Sudan University of Science & Technology

College of Graduate Studies

بسم الله الرحمن الرحيم

جامعة السودان للعلوم والتكنولوجيا

كلية الدراسات العليا

كلية الدراسات العليا

Approval Page

(To be completed after the college council approval)

Name of Candidate: elaraby | abdou | salman | elgallad

Thesis title: The Use of Neural Networks for Identity Recognition Through Palm Print and Voting Technique

استخدام الشبكات العصبية للتعرف على الهوية من خلال راحة اليد وتقنية التصويت

Degree Examined for: دكتوراه بالمقررات والبحث

Approved by:

1. External Examiner

Name: Mohamed Izzeldin A. Basher

Signature: ..... Date: 5/8/2019

2. Internal Examiner

Name: Howida A. M.

Signature: Howida A. M. Date: 5/8/2019

3. Supervisor

Name: Dr. Adel Alimi

Signature: ..... Date: 5/8/2019

Sudan University of Science and Technology

College of Graduate Studies

The Use of Neural Networks for Identity Recognition Through Palm Print and Voting Technique

استخدام الشبكات العصبية للتعرف على الهوية من خلال راحة اليد وتقنية التصويت

A Dissertation Submitted to the College of Graduation Studies,

Faculty of Computer Science and Information Technology,

Sudan University of Science and Technology

In Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major Subject: Image Processing

By

Elaraby Abdou Elgallad

Supervisor

Adel M. Alimi, Professor

August, 2019

Abstract

Biometry 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, palm print, iris, and hand-geometry have been well explored; and matured approaches are available to perform personal identification.

The human hand is one of the body parts with special characteristics that are unique to every individual. The distinctive features can give some information about an individual, thus, making it a suitable body part that can be relied upon for biometric identification and, specifically, gender recognition. Several studies have suggested that the hand has unique traits that help in gender classification. Human hands form part of soft biometrics as they have distinctive features that can give information about a person. Nevertheless, the information retrieved from the soft biometrics can be used to identify an individual's gender. Furthermore, soft biometrics can be combined with the main biometrics characteristics that can improve the quality of biometric detection. Gender classification using hand features such as palm contributes significantly to the biometric identification domain and, hence, presents itself as a valuable research topic.

Despite a period of remarkable evolution, no extensive comparison and evaluation have been performed up till now to study the effect of the representation of data through the descriptors on palmprint recognition problem. Motivated by this statement, this research aims to fill this gap and provide a comprehensive comparative study of the performance of a large number of recent state-of-the-art texture descriptors in palmprint recognition.

The research emphasizes the opportunities for features representation and analysis from a palmprint image using handcrafted (Curvelet, Wavelet, Wave Atom, SIFT, Gabor,

LBP) and neural approaches based convolutional neural network. All previous features were merged at the decision level by a proposed voting method in order to enhance the identification of a person. The proposed approach was tested in a number of experiments on the CASIA, IITD, and 11k palmprint databases. The testing yielded positive results supporting the use of the described voting technique for human recognition purposes.

This research also explores the use of Discrete Wavelet Transform (DWT) in gender identification, with SqueezeNet acting as a tool for unsheathing features, and Support Vector Machine (SVM) operating as a discriminative classifier.

In this research, we aim also to apply fusion at score level on predictive labels obtained from different descriptors rather than the labels obtained from different classifiers. We study also the effect of using fusion at decision level through Mode Voting Technique (MVT) to achieve a good performance of our proposed system for identity recognition.

From the results, it is clear that the fusion at decision level using the Mode Voting Technique guarantees an excellent recognition rate regardless of low recognition rate of some datasets. The mode voting technique ranks top of the list of SVM classifiers used for each database.

المستخلص

تعتبر القياسات الحيوية من أفضل السمات المستخدمة لتحديد هوية الأشخاص وحماية المعلومات من الاختراقات الخارجية. وتم استخدام هذه السمات مثل البصمة والوجه وراحة اليد وقزحية العين كسمات اساسية في انظمة التعرف على الهوية الشخصية.

