

Pose Invariant Palm Vein Identification System using Convolutional Neural Network

Nidaa Flaih Hassan

Husam Imad Abdulrazzaq

Received 24/4/2018, Accepted 15/10/2018, Published 9/12/2018



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

Abstract:

Palm vein recognition is a one of the most efficient biometric technologies, each individual can be identified through its veins unique characteristics, palm vein acquisition techniques is either contact based or contactless based, as the individual's hand contact or not the peg of the palm imaging device, the needs a contactless palm vein system in modern applications rise tow problems, the pose variations (rotation, scaling and translation transformations) since the imaging device cannot aligned correctly with the surface of the palm, and a delay of matching process especially for large systems, trying to solve these problems. This paper proposed a pose invariant identification system for contactless palm vein which include three main steps, at first data augmentation is done by making multiple copies of the input image then perform out-of-plane rotation on them around all the X,Y and Z axes. Then a new fast extract Region of Interest (ROI) algorithm is proposed for cropping palm region. Finally, features are extracted and classified by specific structure of Convolutional Neural Network (CNN). The system is tested on two public multispectral palm vein databases (PolyU and CASIA); furthermore, synthetic datasets are derived from these mentioned databases, to simulate the hand out-of-plane rotation in random angels within range from -20° to $+20^{\circ}$ degrees. To study several situations of pose invariant, twelve experiments are performed on all datasets, highest accuracy achieved is $99.73\% \pm 0.27$ on PolyU datasets and $98\% \pm 1$ on CASIA datasets, with very fast identification process, about 0.01 second for identifying an individual, which proves system efficiency in contactless palm vein problems.

Key words: biometrical identification, contactless Palm vein, convolutional neural network, region of interest

Introduction:

The traditional personal identification or verification systems like (passwords, ID cards, etc.) have become inefficient and unable to meet the needs of current society, since it can be stolen or lost. For these reasons the biometrics identification systems becoming the focus of the research in recent years.

Palm vein is a one of the most interesting type of biometric technologies, it is like a palmprint, but rather than using visible light spectrum for capturing palmprint, palm vein needs Near Infra-Red (NIR) illumination for capture the vein pattern that hidden under palm skin. Each person has unique characteristics of the veins that can be used for identification or verification. Comparing with the other biometric techniques, palm vein is (1) (2):

- less costly
- Easy acquiring

Department of Computer Science, University of Technology, Baghdad, Iraq

*Corresponding

0111545@student.uotechnology.edu.iq

author:

Difficult to fake or change because it's unique pattern inside the human's body.

There are two types of palm vein acquisition techniques: contact based and contactless based, the contactless is more convenient for individuals, but without the peg guide the individual hand in contactless acquisition method, there will be pose variations (rotation, scaling and translation transformation) in palm vein image affected by movement of the hand.

There are several researches involved palm vein and contactless palmprint, Methani C, Namboodiri A. in 2009 (3) proposed a method involved with variations in pose (out-of-plane rotations, translation) in contactless and unconstrained palmprint imaging, a similarity measure between the reference image and corrected test image is computed. Image alignment technique is proposed to distinguish the orientation of the palm. Experiments are applied on a real and a synthetic dataset, Mirmohamadsadeghi L, Drygajlo A. in 2014 (4) presented a new approach depends

