From Synthetic Data to Real Palm Vein Identification: a Fine-Tuning Approach

Ruber Hernández-García
CIEAM Research Center, LITRP Laboratory
Universidad Católica del Maule
Talca, Chile
rhernandez@ucm.cl

Edwin H. Salazar-Jurado
DMMA PhD Program, LITRP Laboratory
Facultad de Ciencias Básicas
Universidad Católica del Maule
Talca, Chile

Ricardo J. Barrientos
LITRP Laboratory, Dept. DCI
Faculty of Engineering Sciences
Universidad Católica del Maule
Talca, Chile

Francisco Manuel Castro
Dept. of Computer Architecture
Universidad of Málaga
Málaga, Spain

Julián Ramos-Cózar
Dept. of Computer Architecture
Universidad of Málaga
Málaga, Spain

Nicolás Guil
Dept. of Computer Architecture
Universidad of Málaga
Málaga, Spain

***Abstract*—Palm vein recognition has relevant advantages in comparison with most traditional biometrics, such as high security and recognition performance. In recent years, CNN-based models for vascular biometrics have improved the state-of-the-art, but they have the disadvantage of requiring a larger number of samples for training. In this context, the generation of synthetic databases is very effective for evaluating the performance of biometric systems. The present study proposes a new perspective of a transfer learning approach for palm vein recognition, evaluating the use of Synthetic-sPVDB and NS-PVDB synthetic databases for pre-training deep learning models and validating their performance on real databases. The proposed methodology comprises two different branches as inputs. Firstly, a synthetic database is used to train a CNN model, and in the second branch, a real database is used to finetune and evaluate the performance of the resulting pre-trained model. For the feature learning process, we implemented two end-to-end CNN architectures based on AlexNet and Resnet32. The experimental results on the most representative public datasets have shown the usefulness of using palm vein synthetic images for transfer learning, outperforming the state-of-the-art results.**

***Index Terms*—Biometrics, Convolutional neural networks, Palm vein recognition, Synthetic datasets, Transfer learning**

I. Introduction

During the last decade, vascular-based biometrics has gained the attention of the scientific community due to its advantages, such as high security and recognition performance. However, in comparison with other traditional biometric techniques, palm vein recognition still has several challenges to be studied [1]. Especially its applications in the massive identification of individuals are very limited, mainly because publicly available databases have a very small number of subjects.

Table I provides an overview of the publicly available databases of palm vein images, comparing the number of subjects, captured samples per hand and acquisition sessions, and total images. It is worth noting that all databases from human sources are generally composed of a limited number of individuals, with no more than 300 individuals. Usually, the proposed approaches for palm vein recognition adopt the *nom L & R* protocol [2], where both hands are considered as different subjects, doubling the number of subjects in the dataset. However, more individuals are still required to allow scalability studies of the proposed methods on massive identification tasks [3]. A possible solution to address this drawback could be using images from several databases, but this would not solve the problem due to two main reasons. First, this would only yield a total of 2,510 subjects, which is still insufficient to perform massive identification tasks. Secondly, and most notably, variations in the images resulting from different acquisition systems (resolution, illumination, contrast, etc.) can cause problems in training and adjusting the parameters of the proposed algorithms. To address the above limitations, Synthetic-sPVDB [4] and NS-PVDB [5] were introduced as two synthetic datasets and later extended in [3], reaching a total of 20,000 and 16,000 subjects, respectively.

TABLE I
General Overview of Public Databases of Palm Vein Images. The Number of Samples are Represented as Acquisition Sessions × Hands × Samples.

Dataset	Subjects	Samples	Total Images	Capturing Source
CASIA [6]	100	$2 \times 2 \times 3$	7,200	Human
VERA [2]	110	$2 \times 2 \times 5$	2,200	Human
PUT [7]	50	$3 \times 2 \times 4$	1,200	Human
PolyU [8]	250	$2 \times 2 \times 6$	6,000	Human
Tongji* [9]	300	$2 \times 2 \times 10$	12,000	Human
IITI* [10]	185	$1 \times 2 \times 6$	2,220	Human
FYO [11]	160	$2 \times 2 \times 1$	640	Human
Synthetic-sPVDB [4]	20,000	$1 \times 1 \times 6$	120,000	Synthetic
NS-PVDB [5]	16,000	$1 \times 1 \times 6$	96,000	Synthetic

*These datasets are available by request to the principal investigator.