تعد راحة اليد من اهم أجزاء الجسم ذات الخصائص المميزة والفريدة لكل فرد والذي يمكن الاعتماد عليها لتحديد هوية الشخص، كما يمكن استخدام هذه الخصائص لتحديد جنس الفرد، علاوة على ذلك، يمكن الجمع بين القياسات الحيوية المختلفة لتحسين جودة الأنظمة التي تستخدم للتعرف على الأشخاص. وعلى الرغم من التطور الملحوظ في هذا المجال، لم يتم إجراء مقارنة وتقييم فعال لدراسة تأثير الخوارزميات المختلفة المستخدمة في تمثيل هذه البيانات المستخلصة من القياسات الحيوية للتعرف على الأشخاص، ولذا يهدف هذا البحث الى تقديم مقارنة شاملة لأداء هذه الخوارزميات. ومن هذه الخوارزميات المطروحة:

Curvelet, Wavelet, Wave Atom, SIFT, Gabor, LBP

جنبا الى جنب مع استخدام الشبكات العصبية Convolutional Neural Network

ومن اهم الاسهامات التي قدمت من خلال هذه الدراسة هي دمج النتائج النهائية - والتي تم الحصول عليها من خلال Support Vector Machine - باستخدام Mode Voting Technique لتحديد هوية الشخص، كما قدمت الدراسة نظاما فريدا للتعرف على جنس الفرد باستخدام Discrete Wavelet Transform وConvolutional Neural Network

تم اختبار الانظمة السابقة من خلال قواعد البيانات المختلفة لراحة اليد مثل:

CASIA, IITD, and 11k palmprint databases

وقد اظهرت النتائج فاعلية هذه الانظمة ودورها في رفع كفاءة التعرف على الاشخاص باستخدام راحة اليد.

Table of Contents

List of Tables

List of Figures

CHAPTER 1. **Introduction**

1.1 Overview

The fraud attempts continue to grow constantly in our society. So, it has become necessary to verify the identity of a person during his access to local or crossing frontiers. Although traditional means such as badge and password may control the identity of a person, however, they have some disadvantages. For example, access card can be falsified or duplicated, ID card can be omitted or lost and passwords may be forgotten or compromised. Biometry is a technology that allows the verification of the actual identity of a person using physiological features which are specific to each individual. It uses several modalities to demonstrate the identity of a person, including biological traits like saliva, urine, odor or DNA; behavioral traits like dynamic signature, gait or speech and physical traits like iris, hand shape and palm print (Charfi et al, 2016).

Palmprint trait is thus, another personal physiological modality which presents different kinds of features such as principal lines, wrinkles and ridges. These features are invariant over time and claimed to be unique to each individual. So, they may be exploited in personal recognition process. Several methods were developed in order to identify a person via his palmprint features. In this research, palmprint features are extracted using standard transforms such as Gabor, Wavelet, Wave Atom, Curvelet, SIFT, CNN, and LBP.

Features classification is the last step in recognition process. Many methods are offered in literature to validate the identity of users. Classification based similarities is the simplest method for classification which computes similarity between user to identify and users enrolled in database. It is based on measure similarity which can be based on many types of distances like Euclidean, Cosine, MahCosine, Hausdroff, etc. Classification based on probabilities computes in this case the probability of the

membership of object to such class such as Naive Bayesian and Parzen classifiers. Classification based on Decision Boundary. Many types of classifiers are used to classify object as example Neural Network, Binary Decision Tree, RBF, and Support Vector Machines, etc. In this research the SVM is used as a classifier, due to its advantages that it has a strong founding theory, reaches the global optimum due to quadratic programming, has no issue for choosing a proper number of parameters, it is less prone to overfitting, needs less memory to store the predictive model and yield more readable results and a geometrical interpretation (Vapnik, 1995).

The fusion after matching step is based principally on classifiers fusion. In fact, this fusion type is the most studied one by researchers. Such a fusion may be performed at score level, at rank level or at decision level. Fusion at score level is referred to as the combination of similarity scores derived from different classifiers. This type of fusion is the most commonly used one since it may be applied to all types of systems (Ross et al., 2006).

The main contribution in this research is Mode Voting Technique (MVT) that 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. MVT is used to establish the common non-repeated values in the predict label array for the purpose of biometric identification and the Inference is made using mode voting approach.

Soft biometry is a new trend in security and video surveillance systems. It provides additional information to recognize identity of person and animal. The most motivation towards soft biometrics integration into traditional biometric system is that the classification of soft biometric traits can be done by using same techniques carried out in monomodal biometric system (Jain et al., 2004). This research focuses on gender

recognition in biometric identification and explains how it can be used to enhance the accuracy of biometric identification processes.

Recently, convolutional neural networks (CNNs) have been largely outperforming hand-engineered features-based methods. On the effectiveness of using the trained CNN as a generic feature extractor - handcrafted features can be substituted by deep transfer learning features achieving highly accurate results for image classification.

Dense Hand-CNN palm identification system and New Convolution Wavelet Neural Network (CWNN-Net) for Gender Classification system are proposed.

1.2 Historical Background

Table 1-1 summarized some related works focused on palm print identification and gender recognition systems using different descriptors and several classifiers.