on local texture patterns for palm vein patterns using histograms and operators of multi-scale Local Binary Patterns (LBPs), then higher-order local pattern descriptors is investigated by using histogram of Local Derivative Pattern (LDP), Abbas M, George E. in 2014 (5) suggested a palm vein recognition system using spatial energy distribution of wavelet sub-bands, Discrete Haar wavelet (DHW) is applied, then the average energy distribution for each sub-band is computed, feature vector is gained by concatenated these sub-bands, Wang R, *et al.* in 2014 (6) presented palm vein identification method depending on Gabor wavelet, in the beginning, contrast limited adaptive histogram equalization (CLAHE) used for enhances the contrast and image skeletonization for vein thinning, then Gabor wavelet transform-based method used for feature extraction. Jalali A, Lee M. In 2015 (7) proposed a contactless palmprint biometric system using features of palm texture from an image acquired by ordinary digital camera. Meanwhile convolutional neural network (CNN) is applied for palmprint recognition. Dian L, Dongmei S. in 2016 (8) a palmprint recognition approach using CNN are proposed, at first improved fuzzy enhancement algorithm is applied for preprocessing; then the features are extracted using the AlexNet (a specific structure of CNN), finally, the features are matched using the hausdorff distance. Ali H, Razuqi N. in 2017 (2) proposed a Palm Vein Recognition System which is based on Centerline. The proposed system extracted centerline of vein depending on distance from boundary (DFB), the features are set of key points which generated based on the Difference-of-Gaussian (DoG), the absolute difference between two features is used for matching, the system is tested with PolyU multi-spectral palmprint database.

This paper proposed a new pose invariant contactless palm vein identification system, intended to identify palm vein images captured with pose variations in 3D space which cause an out-of-plane rotation, scale and translation variations, at first data augmentation is done by making multiple copies of the input image then perform on them out-of-plane rotation around all the X, Y and Z axes; this step would reduce the samples of palm needed to be captured from an individual, then a proposed fast extract region of interest ROI algorithm is applied on palm vein images for cropping palm region, histogram equalization used to enhance appearance of images, features are extracted and classified by special structure of CNN. To evaluate the proposed system, two public palm vein databases (PolyU and CASIA) are used, each one is divided into training and testing portions; furthermore, for simulating

hand out-of-plan rotation, synthetic datasets are derived from the testing portion. These testing datasets are used to measure the validation accuracy of the proposed system.

Convolutional Neural Network (CNN)

CNN is a class of deep, feed-forward artificial neural networks, the most important characteristics of CNN is the local connectivity and the using of shared weights; thus, it can learn local features of the input image (8).

In image classification, the CNN is very effective and fast method in recognition phase, even in partly rotation, distortion, scaling and translation variations, but it is slow in training phase, so training some thousands of images could take several hours.

Back propagation is used for training the CNN by adjusting the weights of filters (kernels). The weights are updated using Eq. (1):

$$(w_{jk})_{new} = (w_{jk})_{old} - \eta * \Delta w_{jk} \quad \dots \dots \dots (1)$$

Where w_{jk} is the weights of the filters, η is the learning rate equal to 0.0001, and Δw_{jk} is the derivative of error regarded to weights (7).

The typical CNN structure consists of the following layers (9):

- **Convolutional Layers:** takes several feature maps as input and produce n of feature maps as output, where n is the number of learnable filters in the convolution layer. Filter weights are adapted using back propagation depending on training data. The number n of filters as well as the filter's size $k_w \times k_h$ are hyper parameters of convolutional layers. The outputs of this layer go through Rectified Linear Unit (ReLU) that is the common choices for non-linear activation function.
- **Pooling Layers:** Pooling summarizes a $p \times p$ area of the input feature map. Pooling can be used with a stride of $s \in \mathbb{N}_{\geq 2}$, pooling layers are sometimes also called *subsampling layers* is applied for three reasons: To get local translational invariance, to get invariance against minor local changes and, most important, for data reduction to $(1/s^2)$ th of the data.
- **Fully-Connected Layer:** Fully Connected layer is a traditional multi-layer perceptron where every neuron in a layer is connected to each neuron on the following layer. The work of the fully connected layer is to classify the features extracted by previews (convolutional and pooling layers) to the proper classes. Dropout technique can be applied over this layer, it is used to prevent over fitting by setting the output of any neuron to zero with probability p .

The Problem of Contactless Palm Vein

Some of the existing palm vein recognition system employed a special peg to keep the hand of the user stable, to take well aligned palm images, for modern biometric applications, such as mobile devices the pegs can't be used, so the palm images will be captured with variations in 3D space. The position, direction, out-of-plane rotation and degree of stretching may affect the image accuracy. Thus, it is essential for any system to be robust to the translation, position, direction, stretching, and

rotation variations (3)(7). Fig. 1 shows a multiple spectral imaging device without peg guide where the hand is free in 3D space (10).

