

Dense Hand-CNN: A Novel CNN Architecture based on Later Fusion of Neural and Wavelet Features for Identity Recognition

Elaraby A. Elgallad¹
Deanship of Information Technology
Tabuk University, KSA

Wael Ouarda², Adel M. Alimi³
Research Groups in Intelligent Machines
ENIS, BP 1173, Sfax, 3038, Tunisia^{2,3}

Abstract—Biometric recognition or biometrics has emerged as the best solution for criminal identification and access control applications where resources or information need to be protected from unauthorized access. Biometric traits such as fingerprint, face, palmprint, iris, and hand-geometry have been well explored; and matured approaches are available in order to perform personal identification. The work emphasizes the opportunities for obtaining texture information from a palmprint on the basis of such descriptors as Curvelet, Wavelet, Wave Atom, SIFT, Gabor, LBP, and AlexNet. The key contribution is the application of mode voting method for accurate identification of a person at the fusion decision level. The proposed approach was tested in a number of experiments at the CASIA and IITD palmprint databases. The testing yielded positive results supporting the utilization of the described voting technique for human recognition purposes.

Keywords—Deep learning; fusion; palmprint; squeezeNet; voting

I. INTRODUCTION

Biometrics is an authentication method that uses human characteristics to identify a person. Based on the traits, biometric can be divided into two broad types as in [1]: physical and behavioral. Physical biometric is a biometric system that evaluates the physical characteristic of a human body to recognize a person, such as fingerprint, face, retina, etc. On the other hand, behavioral characteristic analyzes the human behavioral traits, such as gait, signature, keystroke, etc. Behavioral biometric is less secured than physical biometric because people can change their behavior anytime they want. For example, people can adjust their signature, keystroke, or walking pattern easily. Today, multimodal biometric systems, which incorporate more than one biometric, with appropriate security measures are acknowledged as more robust and more accurate than unimodal biometrics, because even when the score of one biometric recognition is poor due to environmental conditions, the final outcome can be positive because the score from another biometric recognition is considered.

In the recent years, the computer vision has been increasingly dominated by deep learning, which has proven to have notable capacities to achieve top scores across various tasks and contests. ImageNet stands out as the most recognized among such contests. The ImageNet competition

tasks researchers with creating a model that most accurately classifies the given images in the dataset. Back in 2012, a paper from the University of Toronto was published [2]. The paper proposed to use a deep Convolutional Neural Network (CNN) for the task of image classification. It was relatively simple compared to those that are being used today. The main contributions that came from this paper were using a deep for large scale image classification. This was made possible because of the large amounts of labelled data from ImageNet, as well as training the model using parallel computations on two GPUs. They used ReLU for the non-linearity activation functions, finding that they performed better and decreased training time relative to the tanh function. The techniques of data augmentation that consisted of image translations, horizontal reflections, and mean subtraction are used. These techniques are very widely used today for many computers. Their proposed style of having successive convolution and pooling layers, followed by fully-connected layers at the end is still the basis of many state-of-the-art networks today. Basically, AlexNet set the bar, providing the baseline and default techniques of using CNNs for computer vision tasks.

The VGGNet paper came out in 2014, Simonyan and Zisserman [3], their main idea was that you didn't really need any fancy tricks to get high accuracy. Just a deep network with lots of small 3x3 convolutions and non-linearities will do the trick. They use of only 3x3 sized filters instead of the 11x11 used in AlexNet. The GoogleNet architecture [4] was the first to really address the issue of computational resources along with multi-scale processing.

Through the use of 1x1 convolutions before each 3x3 and 5x5, the inception module reduces the number of feature maps passed through each layer, thus reducing computations and memory consumption. GoogleNet introduces a new idea that CNN layers didn't always have to be stacked up sequentially. The authors of the paper showed that you can also increase network width for better performance and not just depth. Since its initial publication in 2015, ResNets have created major improvements in accuracy in many computer vision tasks [5].

The main contribution of residual learning in ResNet architecture is to show that a naive stacking of layers to make the network very deep won't always help and can actually make things worse. To address the above issue, they introduce residual learning with skip-connections. The idea is that by

using an additive skip connection as a shortcut, deep layers have direct access to features from previous layers. This allows feature information to more easily be propagated through the network. It also helps with training as the gradients can also more efficiently be back-propagated.

At the same time, the development of DenseNets significantly expanded the perspective of shortcut connections [6]. The above networks implement comprehensive cross-layer feed-forward pathway connections. Such an arrangement made it possible to outperform ResNets, as it enables each respective layer to employ as inputs the entire range of feature maps from prior layers, whereas the emerging maps serve as inputs for subsequent layers. To this end, DenseNets are associated with an ability to mitigate the vanishing gradient issue coupled with significant reductions in the parameter numbers, incentives for repeat feature use, and enhanced feature propagation. Moreover, a relatively small convolutional neural network (CNN) titled SqueezeNet designed by authors in [7]. They proved to require fifty times fewer parameters to produce a level of accuracy equal to that of AlexNet. On top of such an accomplishment, this CNN can be reduced to 0.5MB. This size is 510 times smaller than that of AlexNet. Correspondingly, such compressed architecture has a number of advantages, as it is applicable for less bandwidth, uses less inter-server communication in the course of training, and has more feasibility for implementation on field-programmable gate arrays (FPGAs), as well as other limited memory hardware.

The paper emphasizes the opportunities for obtaining texture information from a palmprint on the basis of such descriptors as Curvelet [8], Wavelet [9], Wave Atom [10], SIFT [1], Gabor [11], LBP [12], and AlexNet [7]. The key contribution is the application of mode voting method for accurate identification of a person at the fusion decision level. The experiments on IITD and CASIA databases have presented the efficiency of the proposed approach.

This paper is designed as follows. Section 2 describes the related works of palmprint recognition are obtainable and examined. In Section 3, we describe the methodology of the proposed approach. Section 4 explains the experiments and results of the suggested method. Section 5 assesses the results and feeds the conclusion.

II. RELATED WORK

Palmprint is emerging as alternative hand-based biometrics with user friendliness, flexibility in adapting the environment, and power of discrimination. The uniqueness and stability of palmprints make them a powerful source for ensuring sound criminal identification and access control. Research has reported progress in overcoming the limitations of wavelet analysis and synthesis by the application of composite dual-tree complex transforms coupled with Fourier transform for the purpose of extracting texture features for Support Vector Machine (SVM) detector.

Han [13] calculated seven detailed lines shapes from palmprint with three fingers using the low frequency data obtained from wavelets. This new feature vector is reduced its dimensionality using PCA. Optimum positive Boolean

function and Global learning vector quantization are used to construct the final decision.

Mu, Ruan and Shen [14] claimed a new approach to palmprint representation, which encompasses differentiation of a palmprint into separate areas of equal size. Discriminative local binary patterns statistic (DLBPS) are then utilized to identify the palmprint texture characteristics by the means of examining the distribution of patterns.

