Chahir, 2018) were implemented using the same datasets to compare corresponding results to the proposed method. These methods are some of the most efficient texture-based methods in the literature. For further comparison, we implemented the CNN models which takes the ROI as a whole as used in (Toygar et al., 2020, Babalola et al., 2021) to compare our method where separate training of five regions were carried out. For all these implementations, the aforementioned training and test datasets were used.

The results (accuracies) shown in Table 3 substantiates that our proposed system earned the best accuracy for each of the datasets. The table present systems where the standard method of extracting features in one training set from the ROI is used. These systems can be divided into two types; handcrafted descriptors (BSIF and M-BSIF) and CNN models (AlexNet-M1, AlexNet-M2, VGG16-M1). A comparison of these two types show that CNN models mostly performed better with the highest average accuracy of 99.19 percent reached by the AlexNet-M1 model. Furthermore, we compared the proposed method with the other methods, the proposed method produced the highest average accuracy of 99.79 percent when the AlexNet-M1 model is used in the proposed system, where five sub-regions were trained individually and the individual predictions are fused at the score level. When we compare the lowest average accuracy, we find that the proposed system has the highest one in this category as well (99.07 percent).

Table 3. Proposed system comparison (Accuracy (%)).

Method		Badawi	Bosphorus	FYO	Average
Handcrafted methods	BSIF	98.20	92.42	95.00	95.21
	M-BSIF	98.20	94.41	96.88	96.50
CNN models	AlexNet-M1	99.00	99.67	98.90	99.19
	AlexNet-M2	99.00	97.33	96.56	97.63
	VGG16-M1	97.60	98.50	94.37	96.82
Proposed method	AlexNet-M1	99.70	99.83	99.84	99.79
	AlexNet-M2	99.00	98.67	99.53	99.07
	VGG16-M1	99.60	99.83	99.38	99.60

Furthermore, in order to further check the suitability of the CNN models used, Equal Error Rates (EER) of the models were examined and presented in Table 4. The results show that the EERs are very low, therefore confirming the suitability of the models. GAR at 0.01% FAR and GAR at 0.1% FAR were also checked and 100% was obtained in all cases.

Table 4. Equal Error Rates (EER (%)) of the CNN models.

	Badawi	Bosphorus	FYO
AlexNet-M1	0.023	0.105	0.040
AlexNet-M2	0.019	0.100	0.031
VGG16-M1	0.007	0.097	0.040

Additionally, we explored other fusion techniques that have been used in literature for score level fusion rather than the traditional techniques where values are only normalized and appended. Cheniti, Eddine, and Akhtar (2017) proposed symmetric addition of scores while Herbadji, Guermat, Lahcene, Akhtar, and Dasgupta (2020) proposed Weighted Quasi-Arithmetic Mean (WQAM) for fusion at the score stage in multi-biometric systems. WQAM involve the mathematical functions such as trigonometry functions, cosine, sine and tangent along with some defined weight for score level fusion. General formula of M for this method is given as follows:

$$M(s) = f^{-1}\left(\sum_{j=1}^{n} w_j f(s_j)\right) \quad (1)$$

where s is the input sample, f is the mathematical function and f^{-1} is its corresponding inverse, while w_j is weight assigned to each score. The weights add up to 1.

Some of the functions are adapted from Herbadji et al. (2020) and used to compare the original SLF. The results are shown in Tables 5, 6 and 7, indicating that the system works effectively across different score level fusion techniques.

Table 5. Accuracy of score level fusion techniques in AlexNet-M1.

Fusion Method	Badawi	Bosphorus	FYO
Normal SLF	99.70	99.83	99.84
WQAM; f=sin($\frac{\pi}{2}$s)	99.80	99.92	100.00
WQAM; f=cos($\frac{\pi}{2}$s)	99.40	99.42	99.06
WQAM; f=tan($\frac{\pi}{2}$s)	99.30	99.08	99.06
WQAM; f= $\boldsymbol{s^r}$	99.40	99.33	98.91
WQAM; f= $\boldsymbol{r^s}$	99.70	100.00	100.00

Table 6. Accuracy of score level fusion techniques in AlexNet-M2.

Fusion Method	Badawi	Bosphorus	FYO
Normal SLF	99.00	98.67	99.53
WQAM; f=sin($\frac{\pi}{2}$s)	99.10	99.83	99.53
WQAM; f=cos($\frac{\pi}{2}$s)	98.70	98.02	98.75
WQAM; f=tan($\frac{\pi}{2}$s)	98.60	98.33	98.59
WQAM; f= $\boldsymbol{s^r}$	98.70	98.47	98.44
WQAM; f= $\boldsymbol{r^s}$	99.20	99.92	100.00

Table 7. Accuracy of score level fusion techniques in VGG16-M1.