Other problem in contactless system is matching speed, especially on the systems that use traditional methods for matching although it has hundreds of enrolled individuals, its maybe not convenient to user still raises his hand waiting for identifying process, so the system should be fast enough.

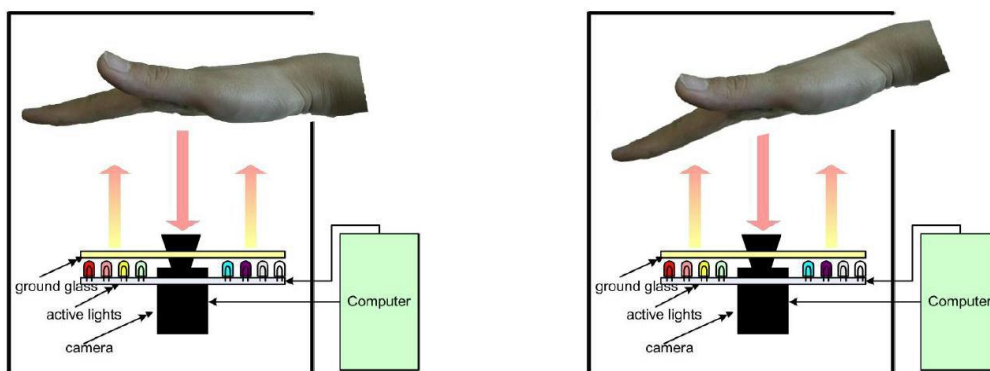


Figure 1. multiple spectral imaging device without peg guide and show hands with different poses (10)

Datasets

In this paper, real and synthetic datasets is derived from two well-known multi-spectral palm veins databases, these datasets are categorized as the following:

- **Real datasets**

1. **CASIA dataset (contactless):** consist of 600 images derived from the CASIA multi-spectral palmprint image database V1.0 (10), it is composed

of palm vein images of 100 individuals (six images for each individual), captured using a multiple spectral imaging device which it without guide pegs so there is a certain degree of variations of hand pose. The 600 images that used here comprise all samples from all individual's right hands captured under 940 nm near-infrared (NIR) illuminations, some of its samples shown in Fig. 2.



Figure 2. Samples from CASIA dataset

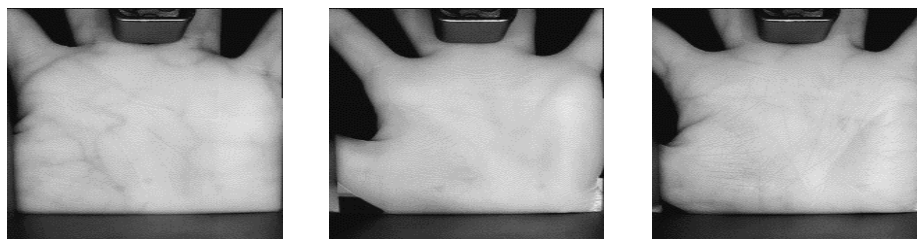


Figure 3. Samples from PolyU dataset

2. **PolyU dataset (contact based):** consist of 3000 NIR illuminations images derived from the PolyU multi-spectral palm print database (11), this database composed of NIR palm samples collected

from 250 volunteers, 24 images for each volunteer, 12 images of right palm and 12 of the left. The dataset which used here comprises all images of all individual's left palms , samples shown in Fig. 3

• **Synthetic Dataset for Testing Phase**

Synthetic datasets are generated for simulating an out-of-plane rotation (which represent hands rotates freely and randomly in 3D space), the palm images in testing datasets which derived from

CASIA and PolyU databases are rotated around all, X and Y axes (Z in plane axis), by a random angle for each axis in the range (-20° to +20°), Fig. 4 shows samples from the synthetic datasets notated with its rotation angels which chosen randomly.