Zeng and Huang [15] solve the problem of linked features which generated because of using PalmCode for distinctive palms. In order to remove this correlation, they used PalmCodes and phase data that obtained from Gabor filters. The phase data is merged to obtain the Fusion Code using a fusion rule.

Conversely, hybrid approaches take the global and local features into consideration, which is arguably to be potentially the best approach. The author in [16] adopted such an approach to detect palmprint features from an image through the combined efforts of Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT). The use of Euclidean distance as a matching metric resulted in improved recognition outcomes as opposed to those produced by a separate implementation of DCT or DWT.

The author in [17] applied Hough transform for an extraction of distinct fingerprint features. This study will consider implementing Dempster-Shafer evidence theory and Bayesian fusion technique [18], majority rule [19], weighted majority algorithm [20], behavior-knowledge space method [21] and disjunction ("AND") vs conjunction ("OR") models [22] for fusion at the decision level.

The texture features extracted using Gabor filters have often performed well for recognition tasks including iris, face, and fingerprint. In the case of palmprint recognition, it has been shown to outperform line based and appearance-based approaches. Several techniques have been proposed for palmprint identification based on binary encoding of quantized Gabor features, including the use of subspace methods to reduce dimension. These approaches have gained popularity due to its efficient and compact representations, which are more suitable for online applications as in [23], [24], and [25].

One study of palmprint recognition [9] established the productivity of wavelet transforms in generating a successful 90% mid-level reference. Another study, Misar and Gharpure [26] described the features of palmprint images post extraction through the utilization of wavelet coefficients.

This study will involve an exception of the target palm area from the palm image based on the palm geometry, whereby the palm will need to be present in any direction opposite to the camera.

Wave Atom has advantage of compression above the other transforms [27]. In [28], the Wave Atom and the bidirectional 2D principal component analysis (B2DPCA) are applied to the cropped image to decrease the feature vectors dimension and they use the Extreme Learning Machine (ELM) as a classifier.

In [8], a comparison between Curvelet transform, Gabor filter, discrete cosine-transform and wavelet is performed which retrieves capabilities of Curvelet transform superior than all of other transforms in this study. For palmprint recognition, authors in [29] firstly implemented digital curvelet transform and the recognition rate of the experiment was up to 95.25%. In [30], the Support Vector Machine (SVM) as a classifier for the curvelet decomposed features of palm-print, the recognition accuracy became 98.5%. In [31], the second frequency band of curvelet coefficients is used to represent the palmprint image. The recognition accuracy of the experiments was up to 99.9%.

Furthermore, researchers Chen and Moon [32] effectively extracted palmprint features using Scale Invariant Feature Transformation (SIFT) descriptions and fusion based on Symbolic Aggregate Approximation (SAX). Zhao, Bu and Wu [33] fused competitive coding and SIFT to enhance palmprint verification.

In their turn in [34], they proposed a new verification system on a basis of palmprint and hand shape fusion screened through SIFT. Such an adjustment improved the effectiveness of SIFT in extracting the features that are invariant to scaling and image rotation across various applications such as object identification and video tracking. The experiments revealed promising matching score findings in the aspect of fusing palmprint and handshape features using the IIT Delhi Touchless Palmprint Database [35].

Reference [36] applied the sparse representation of SIFT to implement a touchless method for palmprint identification by extracting the left and right palms' print features. The SVM probability distribution detector was used to produce the rank level fusion in finalizing a personal identification.

Several studies yielded competitive palmprint identification findings on the bases of R_Egim Sfax Tunisia (REST) hand database [36] and CASIA Palmprint Database [1]. In particular, they developed a bimodal identification approach using SIFT descriptors for obtaining hand shape and palmprint features. The researchers applied a local sparse representation technique to examine images with high discrimination. Additionally, they implemented a cascade fusion at decision and feature levels to reach a notable 99.57% rate of identification, which is among the best related outcomes reported in research literature.

Local Binary Pattern (LBP) is one of the simple techniques to extract an identification feature in use across different computer applications [37]. Researchers in [38] utilized boosted LBP for purposes of palmprint identification. Under this method, scalable sub-sections in the LBP based histograms serve to depict the scanned palmprint features. However, the resulting texture is distorted and involves vague multidirectional ridges and lines. Later in [39], LBP shaped the basis for feature extraction in an enhanced identification approach grounded in the directional shifts of gradient operator. Finally, the above method was further improved in the aspects of speed and precision (Promila and Laxmi [12]).

Deep learning is a well-known machine learning subsegment, which addresses algorithms in the field of

artificial neural networks modeling the neurobiological behavior of the human brain.

The author in [40] utilized CNN to classify handwriting digits using the method of backward error propagation. They accomplished notable progress as indicated by 150,000 testing images, 1.2 million training, and 50,000 validations. In the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) contest from 2015, authors in [2] were able to train a comprehensive CNN capable of categorizing nearly 1.2 million images in high resolution under one thousand distinct categories. This expert achieved impressive 17% and 37.5% scores on the top-five and top-one fault rates in the test phase effectively surpassing the earlier records. Correspondingly, his AlexNet encompasses five convolutional layers, sixty million factors, and 650,000 neurons. Three of the above are full-scale layers with no less than 1000-way SoftMax, whereas the other two are supported by ordinary max pooling layers. Some important suggestions were made to increase the training speed by the means of improving the convolution process with regard to graphics processing unit (GPU) and non-saturate neurons. Moreover, the problem of overfitting in the three connected layers was addressed by a very progressive regularization method of "dropout." As a result, the modified version of AlexNet produced the 15.3% winning score on the top-five fault rate as compared to 26.2% from its closest competitor in the ILSVRC-2012.

Another massive-scale classification of images using a comprehensive 19-weight-layers convolutional network was performed by Simonyan and Zisserman [3]. The experiment indicated the usefulness of illustration depth in enhancing the accuracy of categorization. Overall, the utilized models proved to have high generalizability across different sets of data and tasks, whereby they were capable of outperforming some more sophisticated recognition systems. Furthermore, a substantial deep convolutional architecture with 22 layers entitled GoogleNet [4] was designed in 2015. Its chief distinctive characteristic is the improved process of inside-network computing resources utilization. GoogleNet was also expanded in width and depth at no significant extra costs from the computational budget. The quality optimization was achieved through multi-scale processing intuitions and the Hebbian principle.

In this regard, the steady increase in network depth prompted the development of deep residual learning methodology [5] for training facilitation. In contrast to approaching layers as targeting non-referenced functions, under this framework, they are re-conceptualized to learn residual functions with an emphasis on the input domains. The resulting residual architectures are both deeper and less complex contributing to the increased accuracy and ease of optimization. In the Dense Convolutional Network or DenseNet designed by authors [6] in 2017, all layers are cross-connected in a feed forward pathway. Some significant advantages of such an architecture include a marked reduction in the parameter numbers, enhanced feature propagation, mitigation of the vanishing gradient issue, and incentive for repeat feature use. To this end, a small ImageNet-based CNN entitled Squeezenet [7] proved to require fifty times fewer parameters to accomplish the same level of accuracy than

AlexNet. With the use of compression techniques, it can be reduced to up to 510 size of AlexNet, which is equal to 0.5 MB.