Table 1-1 Related works focused on palm print identification and gender recognition

Author	Features Extractor	Features Classifier	Database	RR (%)
Zhao et al., 2013	SIFT	competitive code algorithm	IITD	Equal Error Rate = 0.49
more robust and accurate classifier can improve the performance of this system				
Varshney et al., 2014	DWT - DCT	Euclidean Distance	IITD, PolyU	94.44, 95.65
Although Euclidean distance is very common in clustering, it has a drawback: if two data vectors have no attribute values in common, they may have a smaller distance than the other pair of data vectors containing the same attribute values. Another problem with Euclidean distance is that the largest-scaled feature would dominate the others.				
Charfi et al., 2014	SIFT	Matching Score	IITD	Palmprint = 94.05 Hand shape + Palmprint = 97.82
Jaswal et al., 2015	2D Gabor filter	Euclidean Distance	CASIA IIT Delhi	90.76 91.4
Misar and Gharpure, 2015	Discrete Wavelet Transform	Neural Network	IITD	75.6
The low recognition rate may be due to the disadvantage of NN as a classifier which is: Local minima, over fitting, the processing of ANN network is difficult to interpret and require high processing time.				
Charfi et al., 2015	SIFT and Gabor	Matching score	IITD	Palmprint = 91.08 Hand shape + Fingers + Palmprint = 98.04
Charfi et al., 2016	SIFT sparse representation	SVM	IITD Bosphorus	IITD: Palmprint = 96.73 Hand shape + Palmprint = 99.57 Bosphorus: Palmprint = 94.95 Hand shape + Palmprint = 97.61
Better performance as compare to other classifier, easy handle complex nonlinear data points and over fitting problem is not as much as other methods				
Afifi, 2017	CNN-features + LBP	SVM	IITD 11k	IITD: CNN Fea. = 90 CNN Fea. + LBP = 94.8 11k: CNN Fea. = 94.8 CNN Fea. + LBP = 96

From this given state of the art, we are absolutely recommended to make experimental study on palm recognition to compare between the performances of different descriptors.

We intend in the following chapters, to validate this state of the art by set of experiments on palm identification and gender recognition to choose the most appropriate techniques used to feature extraction. Recent works done also shows that Linear SVM ranks top the list of classifiers was used so we recommended it as a classifier to our proposed systems. Recently, convolutional neural networks (CNNs) have been largely outperforming hand-engineered features-based methods, which is used as a generic feature extractor - handcrafted features that can achieve highly accurate results for image classification.

1.3 Motivation and Problem Statement

Biometric identification is the process of recognizing individuals using characteristics that can be behavioral or physiological. The palmprint recognition has become one hotspot in the field of identity recognition because it has rich texture information, stable characteristics, low-resolution image, low-cost collection, easy acceptability and high accuracy recognition rate. Many studies show that hand dimensions possess distinctive features that can reveal gender information. This kind of information is one of a set of attributes known as soft biometrics, which can be extracted from human biometric cues. Such soft biometrics can be integrated with primary biometrics traits to boost the accuracy of biometric identification. An advantage of hand images is they are usually captured under a controlled position, unlike face images, which are usually unconstrained. Additionally, hands have less variability compared to, for example, faces, which are usually represented by deformable models due to facial expression changes.

Palm print area is one of the human traits that researchers in biometric fields have given a huge consideration in recent years. Many researches have been proposed in the literature to design palmprint recognition framework using many different descriptors to extract image features in most of these approaches. Despite a period of remarkable evolution, no extensive comparison and evaluation has been performed up till now to study the effect of the representation of data through the descriptors on palmprint recognition problem. Motivated by this statement, this research aims to fill this gap and provide a comprehensive comparative study of the performance of a large number of recent the state-of-the-art texture descriptors in palmprint recognition.

In this research, we aim also to apply fusion at score level on predictive labels obtained from different descriptors rather than the labels obtained from different classifiers. We study also the effect of using fusion at decision level through Mode Voting Technique (MVT) to achieve a good performance of our proposed system for identity recognition. So, we aim in our research to investigate on the utility of hard and soft biometrics in human identity recognition using palm print images.

Most state-of-the-art methods, which use hand images for either gender recognition or biometric identification, follow the traditional pipeline. Firstly, handcrafted features are extracted, followed by either training an off-the-shelf classifier or using a similarity measurement metric to compare such features with templates in the database. However, recently, convolutional neural networks (CNNs) have been largely outperforming hand-engineered features-based methods.

1.4 Research Questions

1- How can the representation of data through the descriptors in the feature extraction step in building the classification model affect the predictive and computationally of the performance of palm identification system?

2- How can the fusion at decision level achieve an interesting classification accuracy and improve the recognition rate of the identification system?

3- What is the role of CNN in the improvement of the performance of the identification system?