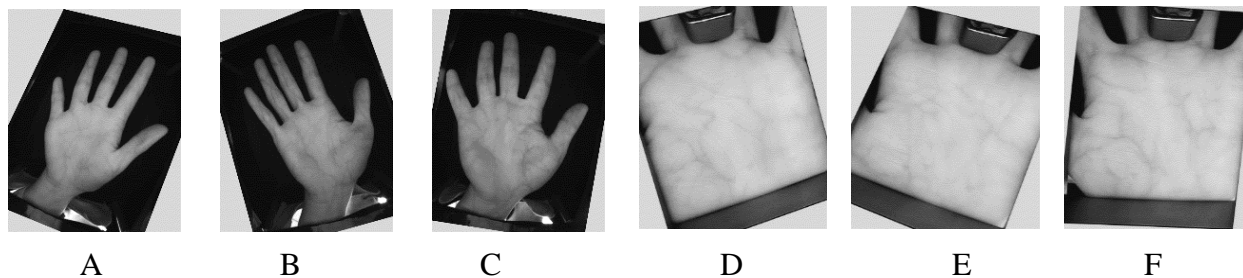


Figure 4. Samples from Synthetic Dataset for testing phase, A: palm image from CASIA database synthetically Rotated (+14° on X, -19° on Y , +18° on Z axes), B: Rotated (-15° on X, +7° on Y , -19° on Z axes), C: Rotated (-7° on X, -1° on Y , +10° on Z axes), D: palm image from PolyU database synthetically Rotated (+17° on X, 0° on Y , -16° on Z axes), E: Rotated (+14° on X, -19° on Y , +18° on Z axes), F: Rotated (-20° on X, +10° on Y , +3° on Z axes).

The Proposed System

In this section, the proposed pose invariant contactless palm vein system is presented, it included two major phases; enrollment phase, & identification phase. Enrollment phase is the training phase: in which, all training palm vein images of all individuals are simultaneously pass in the system for training the CNN. Training phase may take about an hour to produce a trained CNN. Identification phase is the testing phase; in which, the trained system can directly identify any palm vein of enrolled individuals.

The proposed system is composed of the following steps: Data Augmentation, Extract (ROI) (using a proposed algorithm), Histogram equalization,

Feature extraction and Classification (using a proposed CNN Structure)

• **Data Augmentation**

To reduce the number of sample images acquired from the individual and for simulating the hand out-of-plane rotation around all axes. The system preforms data augmentation step on the input image (in training phase only) by automatically generating 7 different poses copies of it, one image without any rotation, three images rotated with (+20°) angle around X, Y and Z axes, and the last three images rotated (-20°) around same axes, this augmentation is shown in Fig. 5.

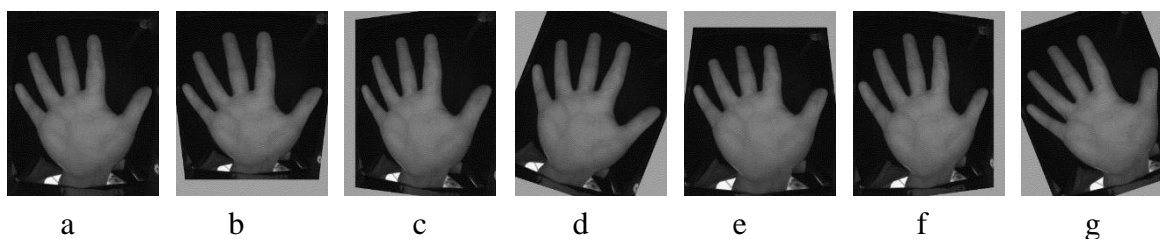


Figure 5. One palm image from CASIA Database after Data Augmentation, (a): original image, (b), (c) and (d) images rotated with (+20°) angle around each X, Y and Z axes, (e), (f) and (g) images rotated with (-20°) angle

• **Extract Region of Interest (ROI) Algorithm**

In this section, a new fast extract ROI algorithm is proposed to detect and crop the most interesting center region of the palm. The proposed algorithm as shown in *Algorithm 1*, is described in the following steps:

1. The input gray scale NIR palm vein image is captured from the imaging device or databases.

2. Convert the input image to binary using simple thresholding parameter

3. Apply Morphological open operation with a small filter size to remove the noise

4. Apply Morphological open operation with a big size filter commensurate with palm region size, which is estimated by computing the arithmetic mean of the binary image, this operation extracted

the exact palm region regardless of the zooming or translation.