Deep learning has been effectively utilized across diverse biometric domains as a breakthrough method in computer-based image processing. Highly satisfactory outcomes have been particularly linked to palmprint recognition. The author in [41] used a three-step process in implementing the above technique for palmprints within a deep confidence architecture. They first developed top-to-down training with no supervision to instruct the selected samples. The researchers further identified optimum parameters to amend the system for an improved performance. Finally, they employed deep learning models to examine the test samples. In the outcome, the deep learning approach proved to be associated with advanced recognition scores for palmprints as

compared with the traditional techniques, including LBP and principal component analysis (PCA).

Minaee and Wang [42] designed a convolutional scattering transform/network for purposes of palmprint recognition through multi-layer representations. The network runs on default wavelet transforms. Its initial layer targets the relevant features for processing through SIFT description, whereas the higher layers extract contents of increased frequency inaccessible for descriptors. PCA contributes to the process of recognition by adjusting its computational complexity. It is specifically used to reduce scattering feature dimensionality following an extraction. Finally, recognition is ensured by two distinct classifiers, including a minimum distance classifier and a multi-class SVM. The described procedure yielded 99.95% to 100% accurate recognitions upon its testing at a recognized palmprint database.

TABLE I. COMPARISON BETWEEN PALMPRINT BIOMETRIC SYSTEMS IN THE RELATED WORK [46]

Ref.	Database	Features Classification	Features Extraction	RR
[13]	Grabbed from a CCD camera	GLVQ approach	Wavelet	FRR = 1.6%, FAR = 36.3%
[29]	PolyU	Euclidian distance classifier	Digital Curvelet Transform	95.25
[47]	PolyU Palmprint Database	SVM	Dual-tree complex wavelet transforms	97
[38]	UST-HK palmprint database	Chi square distances	Boosting Local Binary Pattern	Equal Error Rate = 2%
[32]	PolyU	Symbolic Aggregate approximation	SIFT	Equal Error Rate = 0.37%
[39]	palm print tracking in dynamic environment	Chi square and PNN	Sobel and LBP	PNN: EER=0.74% Chi: EER=1.52%
[30]	PolyU	SVM	Digital Curvelet Transform	98.5
[14]	Fujitsu fi-60F high speed flatbed scanner is used	Nearest neighbor (NN) classifier based on the Euclidean distance	Discriminative local binary patterns statistic (DLBPS)	98
[15]	PolyU Palmprint Database	Euclidean distance and the nearest neighbor classifier	Gabor features	100
[9]	Hong Kong Polytechnic University 2D_3D Database	Mean square error	Wavelets	93
[12]	PolyU	Chi-square test and Pearson correlation test	LBP	99.22
[33]	IITD	Competitive code algorithm	SIFT	Equal Error Rate = 0.49%
[16]	IITD, PolyU	Euclidean Distance	DWT - DCT	94.44, 95.65
[34]	IITD	Matching Score	SIFT	Palmprint = 94.05 Hand shape + Palmprint = 97.82
[24]	CASIA IITD	Euclidean Distance	2D Gabor filter	90.76 91.4
[41]	Beijing Jiao Tong University	DBN	DBN	90.63
[26]	IITD	Neural Network	Discrete Wavelet Transform	75.6
[31]	PolyU	Nearest neighbor method	Digital Curvelet Transform	99.9
[36]	IITD	Matching score	SIFT and Gabor	Palmprint = 91.08 Hand shape + Fingers + Palmprint = 98.04
[1]	IITD Bosphorus	SVM	SIFT sparse representation	IITD: Palmprint = 96.73 Hand shape + Palmprint = 99.57 Bosphorus: Palmprint = 94.95 Hand shape + Palmprint = 97.61
[42]	PolyU	Minimum distance SVM	Scattering Features + PCA	99.95 100
[43]	PolyU	Softmax	CNN-F architecture	100
[45]	IITD 11k	SVM	CNN-features + LBP	IITD: CNN Fea. = 90 CNN Fea. + LBP = 94.8 11k: CNN Fea. = 94.8 CNN Fea. + LBP = 96

A deep CNN was also employed in a project implemented by authors in [43] to extract palmprint features. Because of its capacity to combine features across all levels, the CNN has been marked for its outstanding performance with respect the processing of images, speech, and video. The researchers specifically applied the CNN-F model to identify and verify convolutional features across different architecture layers. The resulting findings on the basis of PolyU palmprint database yielded 0.25% and 100% with respect to verification accuracy and identification score respectively evidencing the reliability and effectiveness of CNN in the aspect of palmprint recognition. Another block approach entitled “Squeeze-and-Excitation” (SE) was introduced in [44] to provide explicit inter-channel modeling for purposes of improving the flexibility of channeled feature reactions. The researchers found that Squeeze-and-Excitation Networks (SENet) developed through a combination of blocks had high generalizability across a wide variety of sets of channeling data. Finally, the extensive dataset of hand images containing empirical evidence for biometric and gender identification suggested by Afifi [45] contributed to the sound CNN training in the aspect of biometric identification. The trained system proved effective in extracting features to yield a range of SVM classifiers.

Table I summarizes the unimodal system palmprint methods, proposed in the literature.

III. METHODOLOGY

Two systems are used in this paper; the block diagram of the first system is shown in Fig. 1. Whereas the second system shown in Fig. 2. The segmented palmprint (ROI) palmprint of CASIA [48] and IITD databases [35] are used for the two systems. IITD database consists of 230 subjects, about 10 images for each, left and right with size 150x150 pixels in grayscale. CASIA database contains 240 subjects, around 10 images for each subject classified left and right with size 192x192 pixels. Images are resized to 227x227 pixel and converted to RGB to extract features for CNN.

A. Features Extraction

In the initial setting, the project uses a variety of feature representations extracted from the same person using such texture-based techniques as Curvelet, Wavelet, Wave Atom, SIFT, Gabor, LBP, and AlexNet. Such an undertaking shapes a new methodology for improving the matching process accuracy in palmprint recognition.

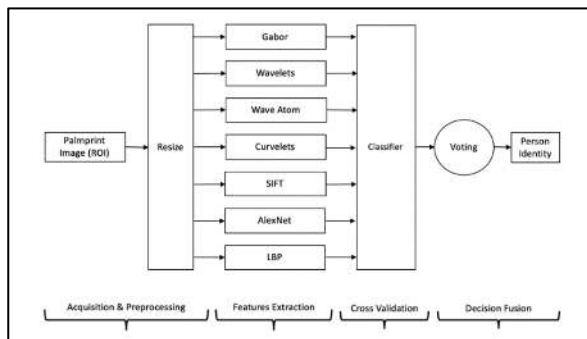


Fig. 1. System 1: Human Identification using Texture based Descriptors for Palmprint Images (Elgallad et al., 2017).