The generation of synthetic databases has been used in some biometric techniques (fingerprints, iris, or face) [12]. Synthetic images are very effective for evaluating the performance of different image processing algorithms, but their use for biometrics is controversial. However, synthetic databases have the advantage of avoiding a time-consuming sample collection process, and they also do not compromise user security at all, which is often a personal concern in terms of privacy violation. Furthermore, although this type of dataset does not fully replace validation with real images, it facilitates the validation of proposed approaches in large-scale databases.

In recent years, CNN-based models for vascular biometrics have gained prominence due to their excellent results, but they have the disadvantage of requiring a larger number of samples for training. As a solution, some approaches have taken advantage of different pre-trained models [13], [14] or applied data augmentation techniques [15]–[17]. It is noticeable that pre-trained models only achieve SOTA results on two databases (i.e., PUT and PolyU). This is caused by the fact that vein images present very particular characteristics of texture, contrast, and blurring, resulting from the NIR acquisition process, which is very different from images on the Imagenet dataset where these models have been initially trained. On the contrary, data augmentation can provide relatively good performance but often require large training samples, which demand high computation and is infeasible for massive identification.

The present study aims to evaluate the use of the proposed synthetic databases Synthetic-sPVDB [4] and NS-PVDB [5] for pre-training convolutional neural network (CNN) models and to validate their performance on real databases. For this, we implemented two end-to-end CNN architectures based on AlexNet [18] and Resnet32 [19]. Firstly, we trained both models on the synthetic databases and later evaluated their performance by using transfer learning with and without a fine-tuning process. Thus, we tackle two main research questions in this work, but not limited to: (i) can we improve the recognition performance of CNN models when pre-trained on synthetic datasets? and (ii) does the recognition performance not decrease as the size of the database increases? In both cases, it would prove the usefulness of using synthetic images to obtain larger datasets than the real ones.

The remaining sections of this article are organized as follows. Section II reviews CNN-based state-of-the-art works using transfer learning approaches. The proposed methodology is formally described in Section III. Then, the obtained experimental results are analyzed in Section IV. Finally, Section V provides the conclusions and future works.

II. Related Works

Table II summarizes the most representative CNN-based approaches of the last five years for palm vein recognition by using transfer learning. It is worth mentioning that, to the best of our knowledge, DNN approaches for palm vein recognition were rarely reported before 2018. During the last few years, DNN-based methods have improved the state-of-the-art (SOTA) results on palm vein databases, achieving an accuracy of 100% on four of seven public databases (i.e., Tongji [9], VERA [2], PolyU [8], and FYO [11]). However, SOTA models based on transfer learning have serious issues related to available training datasets, many parameters' adjusting, and scalability, among others. Although the application of transfer learning techniques avoids optimization of a large number of weights of the model, it is noticeable that pre-trained models only achieve a perfect recognition rate on the PolyU dataset. The used CNN architectures are pre-trained on the Imagenet dataset [20], and different transfer learning approaches are subsequently applied on real datasets. Although training on the Imagenet dataset ensures the inter-class and intra-class variability of the feature representation, the obtained models do not take into account the special effects of NIR palm vein images.

To address de above limitations, the present study proposes a new perspective of a transfer learning approach for palm vein recognition. From the literature review, few publications have focused on providing methodologies for generating and evaluating synthetic images for vein-based biometrics. Contrary to SOTA works, we evaluated the recognition performance of pre-trained CNN models on two synthetic databases [4], [5] and then performed a fine-tuning process on real datasets.

III. Proposed Methodology

Usually, a palm vein recognition system comprises four main processes: image acquisition, image preprocessing, feature extraction/learning, and recognition. Figure 1 depicts the flowchart of the proposed methodology. As input, there are two different branches: initially, a synthetic database is used to train a CNN model, and in the second branch, a real database is used to finetune the resulting pre-trained model. The image preprocessing and recognition performance evaluation are predefined processes that are adopted from the state-of-the-art. Each process is presented in detail in the following sections.