5. Compute upper left point and lowest right point, from contours points of palm region.
6. Crop the palm region rectangle from the input gray level image.
7. Resize the cropped image to (64 pixel * 64 pixel).

This method is very fast, it takes about (0.007) second for one image, it's also invariant to translation, and scale variations, but not for rotation, since the rotation variations are solved next by CNN. Fig.6 displays the output images of each proposed ROI steps.

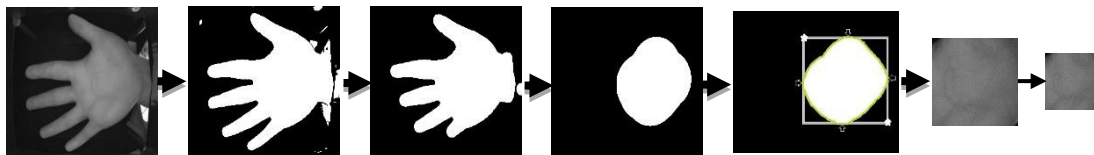


Figure 6. The output images of each ROI steps, each image corresponding to one of the algorithm steps

Algorithm 1. Fast extract ROI

Input: Input_img // (768×576) grayscale 8-bit NIR palm vein image

Output: ROI_img // (64×64) grayscale 8-bit image

Step_1 : Bin = threshold(Input_img)

Step_2 : C_Bin = Morphology_Open(Bin, filter_size(small value))

Step_3 : P_Ratio = Image_mean(C_Bin) \\ mean of all pixels values

Step_4 : R = P_Ratio * specific_rate \\ compute filter size commensurate with palm region size

Step_5 : O_img = Morphology_Open(C_Bin, filter_size(R,R))

Step_6 : Cnts = findContours(O_img)

Step_7 : U_L_point = point(Min_X (Cnts) , Min_Y (Cnts))

Step_8 : L_R_point = point(Max_X (Cnts) , Max_Y (Cnts))

Step_9 : Crop_img = Crop(Input_img , Region(U_L_point, L_R_point))

Step_10 : ROI_img = Resize(Crop_img , (64× 64))

• Histogram Equalization

Normalizes the brightness and increase the contrast of ROI palm image using a traditional Histogram Equalization method. The aim of this stage is to speed up the training phase of CNN, since it was observed that this step reducing the number of epochs needed to get the same results.

• Feature Extraction and Classification using CNN

In this section, a special structure of convolutional neural network CNN is presented as shown in Fig. 7. The CNN combines feature extraction and classification steps in a single trainable module are described in the following:

1. *Feature extraction*, consists of eight succeeded Convolutional and Max pooling layers takes input image (64×64×1) and output (4×4×256) neurons represent the feature map.

2. *Classification*, consists of two fully connected layers, the first one size is 1024 neuron and the second one size is equal to the number of enrolled individuals. The output of the last fully connected layer input to a SoftMax activation function that assigns a probability for each individual class for matching.

Table 1 explains the layers of the proposed CNN structure, the second column in the table contains layer type followed by its activation function (layer 1 to layer 7), while layer 8 and 9 are followed by dropout technique with 0.5 rate, third and fourth columns describe the size of feature map matrix, and the last column shows the kernel (filter) size for each layer with stride. After several training epochs a trained CNN will be ready for identifying any test sample.