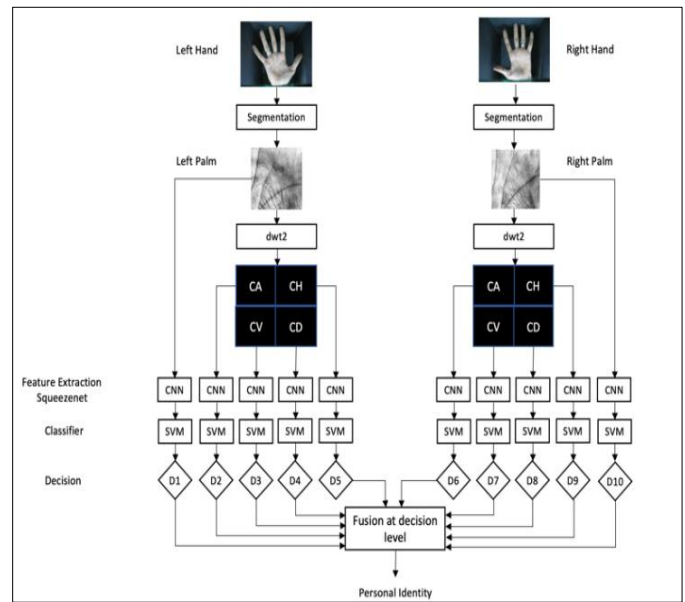


Fig. 2. System 2: Dense Hand-CNN: CNN Architecture based on Later Fusion of Neural and Wavelet Features.

Wave-packet transforms such as Gabor function, wavelets, wave atoms and curvelet are used in this system. 2D wave packets is denoted as $\varphi_{\mu}(\chi_1, \chi_2)$. The main parameters for wave packet architectures which are used as indicator are α and β . If ($\alpha = 1$), this indicates that multiscale decomposition is used. While ($\alpha = 0$) indicates that it is not. Base elements may be either local or weakly directional. β parameter is used to index the base element. When ($\beta = 1$), it is extended and entirely directional when ($\beta = 0$) [8]. When ($\alpha = 1, \beta = 1/2$), this match Curvelet. Wavelets match ($\alpha = \beta = 1$), and the Gabor transform is match to ($\alpha = \beta = 0$). Wave atoms are described as the point $\alpha = \beta = 1/2$.

Using (1) to extract Gabor features from the resized image (64x64 pixels):

$$G(x, y, \theta, u, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{x^2 + y^2}{2\sigma^2}\right\} \quad (1)$$

The features vector is constructed by merging the mean squared energy and mean amplitude matrices. Two factors are examined to get the best possible features, wavelet scales' number and filter orientations' number.

Haar wavelet filter is used to perform Single-level wavelet transform. The resized input image (64x64 pixels) is used to obtain the approximation coefficients matrix using (2a, 2b):

$$w(a, b_1, b_2) = \frac{1}{a} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} I(x, y) \psi\left(\frac{x-b_1}{a}, \frac{y-b_2}{a}\right) dx dy \quad (2a)$$

$$\psi_H(x, y) = \begin{cases} 1, 0 \leq x < \frac{1}{2}, 0 \leq y < 1; \\ -1, \frac{1}{2} \leq x < 1, -1 < y \leq 0; \\ 0, \text{else} \end{cases} \quad (2b)$$

Partitioning of the frequency in parabolic scaling with a single defined oscillation direction and real-valued frame forms the basis of the forward 2D wave atom transform extended by mirror. As described in (3), the wave atom transform is utilized to extract a cell array comprising the related coefficients from the input image with adjusted size of 64x64.

$$|\varphi_\mu(\omega)| \leq C_M \cdot 2^{-j} (1 + 2^{-j} |\omega - \omega_\mu|)^{-M} + C_M \cdot 2^{-j} (1 + 2^{-j} |\omega + \omega_\mu|)^{-M}$$

and

$$|\varphi_\mu(x)| \leq C_M \cdot 2^{-j} (1 + 2^{-j} |x - x_\mu|)^{-M} \text{ for all } M > 0,$$

where

$$x_\mu = 2^{-j} n \quad \omega_\mu = \pi 2^{-j} m \quad C_1 2^j \leq \max_{i=1,2} |m_i| \leq C_2 2^j \quad (3)$$

where C1 and C2 > 0, and will be indirect by the details of the execution.

In DWT, multi-stage filter banks with high-pass (HP) and lowpass (LP) filters are used to perform a series of dilations. detail coefficients are obtained after the HP filters while the approximate coefficients are obtained after the LP filter [26]. Furthermore, the two-dimensional setting involves three distinct classes of detail coefficients situated across diagonal, horizontal, and vertical pathways. To this end, the detail coefficients are represented by the respective subbases LH_j; HL_j, and HH_j; j = 1; 2;...; J, whereby the most coarse or the largest decomposition scale is denoted as J and j identifies the scale. The multilevel wavelet decomposition is depicted at the third level under Fig. 3.

To extract the curvelet coefficients, Fast Discrete Curvelet Transform via wedge wrap is used for 128x128 input images as in (4a,4b):

$$c_{j,\ell,b} = \int \hat{f}(\omega) U_{j,\theta_\ell}(\omega) e^{ib \cdot \omega} d\omega, \quad (4a)$$

where U_{j,θ_ℓ} is a real wedge frame value that expanded to measure j and by the shearing process, parabolic restricted to angles close to θ_ℓ . The value of b in (4a) should be evaluated to discrete curvelet factors which are still a fitted structure:

$$f(x) = \sum_{i,\ell,k} \langle f, \varphi_{j,\ell,k} \rangle \varphi_{j,\ell,k}(x), \quad (\text{conv. in } L^2). \quad (4b)$$

Constant local feature arguments are extracted using the Scale Invariant Feature Transform [1]. To select key locations in scale space, local smallest and highest values of a variance of Gaussian function are used by Comparing each pixel to its neighbors as in (5), (7):

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (5)$$

where the scale space is L(x, y, σ), I(x, y) is the input image, and G(x, y, σ) is defined as:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (6)$$

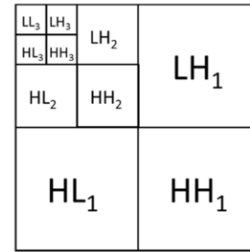


Fig. 3. Level 3 Wavelet Decomposition [49].

which is a variable scale function, and the Gaussian difference scale space is defined as:

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad (7)$$

When extreme points locations are detected, the key points that are invariant to affine transformations and unaffected to noise must be used.

As in (8), (9), to compute the direction, a neighborhood is determined around the key point to find its descriptor using gradient magnitude $m(x,y)$ and the scale. $\theta(x,y)$ is the orientation of the key point.