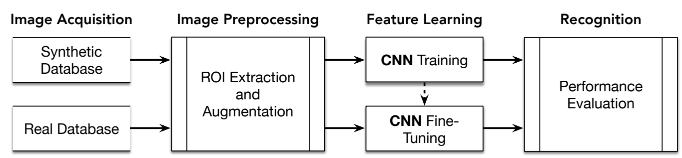

Fig. 1. Flowchart of the proposed methodology.

A. Palm Vein Databases

Vein patterns are visualized as dark lines appearing on the palm images through irradiation of near-infrared (NIR) light and can only be acquired from a living body. Table I summarizes the information of the nine public databases found in the literature, which are used for the evaluation of the proposed models. The first seven were acquired from humans, and the last two were constructed synthetically. The proposed methodology is based on the use of synthetic databases Synthetic-sPVDB [4] and NS-PVDB [5], particularly their second version proposed in [3]. Synthetic-sPVDB was generated by a generative model through the StyleGAN2 network [25]. The StyleGAN2 network was trained using images from the 850nm band of the CASIA dataset [6], which is commonly used for palm vein recognition. On the other hand, NS-PVDB was created by using a mathematical model to simulate the venous plexus of the hand. The mathematical model is based on adaptive dynamics pathfinding, used to simulate the growth of the *Physarum polycephalum* [26], forming vein

TABLE II

OVERVIEW OF THE STATE-OF-THE-ART CNN-BASED APPROACHES ON PALM VEIN RECOGNITION USING TRANSFER LEARNING, FROM LAST FIVE YEARS ON PUBLIC DATABASES. PERFORMANCE RESULTS ARE REPORTED IN CORRECT IDENTIFICATION RATE (CIR), ACCORDING TO ORIGINAL PAPERS. THE BEST RESULTS FOR EACH PUBLICLY AVAILABLE DATABASE ARE HIGHLIGHTED IN BOLD FONT.

Approach	CNN Architecture	Year	Database	Performance
Discrete Wavelet Transform enhancement and evaluation of pre-trained CNNs [13]	AlexNet	2018	PUT	93.92%
	VGG-16			90.83%
	VGG-19			92.27%
	GoogLeNet			76.92%
Convolutional encoder-decoder network and a pre-trained Siamese network using Triplet Loss [21]	U-net *encoder-decoder* + *conv6-fc* triplet Siamese	2019	CASIA	85.16%
			IITI	**97.47%**
			PolyU	98.78%
Transfer learning and fine-tuning on CNN models [14]	*conv6-fc2*	2019	**PUT**	95.16%
	AlexNet			92.16%
	VGG-16			97.33%
	ResNet-50			**99.83%**
	SqueezeNet			91.66%
Pre-trained CNN model with PCA and Random Forest classifier [22]	AlexNet	2019	**PolyU**	**100%**
			CASIA	97.00%
Pre-trained CNN and densely-connected convolutional autoencoder [23]	DenseNet-161	2020	PolyU	99.69%
	ResNeXt-101			98.67%
Evaluation of classic and recent CNN models by using transfer learning [24]*	MobileNet v3	2021	**PolyU**	**100%**
			Tongji	98.67%
	EfficientNet		**PolyU**	**100%**
			Tongji	99.00%
	GhostNet		PolyU	99.60%
			Tongji	96.20%
	ResNeSt-50		PolyU	98.57%
			Tongji	94.92%

*Authors evaluate 17 CNN models; for the sake of clarity, only the last ones are shown.

network-like patterns. The resulting pre-trained models on both synthetic databases were fine-tuned and evaluated on the most representative real databases of the state-of-the-art, i.e., CASIA [6], PolyU [8], PUT [7], VERA [2], Tongji [9], and IITI [10].

B. ROI Extraction and Augmentation

Since public palm vein datasets have a limited number of samples, it is an important limitation to be solved for CNN model training. Unlike other approaches that apply image augmentation from ROI samples during the training process, we adopt the augmentation procedure of ROI samples proposed in [27]. As shown in Fig. 2, the region of interest (ROI) of the palm is determined by the midpoints between the index and middle fingers, and the ring and middle fingers (red square). A central square (blue) is centered in the palm ROI, which is 2/3 the size of the ROI. To increase the number of ROI samples, this central ROI is repeatedly shifted on both axes with $\Delta(X,Y) = \pm 5px$ up to the ROI threshold and is also rotated with $\Delta\theta = \pm 5^{\circ}$ up to $\pm 15^{\circ}$. Therefore, we obtain 87x ROI samples of each palm vein image.