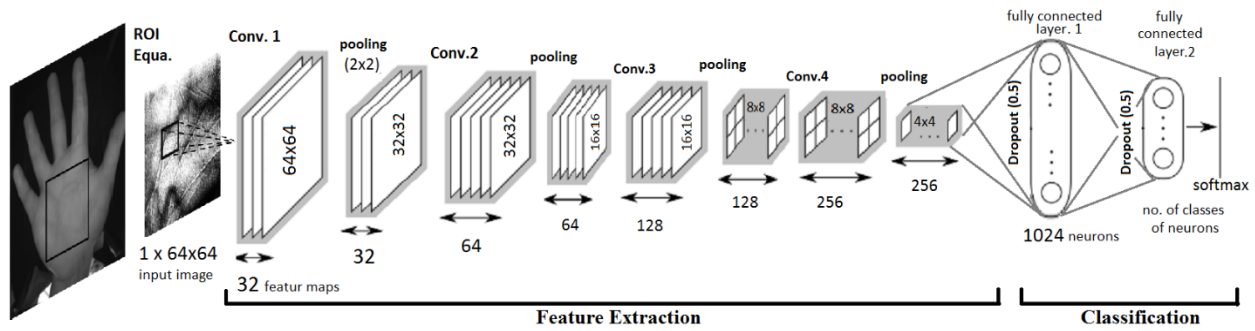


Figure 7. The proposed structure of CNN with size of feature maps and kernels

Table 1. The proposed structure of CNN, the input and output sizes are described in (rows*cols*#kernels). The kernel is specified as rows*cols; stride

no	Layers	Input size	Output size	kernel
1	Convolution 1 + ReLU	64×64×1	64×64×32	3×3; 1
2	max pooling	64×64×32	32×32×32	2×2; 2
3	Convolution 2 + ReLU	32×32×32	32×32×64	3×3; 1
4	max pooling	32×32×64	16×16×64	2×2; 2
5	Convolution 3 + ReLU	16×16×64	16×16×128	3×3; 1
6	max pooling	16×16×128	8×8×128	2×2; 2
7	Convolution 4 + ReLU	8×8×128	8×8×256	3×3; 1
8	max pooling + Dropout (0.5)	8×8×256	4×4×256	2×2; 2
9	Fully connected + ReLU + Dropout (0.5)	4096	1024	
10	Fully connected + SoftMax	1024	number of classes (individuals)	

Results and Discussion:

The system is implemented using Python programming language and TensorFlow open-source framework. The classification accuracy metric (i.e., the proportion of correct predictions) is used for evaluating the proposed system. The considered *training:testing* ratio is 3:1 for PolyU dataset, and 5:1 for CASIA dataset. Six experiments are applied on each PolyU and CASIA datasets, the results are presented in Tables 2 and 3. The first column in these tables describes the training approach of the

system used for each experiment notate if there is data augmentation with specific angle is applied on the input image; the second column notates which dataset categorize are used for testing, last column resulted the mean classification accuracy for validation process \bar{x} standard error.

Fig.8 shows validation accuracy plot which displays the increasing of accuracy during training epochs, of experiment 2 and experiment 6, for both CASIA and PolyU datasets.

Table 2. Result of the six experiments on PolyU datasets.

Exp. No.	Training approach of the system	Testing dataset	Validation accuracy
1	No Augmentation	Real	99.33% \bar{x} 0.13
2	Augmentation (10°)	Real	99.73% \bar{x} 0.27
3	Augmentation (10°)	Synthetic (10°)	99.6 % \bar{x} 0.27
4	No Augmentation	Synthetic (10°)	95.86 % \bar{x} 0.8
5	Augmentation (10°)	Synthetic (20°)	96.8 % \bar{x} 0.9
6	Augmentation (20°)	Synthetic (20°)	99.6 % \bar{x} 0.27

Table 3. Results of the six experiments on CASIA datasets.

Exp. No.	Training approach of the system	Testing dataset	Validation accuracy
1	No Augmentation	Real	93 % \bar{x} 2
2	Augmentation (10°)	Real	98 % \bar{x} 1
3	Augmentation (10°)	Synthetic (10°)	97 % \bar{x} 2
4	No Augmentation	Synthetic (10°)	87 % \bar{x} 2
5	Augmentation (10°)	Synthetic (20°)	87 % \bar{x} 2
6	Augmentation (20°)	Synthetic (20°)	95 % \bar{x} 2