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (8)$$

$$\theta(x, y) = \tan^{-1} \left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \right) \quad (9)$$

AlexNet is utilized as a CNN for purposes of extracting learned image features. The corresponding architecture encompasses a combined activity of two layer types, including the three connected layers and the initial five convolutional layers. As seen in Fig. 4, a 1000-way softmax delivers distributions beyond 1000 class labels serving as the output of the mentioned connected layers [2].

The palmprint identification perspective involves discriminative LBP features [38]. Upon the detection of the dominant pixel in the representation, the pattern code is matched against neighbors to produce the needed calculation as per (10).

$$LBP_{P,R} = \sum_{p=1}^P s(g_p - g_c) 2^{p-1} \quad (10)$$

where

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

The second setting employs such feature extraction instruments as SqueezeNet and discrete wavelet transform. The former is an architecture with default training based on the ImageNet database sample comprising over one million of images. The model can thus categorize images across 1000 respective object categories. There are three key techniques used in SqueezeNet to design CNN systems [7]. Firstly, 3x3 filters need to be replaced by 1x1 filters, as the latter have nine times fewer parameters. Secondly, squeeze layers should be applied to achieve a drop in the number of input channels.

Thirdly, the sample should be downed late in the system to ensure significant maps of activation for convolution layers.

B. Cross Validation

Originally introduced by Vapnik [50], the SVM is utilized as a cross-validation classifier in the mentioned settings. SVM belongs to the class of Maximum Margin Classifiers (MMC) and is linked to Structural Risk Minimization (SRM). It serves as an input vector to the space of the upper dimension featuring the top separating hyperplane. The project particularly involves a multi-class linear SVM to account for the 230 of its subjects or classes [1].

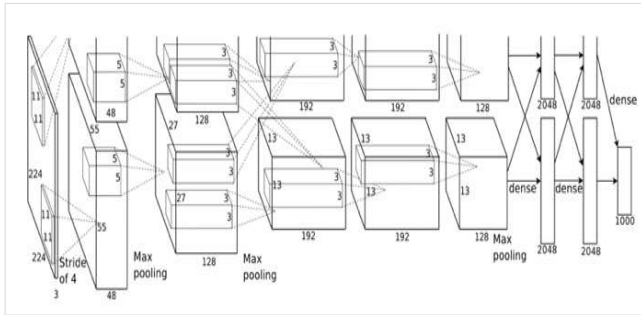


Fig. 4. AlexNet Convolutional Neural Network Architecture [2].

The training data set and its labels is $(x_n, y_n), n=1, \dots, N, x_n \in \mathbb{R}^p, t_n \in \{-1, +1\}$, SVMs learning includes the subsequent controlled optimization:

$$\begin{aligned} \min_{w, \xi_n} &= \frac{1}{2} w^T w + C \sum_{n=1}^N \xi_n \\ \text{s. t. } & w^T x_n t_n \geq 1 - \xi_n \quad \forall_n \\ & \xi_n \geq 0 \quad \forall_n \end{aligned} \quad (11)$$

where ξ_n are the slack variables, w is the vector of coefficients, and C is the capacity constant.

The unconstrained optimization problem in (11) that is recognized as the primal form problem of L1-SVM:

$$\min_w = \frac{1}{2} w^T w + C \sum_{n=1}^N \max(1 - w^T x_n t_n, 0) \quad (12)$$

Meanwhile L1-SVM is not differentiable, the L2-SVM is used to minimize the squared hinge loss as in (12):

$$\min_w = \frac{1}{2} w^T w + C \sum_{n=1}^N \max(1 - w^T x_n t_n, 0)^2 \quad (13)$$

The class label of a test data x is:

$$\text{arg}_t \max(w^T x) t \quad (14)$$

Multiclass SVM uses one-vs-rest approach to represent the output of the k -th SVM.

$$a_k(x) = w^T x \quad (15)$$

the forecast class is

$$\text{arg}_k \max a_k(x) \quad (16)$$

C. Score Fusion

Among the greatest information fusion system challenges is the problem of determining the needed type of data for consolidation under the fusion module. Multiple fusion

strategies are available across all four levels. The discussed settings in [51] involve the match score level fusion as representative of the principal fusion level in biometrics.

The match score identifies similarities between the default biometric feature and the input vectors. In its turn, the match score level fusion is accomplished to make an outcome recognition decision upon the consolidation of output match scores based on the relevant biometric matches [51]. There were many approaches that are used in fusion at decision level such as: Majority Voting and Weighted Majority Voting. In our proposed systems, we introduced a novel approach in decision level technique, the Mode Voting Technique (MVT).

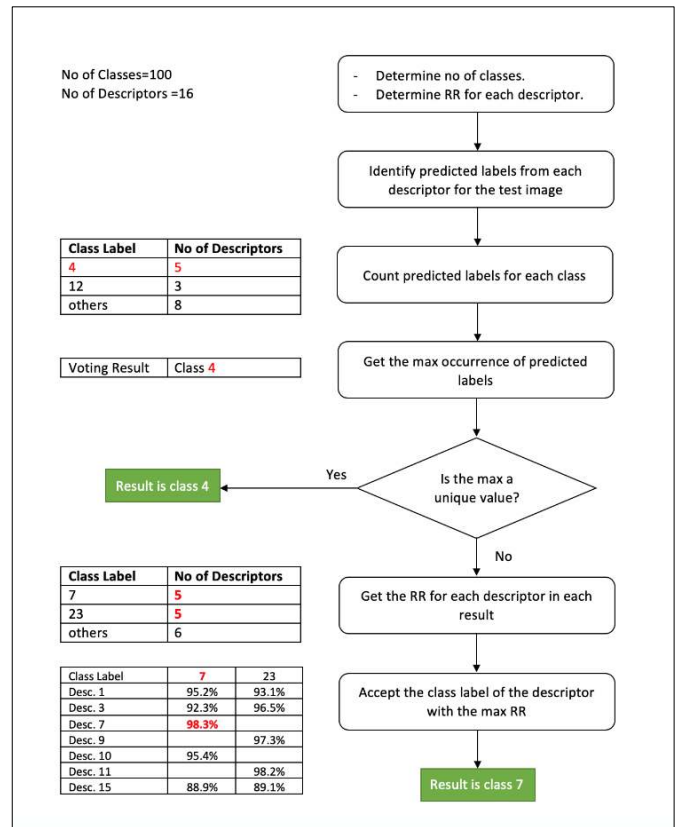


Fig. 5. Flowchart of Mode Voting Technique (MVT).

D. Mode Voting Technique

Mode Voting Technique (MVT) is a novel voting technique that is consolidating information at the decision level. This method utilizes the standard class label values that are retrieved from the predicted label array obtained through the SVM discriminate classifier.

Fig. 5 explains the flowchart of MVT with illustrated example.