1.5 Research Objectives

The physiological modalities are based on the measurement of parts of the human body such as palmprint, footprint, fingerprint, hand shape, face, retina, and iris. The main objective of this research is to propose a new Human Identity Recognition system based on Body's Limbs Images. Palmprint and hand shape are the most confident physiological modalities used in this research. We can summarize the main objectives as follow:

- Propose a new palmprint identification approach which fuses different representations. These representations are extracted from the texture information of palmprint modality such as Gabor, Wavelet, Wave Atom, Curvelet, SIFT, CNN, and LBP.
- Using the sparse autoencoder at the stage of features transformation in order to enhance the performance of the extracted features.
- Perform the fusion of these different features using the mode voting technique (MVT).

- Prove that soft biometrics can be integrated into traditional recognition system. Propose a new convolution wavelet neural network system for Gender Classification using Palm print.
- Evaluate the performance of the proposed approach using IITD, CASIA and 11k databases.

1.6 Methodology

Figure 1-1 Shows the research framework of the contributed systems.

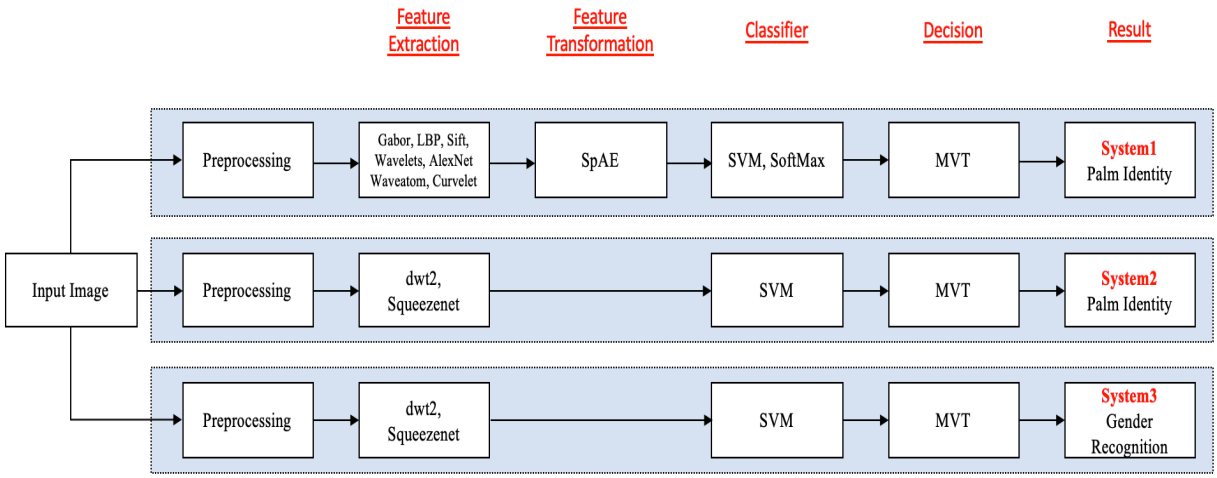

Figure 1-1 Research Framework

1.7 Thesis Contributions

The main goal of biometric system is to recognize person's identity. Two existent architectures still exist: Mono-modal and Multi-modal Biometric systems. We detailed in previous section the advantages and the drawbacks of both architectures. The current research aims at researching the palm print identification system and the integration of soft biometrics into traditional recognition systems for gender recognition. Figure 1-2 summarizes below all contributions of our dissertation.

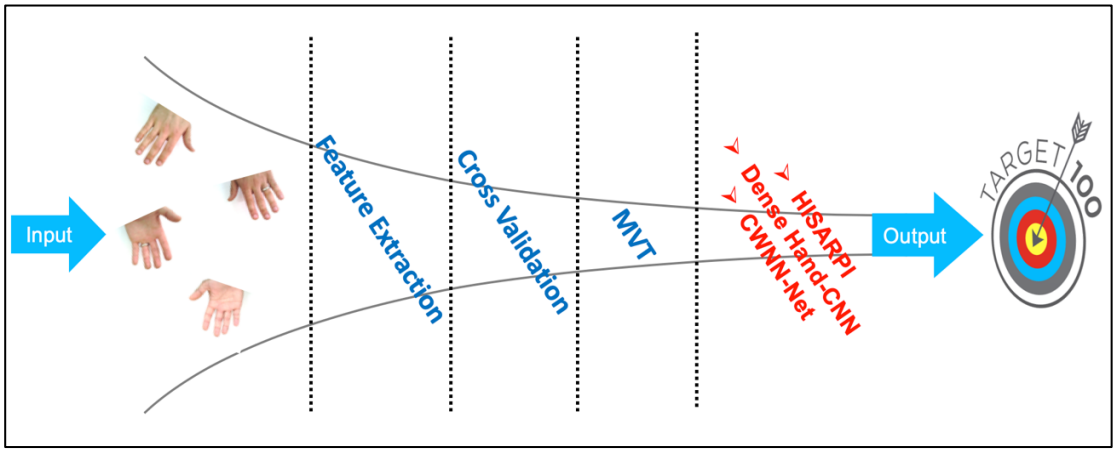

Figure 1-2 Summary of Research Contributions

1.8 Thesis Outline

The thesis is organized into seven chapters.