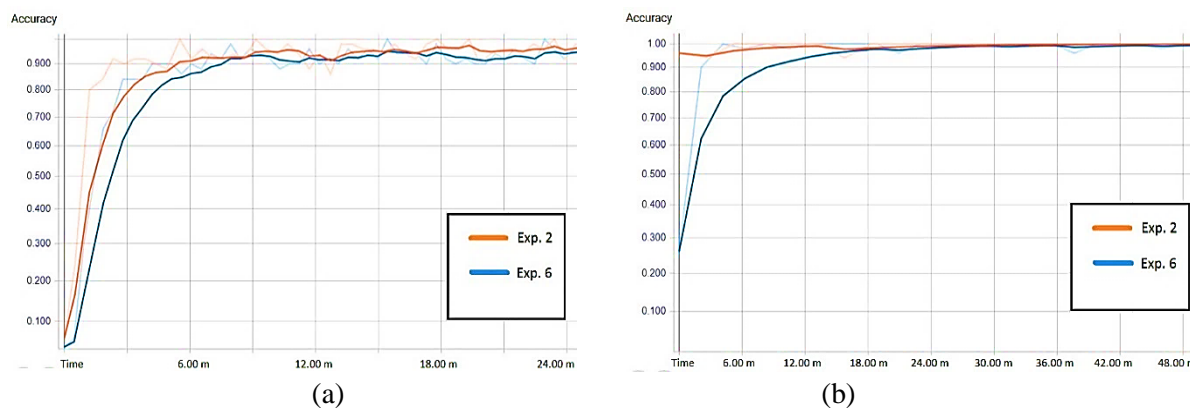


Figure 8. Validation accuracy plot, display the increasing of accuracy during the training time, comparing between experiment 2 and experiment 6, (a) tensor-board of CASIA dataset, (b) tensor-board of PolyU dataset

Experimental results demonstrate that the proposed system is very effective for palm vein identification. The first experiment presented in Table 1 and 2, shows the accuracy resulted when the proposed system ordinary implemented without any data augmentation and synthetic contactless simulation. The accuracy in the second experiment achieved best result; this proves that the synthetically augmentation of input palm image with new variant poses would highly improve the accuracy of system. In the sixth experiment, acceptable accuracy is achieved, this indicates that the system handling a sensible pose variation of the tested palms that could be rotated in some angles in all directions.

In general, palm vein recognition was addressed in (2)(4)(5)(6) papers, when traditional methods are used and without simulation for contactless palm. To our knowledge, this is the first time CNN method are used with contactless palm vein, and this is the first time a synthetically simulation for contactless palm vein are performed.

Despite the results of the above mentioned papers may be equal or slightly higher than the result of the proposed system, but the obtained accuracy is very good even it has been experimented on a synthetic dataset that simulates a high degree of pose variation, also the using of proposed fast ROI algorithm with CNN result a very fast identification process, since the overall time needed for identifying one palm from the input to ROI step until the output is about (0.01) second, this time is fixed for all experiment since the proposed structure of CNN is fixed.

Conclusions:

1- The proposed system is a contactless palm vein identification system, which is robust to the translation, scaling, out-of-plane rotation variations.

- 2- Synthetic datasets are derived from CASIA and PolyU palm vein databases for simulating contactless palm, the proposed system achieves high validation accuracy rates 98%, 99.73% respectively, on these databases.
- 3- Using the new proposed fast ROI algorithm to avoid translation and scale variations, and CNN to avoid out-of-plane rotation variations in contactless palm vein, gives the system more advantages over the traditional systems.
- 4- The proposed system is very fast comparing with traditional systems, because the overall time needed for identifying one palm is about (0.01) second only.

Conflicts of Interest: None.