The mode voting technique uses the most common class label values obtained from the predicted label array that was extracted from SVM classifier. In order to identify human, the mode voting technique is implemented to find the most frequent non-repeated scores in the predicted label array X .

$$Z = \text{mode}(X_{k,i}) \quad (17)$$

where Z is the class label of the test image, k is the index of the test image, and i is the index of the descriptor.

IV. EXPERIMENTAL RESULTS

In the proposed systems, two sets of databases are which are CASIA Palmprint Database and Delhi Touchless Palmprint Database IIT version 1.0. CASIA Palmprint database contains 2400 palmprint images for 240 subjects from left and right palms in size 192x192 pixels for the segmented palmprint ROI. The images are 8 bit gray-level JPEG files. IITD database basically contains hand images saved in format of bitmap and contains both of left-right hands images for 230 persons in size 150x150 pixels. The age varied between 14 and 56 years old. The segmented and normalized of palmprint regions are available. For CNN feature extraction, the images are improved to RGB and resized to 227x227 pixels. In both systems, SVM classifier is used.

In the first system, the Gabor features are 120. The recognition rate is 70.29% with IITD, and 87.46% with CASIA.

The Wavelets features are 1024. The recognition rate is 80.94% with IITD, and 92.46% with CASIA.

In wave atom, the output array is obtained and organized to get 1120 elements. The recognition rate is 77.32% with IITD, and 92.39% with CASIA.

In Fast Discrete Curvelet Transform, features are 121. The recognition rate is 79.57% with IITD, and 91.16% with CASIA.

In Scale Invariant Feature Transform, 1024 features are extracted from 150x150 images. The recognition rate is 96.96% with IITD, and 98.91% with CASIA.

AlexNet convolutional neural network is used to extract 4096 features from the last fully-connected layer. The recognition rate is 93.91% with IITD, and 98.55% with CASIA.

LBP output the histogram of 64x64 input images. The output array has 256 features. The recognition rate is 66.81% with IITD, and 82.54% with CASIA.

Table II illustrates the results of the used seven descriptors for both databases. The fusion at decision level using mode voting technique, which depends on each descriptor's predicted label array, achieved a recognition rate equals to 99.57% with IITD database, with processing time for each image 1.88 sec. With CASIA database, the recognition rate is 100%, with 2.36 sec processing time for each image.

Fig. 6 shows CMC Curves of system1, probability of detecting the correct identity within the top K ranks for the descriptors.

The main objective of the second system is to reduce the processing time for each image and maintain the high recognition rate obtained from the first system.

TABLE II. RESULTS OF THE FIRST SYSTEM

Descriptor	Features	IITD	CASIA
Gabor	120	70.29	87.46
Wavelets	1024	80.94	92.46
WaveAtoms	1120	77.32	92.39
Curvelet	121	79.57	91.16
SIFT	1024	96.96	98.91
AlexNet	4096	93.91	98.55
LBP	256	66.81	82.54
Voting		99.57	100
Time / Sec		1.88	2.36

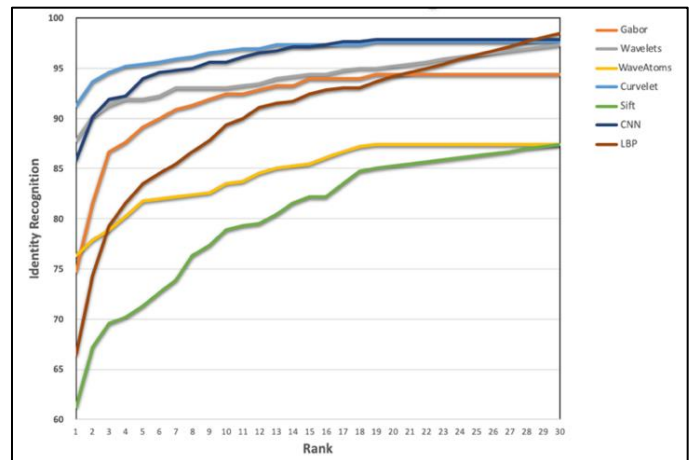


Fig. 6. CMC Curves: Probability of Detecting the Correct Identity within the top K Ranks.

SqueezeNet has 50x fewer parameters compared with AlexNet. It has a model compression technique which can compress SqueezeNet to less than 0.5MB (510 smaller than AlexNet). Due to these advantages, it's used as a feature extractor in the second system.

The left and right palm for each subject is used. The features are obtained from Squeezenet for the palm images, and the single-level 2-D discrete wavelet transform (DWT) of the images using the Haar wavelet filter. The DWT yields the approximation coefficients matrix cA with the detail coefficients matrices cH (horizontal), cV (vertical), and cD (diagonal).

The predicted label arrays obtained from SVM as a classifier for the obtained features are fused using mode voting technique. The recognition rate equals to 100% with IITD database, with processing time for each image 0.74 sec. With CASIA database, the recognition rate is 99.6%, with 0.67 sec processing time for each image.

Table III illustrates the detailed result of the second system.

TABLE III. RESULTS OF THE SECOND SYSTEM

Descriptor	IITD				CASIA			
	L		R		L		R	
	RR	t/s	RR	t/s	RR	t/s	RR	t/s
SqueezeNet	96.96	0.06	97.61	0.07	97.92	0.09	96.25	0.07
CA	88.70	0.06	89.57	0.07	92.92	0.07	90.42	0.06
CH	87.83	0.07	90.00	0.07	64.79	0.06	46.46	0.06
CV	93.91	0.07	94.13	0.07	87.08	0.06	82.71	0.06
CD	62.46	0.09	53.91	0.09	5.00	0.07	5.42	0.07
Voting %	100				99.6			
Time/sec.	0.74				0.67			

V. DISCUSSION

The results show that the fusion at decision level promises an outstanding recognition rate regardless of low recognition rate of some descriptors and filters. The mode voting technique positions top of the list of SVM classifiers used for each descriptor.

Table IV displays the performance of the proposed palmprint recognition systems vs. current systems in the related work that use IITD and CASIA databases.

A significant issue in this comparison table is the processing time that is explained in Table V. in System1, the processing time for an image of IITD database is 1.88s, while in CASIA is 2.36s. In the second system, the process time is 0.74s in IITD, and 0.67 in CASIA database.