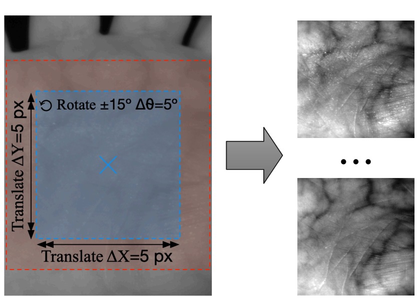

Fig. 2. Scheme of the ROI sample augmentation process.

This augmentation procedure was implemented in a MATLAB script, which allowed the augmentation of all the training sets used, both for the initial training on the synthetic databases and for the fine-tuning process on the real databases. Finally, all resulting images were applied a gray level normalization and then a Contrast Histogram Equalization (CLAHE) to enhance the vein pattern details.

C. CNN Feature Learning

For the feature learning process, we adopted two end-to-end CNN architectures proposed in [27]. Both models differ in depth, and their architectures are presented in tables III and IV. The first model is based on the AlexNet architecture [18], which was referred to as SingleNet. The second one is based on the ResNet32 [19] architecture, taking into account the application of a more complex classification scheme with residual operations. Contrary to SOTA works, we aimed to validate the performance of low-complexity end-to-end CNN models without using pre-trained models on the Imagenet dataset. The details of the implementations of each model evaluated are described below.

1) SingleNet model: The model comprises five convolutional layers followed by three dense layers (*fully-connected*). All convolutional layers use *ReLU* activations and are followed by batch normalization (*BatchNorm*) and pooling (*max-pooling*) operations. In addition to the above normalizations, after the first two fully-connected layers, we applied dropout operations (*dropout*) in order to reduce network overfitting. Finally, the last dense layer has as many neurons as classes (individuals) in the dataset and uses the *Softmax* loss function. For training, the CNN model was compiled using the *Categorical Crossentropy* loss function and the *SGD* optimization function.

TABLE III
SINGLENET ARCHITECTURE: P=POOLING SIZE; BATCHNORM=BATCH NORMALIZATION; DR=DROPOUT; C=OUTPUTS (CLASSES).

Layers	Specifications	Post-processing
Input	64×64	-
Conv01	$7 \times 7 \times 96$	BatchNorm; P: 2×2
Conv02	$5 \times 5 \times 192$	BatchNorm; P: 2×2
Conv03	$3 \times 3 \times 284$	BatchNorm; P: 2×2
Conv04	$3 \times 3 \times 512$	BatchNorm; P: 2×2
Conv05	$2 \times 2 \times 1024$	BatchNorm; P: 2×2
Full01	1024	Dr: 0.65
Full02	512	Dr: 0.65
Full03	C	-

2) ResNet32 model: The ResNet architecture is based on a direct-access connection structure called residual module, where several residual modules are stacked. Residual operations help to prevent information loss along the processing flow without increasing the number of parameters or the computational complexity of the model and improving model performance. The model comprises three groups of residual modules (*ResStack*). At the beginning of each stack, the size of the feature map is reduced by a convolutional layer with $stride = 2$, while the number of filters is doubled. Within each group, all layers have the same number of filters, and the last ReLU activation is performed after the residual connection. After the residual stacks, an average pooling (*AvgPooling*) is computed, and the output dense layer is added for classification. The model was compiled with the *Categorical Crossentropy* loss function and the *Adam* optimization.

TABLE IV
RESNET32 ARCHITECTURE, THE OUTPUTS OF FULL01 LAYER CORRESPOND WITH THE NUMBER OF CLASSES C.

Layers/Stacks	Specifications	Post-processing
Input	64×64	-
Conv2D-BN-ReLU	$3 \times 3 \times 16$	BatchNorm
ResStack01	32×32 conv; 16	-
ResStack02	16×16 conv; 32	-
ResStack03	8×8 conv; 64	-
AvgPooling	8×8	-
Full01	C	-

IV. EXPERIMENTAL RESULTS

The experiments were carried out in a dedicated server consisting of two Intel Xeon Gold 6140 CPUs (36 physical cores), 126GB RAM, and four NVIDIA GeForce GTX 1080Ti GPUs, running on Debian GNI/Linux 11 (bullseye) kernel 5.10.0-19-amd64 x86_64. CNN models were implemented with Python (v.3.10) and libraries OpenCV (v.4.1), TensorFlow (v.1.15), CUDA (v.10.1), and cuDNN (v.7.5).