- In chapter1, we introduced the motivation and the problem statement of the research. We presented the types of biometric systems, multimodal and unimodal biometric systems, then, we detailed the limitations of unimodal biometric systems and the Advantages of multimodal biometric systems. We introduced the role of soft biometrics and convolutional neural network in improving the performance of unimodal biometric systems. The motivation problem statement, research questions and research objectives are presented. This chapter was ended by a summary of this research contributions.
- In chapter 2, we firstly present a state of the art in the field of human identity recognition. This chapter comprehensively explores the existing literature on human identity recognition and gender identification.
- In chapter 3, the research methodology is illustrated and discussed for the related work done in the different systems.
- In chapter 4, we emphasize the opportunities for obtaining texture information from a palmprint on the basis of descriptors in the Dense Hand-CNN system.

We give a comprehensive experimental study on palmprint recognition to justify the use of such descriptors and classifiers in our proposed scheme for features representation.

- In chapter 5, the CWNN-Net system focuses on gender recognition in biometric identification and explains how it can be used to enhance the accuracy of biometric identification processes.
- In chapter 6, the experimental results on palmprint recognition is stated in details foreach descriptors and classifiers in our proposed systems.
- Lastly, in chapter 7, we discuss and summarize all research achievements and findings of the systems' results and eventually provides the conclusion of this research. We finish by giving some lines of future work.

CHAPTER 2. **Related Work**

2.1 Introduction

In this chapter, we will present a comprehensive state of the art of all fields concerned in our dissertation. We focused on recent related work on human identity recognition as well as we studied in other hand the soft biometrics classification. In addition, we discussed the use of different techniques performed in literature to recognize identity and to classify.

2.2 Biometric Systems

Biometrics is a verification technique that incorporates science (biology) as well as technology. it is an authentication method that uses human characteristics to identify a person. It is highly applicable in information assurance. The method uses human biological data, including DNA and fingerprints to ensure safe entry, data access, and protection. The systems normally have high-tech elements that are connected to produce powerful performance. Based on the traits, biometric can be divided into two broad types (Charfi et al, 2016): 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 secure than physical biometric because people can change their behavior anytime they want. For example, people can adjust their signature, keystroke, or walking pattern easily.

Biometric systems are popular among institutions that have security systems as well as replacement systems including PIN, and ID replacement. The main distinction between biometric system and other conventional systems is the fact that the former requires the physical presence of the individual who is using the system. This physical presence

creates an extra level of security and makes it difficult for identity thieves to use a false ID card or any other stolen mode of identification.

2.2.1 Multimodal Biometric Systems Vs Unimodal Biometric Systems

Unimodal systems are biometric identification systems in which a single biometric trait of the individual is used for identification and verification. Biometric identification systems which are accomplished to use a combination of two or more biometric modalities to identify an individual are called multimodal biometric systems. The main advantage of using multimodal biometric systems is to increase the recognition rate.

When implementing any biometric system for identification, the most significant question is whether to choose a unimodal or multimodal biometric system.

2.2.2 The Limitations of Unimodal Biometric Systems

One of the most important limitation of unimodal biometric systems is the susceptibility of the biometric sensor to noisy or bad data. The captured biometric trait might be distorted because of unsatisfactory acquisition conditions. This limitation can be seen in applications which use facial recognition. The quality of the captured facial images might get affected by illumination conditions and facial expressions. Another example could be in fingerprint recognition where a scanner is unable to read dirty fingerprints clearly and leads to false database matches. An enrolled user might be incorrectly rejected whereas an impostor might be falsely accepted.

It might not be compatible with certain groups of population. Because of faded fingerprints or underdeveloped fingerprint ridges, fingerprint images might not be correctly captured for the ageing and young children. Though the biometric traits are expected to exist among every individual in a given population, there could be some

exceptions where an individual is unable to provide a particular biometric. For example, if the subject has a pathological eye condition, iris images might not be acquired.

Within a large population, unimodal biometrics is disposed to inter-class similarities. Facial recognition may not work correctly for identical twins as the camera might not be able to distinguish between the two subjects leading to inaccurate matching

Unimodal biometric systems are quite vulnerable to spoof attacks where the data can be imitated or forged. For example, fingerprint recognition systems can be easily spoofed using rubber fingerprints.