TABLE IV. PERFORMANCE COMPARISON OF PROPOSED PALMPRINT RECOGNITION SYSTEM VS. EXISTING SYSTEMS USING IITD AND CASIA DATABASES

	System1		System2	
Palmprint Database	IITD	CASIA	IITD	CASIA
Hand	Left	Left	Left & Right	Left & Right
No of Subjects	230	230	230	240
Feature extraction	7 descriptors		SqueezeNet & dwt	
Classifier	SVM		SVM	
Result by Fusion	99.57	100	100	99.6
Time/image in sec.	1.88	2.36	0.74	0.67

TABLE V. PROCESSING TIME COMPARISON

Ref.	Features Extractor	Features Classifier	Database	RR(%)
[33]	SIFT	competitive code algorithm	IITD	Equal Error Rate = 0.49%
[16]	DWT - DCT	Euclidean Distance	IITD, PolyU	94.44, 95.65
[34]	SIFT	Matching Score	IITD	Palmprint = 94.05 Hand shape + Palmprint = 97.82
[24]	2D Gabor filter	Euclidean Distance	CASIA IIT Delhi	90.76 91.4
[26]	Discrete Wavelet Transform	Neural Network	IITD	75.6
[36]	SIFT and Gabor	Matching score	IITD	Palmprint = 91.08 Hand shape + Fingers + Palmprint = 98.04
[1]	SIFT sparse representation	SVM	IITD Bosphorus	IITD: Palmprint = 96.73 Hand shape + Palmprint = 99.57 Bosphorus: Palmprint = 94.95 Hand shape + Palmprint = 97.61
[45]	CNN-features+LBP	SVM	IITD 11k	IITD: CNN Fea. = 90 CNN Fea. + LBP = 94.8 11k: CNN Fea. = 94.8 CNN Fea. + LBP = 96
Proposed system 1	7 Descriptors	SVM	IITD CAISA	99.57 100
Proposed system 2	SqueezeNet+dwt	SVM	IITD CAISA	100 99.6

VII. CONCLUSION

This work statements two palmprint recognition systems depending on the mode voting technique, and compares the performance of the systems for image processing time. The novelty comes from using mode voting technique at decision level. Our experimental results demonstrate the effectiveness of the suggested systems.

The selection of SqueezeNet and DWT in the second system depends on the result of the first system. As Alexnet achieved high recognition rate, we looking forward to deep learning especially for SqueezeNet due to its advantage. Also, for DWT, the four filters are used, compared with one filter in the first system, to increase the number of predicted label arrays that are needed for mode voting technique.

For future work, MVT can be tested on other experiments in [52:56] to enhance the result and ensure the quality of MVT.

REFERENCES

- [1] N. Charfi, H. Trichili, A.M. Alimi and B. Solaiman, "Bimodal biometric system for hand shape and palmprint recognition based on SIFT sparse representation". *Multimedia Tools and Applications*, 2016, pp.1-26.
- [2] A. Krizhevsky, I. Sutskever and G.E. Hinton, "ImageNet classification with deep convolutional neural networks", In *Advances in neural information processing systems*, 2012, pp. 1097-1105.
- [3] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition", 2014, arXiv preprint arXiv:1409.1556.
- [4] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich, "Going deeper with convolutions", In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1-9.
- [5] 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.
- [6] G. Huang, Z. Liu, L. Van Der Maaten and K.Q. Weinberger, "Densely Connected Convolutional Networks", In *CVPR*, 2017, Vol. 1, No. 2, p. 3.
- [7] F.N. Iandola, S. Han, M.W. Moskewicz, K. Ashraf, W.J. Dally and K. Keutzer, "Squeezenet: Alexnet-level accuracy with 50x fewer parameters and < 0.5 mb model size", arXiv preprint arXiv:1602.07360, 2016.
- [8] I.J. Sumana, M.M. Islam, D. Zhang and G. Lu, "Content based image retrieval using curvelet transform", In *Proceedings of 2008 International Workshop on Multimedia Signal Processing*, pp. 11-16.
- [9] H. Kekre, B. Tanuja, K. Sarode and A.A. Tirodkar, "A study of the efficacy of using Wavelet Transforms for Palm Print Recognition", *International Conference on Computing, Communication and Applications*, Dindigul, Tamilnadu, 2012, pp. 1-6.
- [10] Demanet and Ying, "Wave atoms and sparsity of oscillatory patterns", Elsevier, 2007, pp. 368-387.
- [11] W. Ouarda, H. Trichili, A.M. Alimi and B. Solaiman, "Face recognition based on geometric features using Support Vector Machines", *Soft Computing and Pattern Recognition (SoCPar)*, 6th International Conference of, Tunis, 2014, pp. 89-95.
- [12] Promila and V. Laxmi, "Palmprint Matching Using LBP", *International Conference on Computing Sciences*, Phagwara, 2012, pp. 110-115. doi: 10.1109/ICCS.2012.55.
- [13] C.C. Han, "A hand-based personal authentication using a coarse-to-fine strategy", *Image and Vision Computing*, 2004, 22 (11) 909-918.
- [14] M. Mu, Q. Ruan and Y. Shen, "Palmprint Recognition Based on Discriminative Local Binary Patterns Statistic Feature", *International Conference on Signal Acquisition and Processing*, Bangalore, 2010, pp. 193-197.
- [15] Z. Zeng and P. Huang, "Palmprint recognition using Gabor feature-based two-directional two-dimensional linear discriminant analysis", *International Conference on Electronic & Mechanical Engineering and Information Technology*, Harbin, Heilongjiang, China, 2011, pp.191-1921.
- [16] V. Varshney, R. Gupta and P. Singh, "Hybrid DWT-DCT based method for palm-print recognition", *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, Noida, 2014, pp. 000007-000012.
- [17] A.K. Jain, S. Prabhakar and S. Chen, "Combining Multiple Matchers for a High Security Fingerprint Verification System", *Pattern Recognition Letters*, 1999, 20(11-13): 1371-1379.
- [18] L. Xu, A. Krzyzak and C.Y. Suen, "Methods for Combining Multiple Classifiers and their Applications to Handwriting Recognition. *IEEE Transactions on Systems, Man, and Cybernetics*, 1992, 22(3): pp.418-435.
- [19] L. Lam and C.Y. Suen, "Application of Majority Voting to Pattern Recognition: An Analysis of its Behavior and Performance", *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, 1997, 27(5):553-568.
- [20] L.I. Kuncheva "Combining Pattern Classifiers-Methods and Algorithms", Wiley, 2004.
- [21] Y.S. Huang and C.Y. Suen, "Method of Combining Multiple Experts for the Recognition of Unconstrained Handwritten Numerals", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1995, 17(1):90-94.
- [22] J. Daugman, "Combining Multiple Biometrics", 2000, Available: <http://www.cl.cam.ac.uk/users/jgd1000/combine/combine.html>. [Accessed 15 May 2017].
- [23] I. Fogel and D. Sagi, "Gabor filters as texture discriminator". *Biol. Cybern.* 1989, 61, 103-113.
- [24] G. Jaswal, R. Nath and A. Kaul, "Texture based palm Print recognition using 2-D Gabor filter and sub space approaches", *International Conference on Signal Processing, Computing and Control (ISPC)*, Wagnaghat, 2015, pp. 344-349.
- [25] A. Teuner, O. Pichler and B.J. Hosticka, "Unsupervised texture segmentation of images using tuned matched Gabor filters", *IEEE Trans. Image Process*, 1995, 4 (6), 863-870.
- [26] M. Misar and D. Gharpure, "Extraction of feature vector based on wavelet coefficients for a palm print based biometric identification system", *2nd International Symposium on Physics and Technology of Sensors (ISPTS)*, Pune, 2015, pp. 113-119.
- [27] Z. Haddad, A. Beghdadi, A. Serir and A. Mokraoui, "A new fingerprint image compression based on wave atoms transform". *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, Ajman, 2009, pp. 89-94.
- [28] A.A. Mohammed, Q.M. Jonathan Wu and M.A. Sid-Ahmed, "Application of Wave Atoms Decomposition and Extreme Learning Machine for Fingerprint Classification", In: A. Campilho and M. Kamel (eds) *Image Analysis and Recognition. ICIAR 2010. Lecture Notes in Computer Science*, vol 6112. Springer, Berlin, Heidelberg, 2010.
- [29] K. Dong, G. Feng and D. Hu, "Digital curvelet transform for palmprint recognition, *Advances in Biometric Person Authentication*", Springer Berlin Heidelberg, 2005, pp. 639-645.
- [30] X. Xu, D. Zhang, X. Zhang and Y. Cao, "Palmprint Recognition Based on Discrete Curvelet Transform and Support Vector Machine", *Journal of Infrared and Millimeter Waves*, 2009, vol. 28, no. 6, pp. 456-460.
- [31] F. Liu, L. Zhou, Z.M. Lu and T. Nie, "Palmprint Feature Extraction Based on Curvelet Transform", *Journal of Information Hiding and Multimedia Signal Processing*, 2015, 6(1), pp.131-139.
- [32] J. Chen and Y. Moon, "Using SIFT features in palmprint authentication", *19th International Conference on Pattern Recognition*, 2008, pp. 1-4.
- [33] Q. Zhao, W. Bu and X. Wu, "SIFT-based image alignment for contactless palmprint verification", in: *Proceedings of International Conference on Biometrics*, 2013, pp. 16.
- [34] N. Charfi, H. Trichili, A.M. Alimi and B. Solaiman, "Bimodal biometric system based on SIFT descriptors of hand images". *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, San Diego, CA, 2014, 2014, pp. 4141-4145.
- [35] IIT Delhi Touchless Palmprint Database version 1.0, http://web.iitd.ac.in/~ajaykr/Database_Palm.htm. [Accessed 10 October 2018].