A. Experimental Setup

The applicability and quality of the Synthetic-sPVDB [4] and NS-PVDB [5] datasets were validated in [3]. In the present work, we propose replacing the pre-training of CNN models on the Imagenet database using synthetic palm vein datasets. To carry out the experiments, a closed-set scheme was used. Firstly, we trained both CNN models on the synthetic datasets using 80% of the samples per individual and varying the number of subjects (i.e., 2000, 5000, 10000, 15000, and 20000). Thus, we can know what is the recognition performance while the size of the database increases, testing the obtained model on the remaining 20% of samples. Later, we used the resulting pre-trained models as a feature extractor for a KNN classifier ($k = 1$) to evaluate their recognition performance on real databases. In all cases, an inter-session protocol was implemented using the samples from the first session for training (fine-tuning) and the other for testing. Therefore, the first acquisition session samples were divided into 50% for training and 50% as a biometric gallery, whereas the second session was used as a testing set. For the IITI dataset that has only one session, the same protocol proposed by the database authors was used. To measure the recognition performance, we applied the standard Rank-1 recognition rate, i.e., the percentage of true classified samples.

From the training set, 90% of samples were randomly separated for training and 10% for validation. The training process of the CNN models was carried out for 100 epochs with batch size $batchsize = 32$, and an early stopping criterion was set in case the validation error does not improve for seven epochs. We started from scratch using a learning rate $lr = 0.01$ with a reduction by a factor of 0.2 when the loss function remains unchanged for five epochs. For the final fine-tuning process on real databases, the initial learning rate was reduced by a factor of 0.1. For SGD optimization, the *momentum* = 0.9 and the *weight decay* = 1×10^{-5} were used.

B. Performance Results on Synthetic Databases

Tables V and VI show the results of the training, validation, and testing stages of each model in both synthetic datasets. The size of the databases varied with respect to the number of subjects. For further comparisons on the real datasets, partitions of similar size were established, i.e., 2000, 5000, 5000, 10000, 15000, 15000, 16000, and 20000. The last partition corresponds to the total size of each database. It can be appreciated the high recognition rates achieved at all stages, showing high generalization in validation and testing. These results were somewhat expected since the databases' intra-class variability is low because they do not simulate different acquisition sessions. Comparing both models, slightly better results were achieved with the SingleNet model on Synthetic-sPVDB, whereas ResNet was slightly better on NS-PVDB. Regarding the result variations in relation to the database size, in both cases and for both models, the best performance was obtained for 2000 individuals. However, no significant differences are seen as the number of subjects increases.

C. Transfer Learning Results on Real Databases

The resulting pre-trained models on both synthetic databases were used as feature encoders to evaluate their performance on the real databases using a KNN classifier. The feature descriptors were obtained from the output of the last convolutional layer, eliminating all the *fully-connected* layers of the network. The KNN classifier was implemented using the Euclidean

TABLE V
RECOGNITION PERFORMANCE OF BOTH CNN ARCHITECTURES BY VARYING THE SIZE OF THE SYNTHETIC-SPVDB DATABASE.

Subjects	SingleNet			ResNet		
	Train	Val	Test	Train	Val	Test
2,000	99.90	99.94	99.92	99.71	99.63	99.90
5,000	99.89	99.92	99.91	99.70	99.65	99.89
10,000	99.87	99.91	99.90	99.68	99.60	99.84
15,000	99.85	99.88	99.89	99.65	99.59	99.82
20,000	99.82	99.84	99.86	99.61	99.55	99.78

TABLE VI
RECOGNITION PERFORMANCE OF BOTH CNN ARCHITECTURES BY VARYING THE SIZE OF THE NS-PVDB DATABASE.

Subjects	SingleNet			ResNet		
	Train	Val	Test	Train	Val	Test
2,000	99.48	99.85	99.80	99.87	99.88	99.80
5,000	99.46	99.79	99.71	99.85	99.86	99.79
10,000	99.45	99.80	99.75	99.88	99.90	99.85
15,000	99.43	99.78	99.72	99.84	99.89	99.83
16,000	99.45	99.84	99.74	99.86	99.85	99.82

distance and a $k = 1$. The experiment allows knowing the degree of generalization of the models obtained by transfer learning and the validation of the work's main objective.