2.2.3 Advantages of Multimodal Biometric Systems

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.

2.2.4 Soft Biometrics

Soft biometrics has attracted an enormous range of researchers as a way to enhance the performance of the unimodal biometric system while not the requirement to multi-modal system. In this approach, soft biometrics have distinct characteristics that divide persons into non-overlapping groups using soft traits like gender, age, ethnicity, skin color, hair color, etc. several existing biometric systems use soft data like eye color, height, gender and age of users throughout enrollment.

However, in scientific literature, most biometric systems use one attribute for recognition, for this reason, they're additionally known as unimodal biometric systems.

In alternative words, these systems of human recognition use, in general, solely the first biometric identifier (fingerprint, face, hand- geometry, etc.) and also the soft data is rarely used.

These types of systems are meet totally different issues like noise information capture, lack of permanency of the biometric attribute, and attacks. A number of these issues of unimodal biometric systems is resolved by adding Soft Biometrics to the unimodal biometric systems and also the use of multi-modal biometric systems. So, to enhance the accuracy of biometric identification processes, the gender recognition system is proposed in this research.

Figure 2.1 shows associate example of recognition system based on hard and soft biometrics.

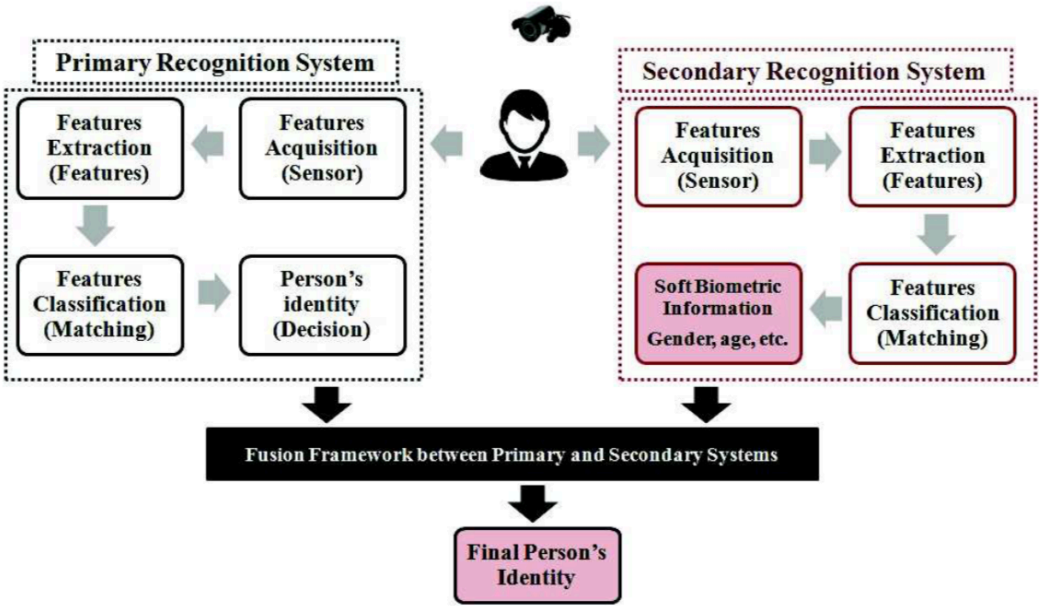

Figure 2-1 Soft and Hard Biometric Systems Architecture (Jain et al., 2004)

2.3 Convolutional Neural Network

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 (Krizhevsky et al, 2012). 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 came out in 2014 (Simonyan and Zisserman, 2014), 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 AlextNet. The GoogleNet architecture (Szegedy et al, 2015) 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 (He et al, 2016). 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. The 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 (Huang et al, 2017). The above networks implement comprehensive cross-layer feed-forward pathway connections. Such an arrangement made it possible to outperform RestNets, 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 Iandola et al (2016). 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 a

training, and has more feasibility for implementation on field-programmable gate arrays (FPGAs), as well as other limited memory hardware.

2.4 Related Work on Palmprint Identity Recognition

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.

2.4.1 Hand Crafted Approaches

Han (2004) 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.

Meiru et al. (2010) 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 et al. (2011) solve the problem of linked features which generated because of using PalmCode for distinctive palms. In order to remove this correlation, Zeng 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. Varshney et al. (2014) 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.