Fusion Method	Badawi	Bosphorus	FYO
Normal SLF	99.60	99.83	99.38
WQAM; f=sin($\frac{\pi}{2}$s)	100.00	100.00	99.53
WQAM; f=cos($\frac{\pi}{2}$s)	98.90	98.70	99.06
WQAM; f=tan($\frac{\pi}{2}$s)	98.20	98.20	98.44
WQAM; f= $\boldsymbol{s^r}$	98.90	98.70	98.59
WQAM; f= $\boldsymbol{r^s}$	100.00	100.00	99.69

On the other hand, earlier human recognition systems, where dorsal vein pattern was proposed as trait are compared to the proposed method in Table 8. The table shows the most essential constituents of all biometric authentication studies, including pre-processing approach, feature description and extraction approach, feature matching and classification approach, dataset used for testing and the experimental outcome. Comparing some similar studies between years 2010 and 2021, it can be deduced that the proposed system compared favorably against the most successful systems in literature.

Table 8. Comparison with similar studies on dorsal vein recognition system.

Dataset	Author / Year/Ref.	Pre-processing	Feature extraction method	Matching / Classification approach	Accuracy (%)
Badawi database	Trabelsi et al. (2013)	Gaussian filter	CDSDP	ANN+FMNN	98.80
	Al-johania and Elrefaei (2019)	None	AlexNet	ECOC plus KNN	97.00
			VGG16		96.00
			VGG19		99.00
			AlexNet	ECOC plus SVM	100.00
			VGG16		96.00
			VGG19		100.00
	Toygar et al. (2020)	Histogram equalization	BSIF	NN Classifier	95.50
			AlexNet		98.67
			Proposed CNN model		98.67
	Proposed method	Histogram equalization	AlexNet-M1	NN Classifier	99.70
			AlexNet-M2		99.00
			VGG16-M1		99.60
Bosphorus database	Al-johania and Elrefaei (2019)	None	AlexNet	ECOC plus KNN	90.00
			VGG16		89.50
			VGG19		77.00

			AlexNet	ECOC plus SVM	97.50
			VGG16		98.50
			VGG19		93.50
	Çimen, Boyraz, Yildiz, and Boz (2021)	CLAHE, median filtering, adaptive thresholding, morphological processes	Fractal dimension box counting method	SVM with linear kernel function (LKF)	99.00
	Toygar et al. (2020)	Histogram equalization	BSIF	NN Classifier	80.17
			AlexNet		100.00
			Proposed CNN model		99.00
		Histogram equalization	AlexNet-M1	NN Classifier	99.83
	Proposed method		AlexNet-M2		98.67
			VGG16-M1		99.83
FYO database	Toygar et al. (2020)	Histogram equalization	BSIF	NN Classifier	95.31
			AlexNet		99.38
			Proposed CNN model		98.90
	Proposed method	Histogram equalization	AlexNet-M1	NN Classifier	99.84
			AlexNet-M2		99.53
			VGG16-M1		99.38

Conclusion

A CNN-based architecture for dorsal hand vein recognition is proposed in this study. It combines information from overlapping regions of dorsal images. The proposed method creates five overlapping regions from all images in each database, forming five separate datasets which are trained separately to exploit the superior efficiency of multimodal biometric systems over unimodal systems. Scores generated from CNN predictions are concatenated to acquire the overall decision of the proposed method. To compare this system with the state-of-the-art methods, other CNN models were experimented on as well as texture based models. The results showed that CNN models outperformed texture-based methods while the proposed model performed the best when compared with the state-of-the-art methods. Furthermore, our subsequent work will be directed at implementing the proposed method on other biometric traits for authentication while also expanding our research on dorsal vein authentication system. We will also study how other deep learning based models will perform when used in-line with the proposed architecture.

References

Abed, M. H. (2017). Wrist and palm vein pattern recognition using Gabor filter. *Journal of Al-Qadisiyah for Computer Science and Mathematics*, *9*(1), 49-60.

Abderrahmane, H., Guermat, N., Lahcene, Z., Akhtar, Z., & Dasgupta, D. (2020). Weighted quasi-arithmetic mean based score level fusion for multi-biometric systems. *IET Biometrics*, *9*(3), 91-99. DOI: https://doi.org/10.1049/iet-bmt.2018.5265

Al-johania, N., & Elrefaei, L. (2019). Dorsal hand vein recognition by convolutional neural networks: feature learning and transfer learning approaches. *International Journal of Intelligent Engineering and Systems*, *12*(3), 178-191. DOI: https://doi.org/10.22266/ijies2019.0630.19

Attia, A., Chaa, M., Akhtar, Z., & Chahir, Y. (2018). Finger kunckcle patterns based person recognition via bank of multi-scale binarized statistical texture features. *Evolving Systems*, *11*(4), 625-635.