Tables VII and VIII summarize the performance results of both models pre-trained on Synthetic-sPVDB and NS-PVDB, respectively. For comparison purposes, we show the baseline results obtained in [27], which are the same models trained from scratch on the real databases. Besides, the results of the pre-trained model without fine-tuning (Pre-trained) and after fine-tuning (Fine-tuned) are shown. It is worth noting that for the sake of clarity, only the results of the pre-trained models on the entire synthetic database are shown. Contrary to the previous section, where no significant differences are observed as the number of subjects increases with a slight downward trend, it was found that the greater the number of individuals in the synthetic database, the better the performance of the pre-trained model on the real databases. This behavior is explained by the fact that the higher the number of subjects, the model learns a better feature representation with a higher inter-class variability, which then becomes a higher generalization degree.

In all cases, it can be seen that while the results without fine-tuning are inferior to the baseline, slightly superior results are obtained after fine-tuning. For the pre-trained model on the Synthetic-sPVDB database, the SingleNet model outperformed ResNet in three of the six databases (i.e., PolyU, PUT, VERA). On the other hand, the pre-trained ResNet model on NS-PVDB was slightly better in all databases except VERA.

D. Comparison against Pre-trained Model on Imagenet

In order to compare the results obtained with other state-of-the-art works, Table IX summarizes the results achieved by CNN models using transfer learning on real databases. All of the SOTA works reported used pre-trained models on the Imagenet database and their results were collected from the papers' results. Moreover, as can be seen in Table II, most of them used hybrid models [13], [21]–[23] much more complex than a CNN architecture, and the remaining ones [14], [24] used deeper architectures than those used in the proposed methodology. In all cases, the results of the pre-trained models on the synthetic databases are higher than those pre-trained on Imagenet. These results suggest that the feature representations learned on synthetic palm vein images are more specific to the classification problem, in contrast to the large variety of images contained in Imagenet.

TABLE VII
RECOGNITION PERFORMANCE OF BOTH MODELS PRE-TRAINED ON SYNTHETIC-SPVDB DATABASE AND FINE-TUNED ON REAL DATASETS: (A) SINGLENET ARCHITECTURE AND (B) RESNET ARCHITECTURE.

Database	SingleNet		
	Baseline	Pre-trained	Fine-tuned
CASIA-940	97.75	90.50	98.42
PolyU	100	98.32	100
PUT	99.00	96.25	99.86
VERA	99.77	96.18	99.88
Tongji	99.87	95.76	99.76
IITI	98.38	96.58	98.94

(a)

Database	ResNet		
	Baseline	Pre-trained	Fine-tuned
CASIA-940	98.75	96.50	99.12
PolyU	99.92	99.90	99.94
PUT	96.33	99.00	99.85
VERA	99.32	96.82	99.48
Tongji	99.96	97.67	99.98
IITI	97.93	94.86	98.62

(b)

TABLE VIII
RECOGNITION PERFORMANCE OF BOTH MODELS PRE-TRAINED ON NS-PVDB DATABASE AND FINE-TUNED ON REAL DATASETS: (A) SINGLENET ARCHITECTURE AND (B) RESNET ARCHITECTURE.

Database	SingleNet		
	Baseline	Pre-trained	Fine-tuned
CASIA-940	97.75	92.14	98.26
PolyU	100	98.06	99.90
PUT	99.00	95.70	99.12
VERA	99.77	96.10	99.80
Tongji	99.87	93.44	99.88
IITI	98.38	96.28	98.66

(a)

Database	ResNet		
	Baseline	Pre-trained	Fine-tuned
CASIA-940	98.75	94.32	99.20
PolyU	99.92	99.46	100
PUT	96.33	99.10	99.18
VERA	99.32	96.94	99.60
Tongji	99.96	97.70	99.90
IITI	97.93	96.40	98.90

(b)

V. CONCLUSIONS

In this work, we have proposed a new transfer learning approach for palm vein recognition. Two end-to-end CNN models were trained from scratch on two different synthetic palm vein databases and then evaluated on real databases. It was found that while the performance of the scratch models remains practically unchanged as the size of the synthetic

TABLE IX
COMPARISON OF THE OBTAINED RESULTS AGAINST OTHER STATE-OF-THE-ART WORKS. THE BEST RESULTS ARE HIGHLIGHTED IN BOLD FONT.