- [36] N. Charfi, H. Trichili, A.M. Alimi and B. Solaiman, "Personal recognition system using hand modality based on local features". 11th International Conference on Information Assurance and Security (IAS), Marrakech, 2015, pp. 13-18.
- [37] T. Ojala, M. Pietikainen and D. Harwood, "A comparative study of texture measures with classification based on feature distribution", *Pattern Recognition*, 1996, 29:51-59.
- [38] X. Wang, H. Gong, H. Zhang and Z. Zhuang, "Palmpoint Identification using Boosting Local Binary Pattern. *Pattern Recognition*", 18th International Conference on Pattern Recognition (ICPR'06), Hong Kong, 2006, pp. 503-506.
- [39] G.K.O. Michael, T. Connie and A.B.J. Teoh, "Touch-less palm print biometrics: Novel design and implementation", *Image and Vision Computing*, 2008, 26(12), pp.1551-1560.
- [40] Y. LeCun, B. Boser and JS. Denker, "Backpropagation applied to handwritten zip code recognition", *Neural Comput* 1(4):541-551, 1989.
- [41] D. Zhao, X. Pan, X. Luo and X. Gao, "Palmpoint recognition based on deep learning", In: *International Conference on Wireless, Mobile and Multi-Media*, 2015, pp. 214-217.
- [42] S. Minaee and Y. Wang, "Palmpoint recognition using deep scattering convolutional network", arXiv preprint arXiv:1603.09027, 2016.
- [43] Q. Sun, J. Zhang, A. Yang and Q. Zhang, "Palmpoint recognition with deep convolutional features", In *Chinese Conference on Image and Graphics Technologies*, Springer, Singapore, 2017, pp. 12-19.
- [44] J. Hu, L. Shen and G. Sun, "Squeeze-and-excitation networks", arXiv preprint arXiv:1709.01507, 7, 2017.
- [45] M. Afifi, "11K Hands: Gender recognition and biometric identification using a large dataset of hand images." arXiv preprint arXiv:1711.04322, 2017.
- [46] E.A. Elgallad, N.A. Charfi, W. Ouarda and M. Alimi, "Human identity recognition using sparse auto encoder for texture information representation in palmpoint images based on voting technique". *Sudan Conference on Computer Science and Information Technology (SCCSIT), Elnihood, Sudan*, 2017, pp. 1-8.
- [47] G.Y. Chen, T.D. Bui and A. Krzyzak, "Palmpoint classification using dual-tree complex wavelets", in: *Proceeding of International Conference on Image Processing*, 2006, pp. 2645-2648.
- [48] CASIA Palmpoint Database, <http://biometrics.idealtest.org/>. [Accessed 10 October 2018].
- [49] L. Lei, C. Wang and X. Liu, "Discrete Wavelet Transform Decomposition level determination exploiting sparseness measurement", *International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering*, 2013, 7,1182-1185.
- [50] V. Vapnik, "The Nature of Statistical Learning Theory". NY: Springer-Verlag, 1995.
- [51] A.A. Ross, K. Nandakumar and A. Jain, "Handbook of multibiometrics (Vol. 6)", Springer Science & Business Media, 2006, pp.73-82.
- [52] W. Ouarda, H. Trichili, A. M. Alimi and B. Solaiman, "Towards A Novel Biometric System For Smart Riding Club.", *Journal of Information Assurance & Security* . 2016, Vol. 11 Issue 4, p201-213.
- [53] W. Ouarda, H. Trichili, A. M. Alimi and B. Solaiman, "Bag of face recognition systems based on holistic approaches", 15th International Conference on Intelligent Systems Design and Applications (ISDA), Marrakech, 2015, pp. 201-206. doi: 10.1109/ISDA.2015.7489225.
- [54] I. Jarraya, W. Ouarda and A.M. Alimi, "Deep neural network features for horses identity recognition using multiview horses' face pattern", *Proc. SPIE 10341, Ninth International Conference on Machine Vision (ICMV 2016)*, 103410B (17 March 2017); doi: 10.1117/12.2269064.
- [55] S. M. Eragi, W. Ouarda and A. M. Alimi, "Human identity recognition based on facial images: Via supervised autoencoder features representation," *Sudan Conference on Computer Science and Information Technology (SCCSIT), Elnihood*, 2017, pp. 1-6. doi: 10.1109/SCCSIT.2017.8293056.
- [56] K. Larbi, W. Ouarda, H. Drira, B. Ben-Amor and C. Ben-Amar, "DeepColorFASD: Face Anti Spoofing Solution Using a Multi Channeled Color Spaces CNN", *IEEE International Conference on Systems, Man, and Cybernetics (SMC), Miyazaki, Japan*, 2018, pp. 4011-4016. doi: 10.1109/SMC.2018.0068.