Jain et al. (1999) applied Hough transform for an extraction of distinct fingerprint features. This study will consider implementing Dempster-Shafer evidence theory and Bayesian fusion technique (Xu et al., 1992), majority rule (Lan and Suen, 1997), weighted majority algorithm (Kuncheva, 2004), behavior-knowledge space method (Huang and Suen, 1995), and disjunction ("AND") vs conjunction ("OR") models (Daugman, 2000) 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 (Fogel and Sagi, 1989), (Jaswal et al., 2015), and (Teuner et al., 1995).

One study of palmprint recognition (Kekre et al., 2012) established the productivity of wavelet transforms in generating a successful 90% mid-level reference. Another study (Misar and Gharpure, 2015) 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 (Haddad et al., 2009). In (Mohammed et al., 2010), 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 (Sumana et al., 2008), 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, Dong et al. (2005) firstly implemented digital curvelet transform and the recognition rate of the experiment was up to 95.25%. In (Xu et al., 2009), the Support Vector Machine (SVM) as a classifier for the curvelet decomposed features of palm-print, the recognition accuracy became 98.5%. In (Liu et al., 2015), 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, 2008) effectively extracted palmprint features using Scale Invariant Feature Transformation (SIFT) descriptions and fusion based on Symbolic Aggregate Approximation (SAX). Zhao et al. (2013) fused competitive coding and SIFT to enhance palmprint verification.

In their turn, Charfi et al. (2104) 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 (2007).

Charfi et al. (2015) 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 REgim Sfax Tunisia (REST) hand database (Charfi et al., 2015) and CASIA Palmprint Database (Charfi et al., 2016). In particular, Charfi et al. (2016) 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 (Ojala et al., 1996). Researchers Wang et al. (2006) 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, LBP shaped the basis for feature extraction in an enhanced identification approach grounded in the directional shifts of gradient operator (Michael et al., 2008). Finally, the above method was further improved in the aspects of speed and precision (Promila and Laxmi, 2012).

2.4.2 Neural Approaches

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.

LeCun et al. (1989) 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, Krizhevsky et al. (2012) 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 (2014). 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 was designed by Szegedy et al. (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 budged. 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 (He et al., 2016) 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 Huang et al. (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 (Iandola, 2016) 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 five hundred and ten 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. Zhao et al. (2015) 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 (2016) 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.

A deep CNN was also employed in a project implemented by Sun et al. (2017) 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 by Hu et al. (2017) to

provide explicit inter-channel modeling for purposes of improving the flexibility of channeled feature reactions. The researchers found that Squeeze-and-Excitation Networks (SENets) 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 (2017) 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 2-1 summarizes the unimodal system palmprint methods, proposed in the literature.

Table 2-1 Comparison between palmprint biometric systems in the related work (Elgallad et al., 2017)

Author	Features Extraction	Features Classification	Database	RR%
Han, 2004	wavelet	GLVQ approach	grabbed from a CCD camera	FRR = 1.6, FAR = 36.3
Dong et al., 2005	Digital Curvelet Transform	Euclidian distance classifier	PolyU	95.25
Chen *et al*., 2006	dual-tree complex wavelet transforms	SVM	PolyU Palmprint Database	97
Wang et al., 2006	Boosting Local Binary Pattern	Chi square distances	UST-HK palmprint database	Equal Error Rate = 2
Chen and Moon, 2008	SIFT	Symbolic Aggregate approximation	PolyU	Equal Error Rate = 0.37
Michael et al., 2008	Sobel and LBP	Chi square and PNN	palm print tracking in dynamic environment	PNN: EER=0.74 Chi: EER=1.52
Xu et al., 2009	Digital Curvelet Transform	SVM	PolyU	98.5
Mu et al., 2010	discriminative local binary patterns statistic (DLBPS)	nearest neighbor (NN) classifier based on the Euclidean distance	Fujitsu fi-60F high speed flatbed scanner is used	98
Zeng and Huang, 2011	Gabor features	Euclidean distance and the nearest neighbor classifier	PolyU Palmprint Database	100
Kekre et al., 2012	Wavelets	mean square error	Hong Kong Polytechnic University 2D_3D Database	93