Approach	CNN Architecture	CASIA	PolyU	PUT	VERA	Tongji	IITI
Wulandari *et al.* [13]	AlexNet	-	-	93.92	-	-	-
Thapar *et al.* [21]	U-net + triplet Siamese	85.16	98.78	-	-	-	97.47
Lefkovits *et al.* [14]	ResNet-50	-	-	99.83	-	-	-
Yuan *et al.* [22]	AlexNet	97.00	**100**	-	-	-	-
Kuzu *et al.* [23]	DenseNet-161	-	99.69	-	-	-	-
Jia *et al.* [24]	EfficientNet	-	**100**	-	-	99.00	-
Our Proposal (Synthetic-sPVDB)	SingleNet	98.42	**100**	**99.86**	**99.88**	99.76	98.94
	ResNet-32	99.12	99.94	99.85	99.48	**99.98**	**99.62**
Our Proposal (NS-PVDB)	SingleNet	98.26	99.90	99.12	99.80	99.88	98.66
	ResNet-32	**99.20**	**100**	99.18	99.60	99.90	98.90

database increases, the greater the number of synthetic individuals, the better the learned feature representation on real databases. On the other hand, the proposed methodology improved the recognition performance by using transfer learning on synthetic databases, outperforming the state-of-the-art results obtained by other authors. The results show the usefulness of synthetic databases to quickly estimate the efficiency, scalability, and performance of new biometric recognition methods.

As future work, we plan to investigate the performance of other shallow CNN architectures. In this way, scalability studies on new, larger versions of the synthetic databases could be conducted. Furthermore, it is also required to develop new real databases containing a larger number of subjects.

ACKNOWLEDGMENT

R.H.-G. thanks to the Research Project ANID FONDECYT Iniciación en Investigación 2022 No. 11220693 "End-to-end multi-task learning framework for individuals identification through palm vein patterns", Ministerio de Ciencia, Tecnología, Conocimiento e Innovación, Gobierno de Chile. Portions of the research in this paper used the CASIA-MS-PalmprintV1 collected by the Chinese Academy of Sciences' Institute of Automation (CASIA). Portions of the research in this paper used the VERA-Palmvein Corpus made available by the Idiap Research Institute, Martigny, Switzerland.

REFERENCES

[1] W. Wu, S. Elliott, S. Lin, S. Sun, and Y. Tang, "Review of palm vein recognition," *IET Biometrics*, vol. 9, no. 1, pp. 1–10, 2019.

[2] P. Tome and S. Marcel, "Palm vein database and experimental framework for reproducible research," in *2015 International Conference of the Biometrics Special Interest Group (BIOSIG)*. IEEE, 2015, pp. 1–7.

[3] E. H. S. Jurado, R. Hernández-García, K. Vilches-Ponce, R. J. Barrientos, M. Mora, and G. Jaswal, "Towards the generation of synthetic images of palm vein patterns: A review," *Information Fusion*, 2022.

[4] E. Salazar, R. Hernández-García, R. J. Barrientos, K. Vilches, M. Mora, and A. Vásquez, "Generating style-based palm vein synthetic images for the creation of large-scale datasets," in *11th International Conference of Pattern Recognition Systems (ICPRS 2021)*. Institution of Engineering and Technology, 2021, pp. 182–187(5).

[5] ——, "Automatic generation of synthetic palm vein images: a nature-based approach," in *11th International Conference of Pattern Recognition Systems (ICPRS 2021)*. Institution of Engineering and Technology, 2021, pp. 38–43(5).

[6] Y. Hao, Z. Sun, T. Tan, and C. Ren, "Multispectral palm image fusion for accurate contact-free palmprint recognition," in *2008 15th IEEE International Conference on Image Processing*. IEEE, 2008, pp. 281–284.

[7] R. Kabaciński and M. Kowalski, "Vein pattern database and benchmark results," *Electronics Letters*, vol. 47, no. 20, pp. 1127–1128, 2011.

[8] D. Zhang, Z. Guo, G. Lu, L. Zhang, and W. Zuo, "An online system of multispectral palmprint verification," *IEEE transactions on instrumentation and measurement*, vol. 59, no. 2, pp. 480–490, 2009.