Babalola, F., Bitirim, Y., & Toygar, Ö. (2021). Palm vein recognition through fusion of texture-based and CNN-based methods. *Signal, Image and Video Processing*, *15*(3), 459-466. DOI: https://doi.org/10.1007/s11760-020-01765-6

Bharathi, S., & Sudhakar, R. (2018). Biometric recognition using finger and palm vein images. *Soft Computing*, *23*(4), 1843-1855. DOI: https://doi.org/10.1007/s00500-018-3295-6

Cheniti M., Boukezzoula, N. E., & Akhtar, Z. (2017). Symmetric sums-based biometric score fusion. *IET Biometrics*, *7*(5)1-4. DOI: https://doi.org/10.1049/iet-bmt.2017.0015

Çimen, M., Boyraz, O., Yildiz, M., & Boz, A. (2021). A new dorsal hand vein authentication system based on fractal dimension box counting method. *Optik*, *226*(1), 1-6. DOI: https://doi.org/10.1016/j.ijleo.2020.165438

Dargan, S., & Kumar, M. (2020). A comprehensive survey on the biometric recognition systems based on physiological and behavioral modalities. *Expert Systems with Applications*, *143*, 1-7. DOI: https://doi.org/10.1016/j.eswa.2019.113114

Farmanbar, M., & Toygar, Ö. (2015). Feature selection for the fusion of face and palmprint biometrics. *Signal, Image and Video Processing*, *10*(5), 951-958. DOI: 10.1007/s11760-015-0845-6

Gupta, P., & Gupta, P. (2015). Multi-modal fusion of palm-dorsa vein pattern for accurate personal authentication. *Knowledge-Based Systems*, *81*, 117-130. DOI: https://doi.org/10.1016/j.knosys.2015.03.007

Hamid, H., Narang, V. K., & Singh, D. (2017). Dorsal Hand Vein Analysis for Security Systems. *International Journal of Engineering and Technology*, *9*(4), 3075-3080. DOI: https://doi.org/10.21817/ijet/2017/v9i4/170904044

Huang, D., Zhu, X., Wang, Y., & Zhang, D. (2016). Dorsal hand vein recognition via hierarchical combination of texture and shape clues. *Neurocomputing*, *214*, 815-828. DOI: https://doi.org/10.1016/j.neucom.2016.06.057

Krizhevsky, A., Sutskever, I., & Hinton, G. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84-90. DOI: https://doi.org/10.1145/3065386

Li, W., & Yuan, W. (2018). Multiple palm features extraction method based on vein and palmprint. *Journal of Ambient Intelligence and Humanized Computing*. DOI: https://doi.org/10.1007/s12652-018-0699-1

Mahalakshmi, R., & Purushothaman, R. (2017). Multimodal biometric in human Identification of finger vein patterns using score level and GSA. *International Journal of Engineering and Computer Science*, *6*(3), 20678-20686. DOI: https://doi.org/10.18535/ijecs/v6i3.51

Matkowski, W., Chan, F., & Kong, A. (2019). A study on wrist identification for forensic investigation. *Image and Vision Computing*, *88*, 96-112. DOI: https://doi.org/10.1016/j.imavis.2019.05.005

Qiu, S., Liu, Y., Zhou, Y., Huang, J., & Nie, Y. (2016). Finger-vein recognition based on dual-sliding window localization and pseudo-elliptical transformer. *Expert Systems with Applications*, *64*, 618-632. DOI: https://doi.org/10.1016/j.eswa.2016.08.031

Sayed, M., Taha, M., & Zayed, H. (2021). Real-time dorsal hand recognition based on smartphone. *IEEE Access*, *9*, 151118-151128. DOI: https://doi.org/10.1109/ACCESS.2021.3126709

Simonyan, K., & Zisserman, A. (2014). *Very deep convolutional networks for large-scale image recognition*. Ithaca, NY: Cornell University. DOI: https://doi.org/10.48550/arXiv.1409.1556

Shahin, M., Badawi, A., & Rasmy M. E. (2011). Multimodal biometric system based on near-infra-red dorsal hand geometry and fingerprints for single and whole hands.*World Academy of Science, Engineering and Technology*, *56*, 1107–1122. Retrieved from https://bitlybr.com/JRDxe

Sharma, S., Dubey, S., Singh, S., Saxena, R., & Singh, R. (2015). Identity verification using shape and geometry of human hands. *Expert Systems with Applications*, *42*(2), 821-832. DOI: https://doi.org/10.1016/j.eswa.2014.08.052

Sudha.K, L., Bhanushali, N., Nikam, P., & Tripathi, P. (2015). High security using palm vein recognition technology. *International Journal of Computer Applications*, *119*(1), 39-44. DOI: https://doi.org/10.5120/21035-3141

Toygar, O., Babalola, F., & Bitirim, Y. (2020). FYO: A novel multimodal vein database with palmar, dorsal and wrist biometrics. *IEEE Access*, *8*, 82461-82470. DOI: https://doi.org/10.1109/ACCESS.2020.2991475

Trabelsi, R., Alima, D., & Sellami, D. (2013). A new multimodal biometric system based on finger vein and hand vein recognition. *International Journal of Engineering and Technology*, *5*(4), 3175-3183.

Yadav, V., & Rajankar, S. (2012). Dorsal hand vein biometry by independent component analysis. *International Journal on Computer Science and Engineering, 4*(7), 1338-1344

Yuksel, A., Akarun, L., Sankur, B. (2011). Hand vein biometry based on geometry and appearance methods', *IET Computer Vision*, *5*(6), 398–406. DOI: https://doi.org/10.1049/iet-cvi.2010.0175