References:

1. Noh ZM, Ramli AR, Saripan M. I, Hanafi M. Overview and Challenges of Palm Vein Biometric System. *Int J Biom.* 2016;8(1):2–18.
2. Ali YH, Razuqi ZN. Palm Vein Recognition Based on Centerline. *Iraqi J Sci.* 2017;58(2A):726–34.
3. Methani C, Namboodiri A. Pose Invariant Palmprint Recognition. In: *International Conference on Biometrics [Internet]*. 2009. p. 577–86.
4. Mirmohamadsadeghi L, Drygajlo A. Palm vein recognition with local texture patterns. *IET Biometrics [Internet]*. 2014;3(4):198–206.
5. Abbas AMJ, George LE. Palm Vein Identification and Verification System Based on Spatial Energy Distribution of Wavelet Sub-Bands. *Int J Emerg Technol Adv Eng.* 2014;4(5):727–34.
6. Wang R, Wang G, Chen Z, Zeng Z, Wang Y. A palm vein identification system based on Gabor wavelet features. *Neural Comput Appl.* 2014;24(1):161–8.
7. Jalali A, Lee M. Deformation Invariant and Contactless Palmprint Recognition Using Convolutional Neural Network. In: *Proceedings of the 3rd International Conference on Human-Agent Interaction.* 2015. p. 209–12.
8. Dian L, Dongmei S. Contactless palmprint

recognition based on convolutional neural network.
2016 IEEE 13th Int Conf Signal Process. 2016;1363–
7.
9. Thoma M. Analysis and Optimization of

Convolutional Neural Network Architectures
[Internet]. 2017.
10. CASIA-MS-PalmprintV1. [Internet].
11. Multispectral PolyU database [Internet].

نظام تحديد الهوية من خلال اورددة اليد الثابتة الموضع باستخدام الشبكة العصبية التلافيفية

حسام عماد عبد الرزاق

نداء فليح حسن

قسم علوم الحاسوب، الجامعة التكنولوجية، بغداد، العراق

الخلاصة:

تميز اورددة راحة اليد واحدة من اكفا التقنيات البيولوجية حيث يمكن التعرف على هوية أي شخص من خلال الخصائص الفريدة للأوردة، ان تقنيات تصوير اورددة الكف اما تقنية ملامسة او تقنية الغير ملامسة حسب كون الكف تلامس سطح الجهاز عند التصوير من عدمه، ان حاجة التطبيقات الحديثة لنظم اورددة اليد الغير ملامسة قد ابرز مشكلتان، التغيير في الموضع للكف من ناحية (التدوير، الحجم، الموقع) حيث ان جهاز التقاط صورة الكف لا يمكنه دائما محاذاه سطح الكف الغير ملامس له، ومشكلة تأخر عملية المطابقة خاصة في الأنظمة الكبيرة. هذا البحث يقترح نظام تحديد الهوية من خلال اورددة راحة اليد الثابتة الموضع والذي يتضمن ثلاث خطوات رئيسية ، في البداية يتم تطبيق طريقة زيادة البيانات من خلال تدوير نسخ للصورة المدخلة حول جميع المحاور، ثم اكتشاف المنطقة المهمة في الكف (Region of Interest) من خلال خوارزمية جديدة مقترحة ، أخيرا استخلاص الخصائص والتصنيف يتم بواسطة استخدام هيكلية خاصة للشبكة العصبية التلافيفية (Convolutional Neural Network)، تم اختبار النظام على قاعدتي بيانات لأورددة اليد (PolyU و CASIA) وعلى قواعد بيانات اصطناعية اخرى تم انشائها من قاعدتي البيانات أنفة الذكر لمحاكاة تدوير اليد حول جميع المحاور بزوايا عشوائية ضمن نطاق -20° الى +20° درجة ، لدراسة عدة حالات من تغيير موقع الكف ، تم تطبيق 12 اختبارا على قواعد بيانات انفة الذكر، أظهرت النتائج ان اعلى دقة للنظام كانت $99.73 \pm 0.27\%$ على قاعدة بيانات PolyU و $98 \pm 1\%$ على قاعدة بيانات CASIA مع سرعة فائقة في عملية تحديد الهوية، حوالي 0.01 ثانية لعملية التعرف على شخص، وهذا يثبت كفاءة النظام في مشاكل اورددة راحة اليد الغير ملامسة.

الكلمات المفتاحية: تحديد الهوية البيومترية، اورددة اليد الغير ملامسة، شبكة عصبية تلافيفية، منطقة الاهتمام.