[9] L. Zhang, Z. Cheng, Y. Shen, and D. Wang, "Palmprint and palmvein recognition based on DCNN and a new large-scale contactless palmvein dataset," *Symmetry*, vol. 10, no. 4, p. 78, 2018.

[10] S. Bhilare, G. Jaswal, V. Kanhangad, and A. Nigam, "Single-sensor hand-vein multimodal biometric recognition using multiscale deep pyramidal approach," *Machine Vision and Applications*, vol. 29, no. 8, pp. 1269–1286, 2018.

[11] Ö. Toygar, F. O. Babalola, and Y. Bitirim, "FYO: a novel multimodal vein database with palmar, dorsal and wrist biometrics," *IEEE Access*, vol. 8, pp. 82 461–82 470, 2020.

[12] A. Makrushin, A. Uhl, and J. Dittmann, "A survey on synthetic biometrics," *IEEE Access*, 2023.

[13] M. Wulandari, D. Gunawan *et al.*, "On the performance of pretrained cnn aimed at palm vein recognition application," in *2019 11th International Conference on Information Technology and Electrical Engineering (ICITEE)*. IEEE, 2019, pp. 1–6.

[14] S. Lefkovits, L. Lefkovits, and L. Szilágyi, "Applications of different cnn architectures for palm vein identification," in *Modeling Decisions for Artificial Intelligence: 16th International Conference, MDAI 2019, Milan, Italy, September 4–6, 2019, Proceedings 16*. Springer, 2019, pp. 295–306.

[15] L. Lu, R. Fu, L. Yuan, W. Chen, Y. Liu *et al.*, "Palm vein recognition based on end-to-end convolutional neural network," *Journal of Southern Medical University*, vol. 39, no. 2, pp. 207–214, 2019.

[16] Y.-Y. Chen, S.-Y. Jhong, C.-H. Hsia, and K.-L. Hua, "Explainable ai: A multispectral palm-vein identification system with new augmentation features," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 17, no. 3s, pp. 1–21, 2021.

[17] H. Qin, M. A. El-Yacoubi, Y. Li, and C. Liu, "Multi-scale and multi-direction gan for cnn-based single palm-vein identification," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 2652–2666, 2021.

[18] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, pp. 1097–1105, 2012.

[19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[20] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*. IEEE, 2009, pp. 248–255.

[21] D. Thapar, G. Jaswal, A. Nigam, and V. Kanhangad, "PVSNet: Palm Vein Authentication Siamese Network Trained using Triplet Loss and Adaptive Hard Mining by Learning Enforced Domain Specific Features," in *2019 IEEE 5th International Conference on Identity, Security, and Behavior Analysis (ISBA)*. IEEE, 2019, pp. 1–8.

[22] L. Yuan, M. Lou, Y. Liu, F. Yang, and J. Huang, "Palm vein classification based on deep neural network and random forest," *Laser & Optoelectronics Progress*, vol. 56, no. 11, p. 101010, 10 2019.

[23] R. S. Kuzu, E. Maiorana, and P. Campisi, "Vein-based biometric verification using densely-connected convolutional autoencoder," *IEEE Signal Processing Letters*, vol. 27, pp. 1869–1873, 2020.

[24] W. Jia, J. Gao, W. Xia, Y. Zhao, H. Min, and J.-T. Lu, “A Performance Evaluation of Classic Convolutional Neural Networks for 2D and 3D Palmprint and Palm Vein Recognition,” *International Journal of Automation and Computing*, pp. 1–27, 2021.

[25] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, “Analyzing and Improving the Image Quality of StyleGAN,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 8110–8119.

[26] Y. Liu, C. Gao, Z. Zhang, Y. Wu, M. Liang, L. Tao, and Y. Lu, “A new multi-agent system to simulate the foraging behaviors of physarum,” *Natural computing*, vol. 16, no. 1, pp. 15–29, 2017.

[27] J. I. Santamaría, R. Hernández-García, R. J. Barrientos, F. M. Castro, J. Ramos-Cózar, and N. Guil, “Evaluation of end-to-end CNN models for palm vein recognition,” in *2021 40th International Conference of the Chilean Computer Science Society (SCCC)*. IEEE, 2021, pp. 1–8.