Author	Features Extraction	Features Classification	Database	RR%
Promila and Laxmi, 2012	LBP	Chi-square test and Pearson correlation test	PolyU	99.22
Zhao et al., 2013	SIFT	competitive code algorithm	IITD	Equal Error Rate = 0.49
Varshney et al., 2014	DWT - DCT	Euclidean Distance	IITD, PolyU	94.44, 95.65
Charfi et al., 2014	SIFT	Matching Score	IITD	Palmprint = 94.05 Hand shape + Palmprint = 97.82
Jaswal et al., 2015	2D Gabor filter	Euclidean Distance	CASIA IIT Delhi	90.76 91.4
Zhao et al., 2015	DBN	DBN	Beijing Jiao Tong University	90.63
Misar and Gharpure, 2015	Discrete Wavelet Transform	Neural Network	IITD	75.6
Liu et al., 2015	Digital Curvelet Transform	nearest neighbor method	PolyU	99.9
Charfi et al., 2015	SIFT and Gabor	Matching score	IITD	Palmprint = 91.08 Hand shape + Fingers + Palmprint = 98.04
Charfi et al., 2016	SIFT sparse representation	SVM	IITD Bosphorus	IITD: Palmprint = 96.73 Hand shape + Palmprint = 99.57 Bosphorus: Palmprint = 94.95 Hand shape + Palmprint = 97.61
Minaee and Wang, 2106	Scattering Features + PCA	minimum distance SVM	PolyU	99.95 100
Sun et al., 2017	CNN-F architecture	Softmax	PolyU	100
Afifi, 2017	CNN-features + LBP	SVM	IITD 11k	IITD: CNN Fea. = 90 CNN Fea. + LBP = 94.8 11k: CNN Fea. = 94.8 CNN Fea. + LBP = 96

2.5 Gender Recognition

There literature on gender recognition remains scarce when compared to person identification that has attracted several academicians and experts in the field. The scarcity of related literature can be explained by the fact that person identification, as opposed to gender identification, has a practical security interest for the government and law

enforcement agencies. Nevertheless, gender recognition has the potential to become the next hot topic when it comes to human-computer communication. Furthermore, gender identification can also be a valuable tool for various organizations that need intelligent advertising based on accurate gender identification.

2.5.1 Hand Craft Related Work

Amayeh et al. (2008) is one of the pioneers who has explored the concept of using computers to classify gender based on the distinctive features of the hand. He came up with a system that splits the hand delineation into various fragments; it aligns with the fingers as well as the palms, and defines the features of every section using different techniques comprising boundary descriptors, file descriptors, region descriptors, and ZMs.

Another researcher, Ming et al. (2014), developed a palm geometric-oriented method for identifying whether a person is a male or a female with the support-vector machine. This method is less complicated and has been shown to be effective with regards to gender classification. They note that this method guarantees a high-level of accuracy without using any intricate computations, processes, and procedures that have been used less successfully in other biometric techniques. Instead, it is characterized by few features, including the ability to record hand features without the users having to peg their hands on the device. Its efficiency, simplicity, and convenience make it appropriate for real-world execution.

Meanwhile, Font-Aragonesin and Faundez-Zanuy (2013) propose a technique that uses anthropometric hand data to identify the respective genders of the sample being tested. Using their method, researchers collected data on unvaried number of several men and women in their visual hand database. Majority of the information was collected on

men. A simple approach was then devised to get the measurements of the users' hands. The information was transmitted through a Biometric Dispersion Matcher (BDM) to retrieve the suitable data. BDM functions as a quadratic discriminant classifier. This discriminant classifier begins with filtering out data that will not assist in disclosing the user's gender. It then proceeds to display a vector of the key computations. The technique had a performance rate of 95% where the ratio of men to women was 2:1, with a projection that the accuracy could become higher when data records increase.

Meanwhile, it is apparent that many researchers have focused on Local Binary Patterns (LBP) and the alternatives in computer vision that can be used in gender classification. One of the suggestions shared by researchers, in this case Iglesias et al. (2014), is Binary Robust Independent Elementary Features (BRIEF), Oriented FAST and Rotated BRIEF (ORB) as well as Binary Robust Invariant Scalable Keypoints (BRISK), which have been reported to be reliable and fast in gender classification. When compared to LBP, the latter variants provide speedy detection while still maintaining high performance rates as LBP. This makes them better than LBP when used in systems that require constant gender recognition. Meanwhile, there is no sufficient literature to validate that these options are more suitable than LBP.

Another approach that captured the interest of researchers is gender recognition through boxing action. Wang et al. (2011) tested this technique whereby a period detection procedure is applied and thereafter, an averaged profile is used to denote a boxing sequence of a period. The human classification is done using Nearest Neighbor Classifier (NNC), which basically compares the recorded data with the data set that is being tested. The NNC was based on Euclidian metric. The tests were done on the KTH-Dataset, which has a good performance rate of at least 80%. PCA was used to extract features, and finally SVM was used to allocate the respective genders. Yang et al. (2011)

has also recommended Global Local Feature Fusion (GLFF) as a reliable method for gender recognition. The study relied on findings from psychophysics as well as neurophysiology observations, which asserted that universal and local data is important in image perception. The initial step in the GLFF technique is mining universal and local traits through Active Appearance Model (AAM) and LBP tool respectively. One then proceeds to merge the universal and local features through sequent selection. Eventually, the chosen traits are used to determine the receptive genders in the dataset with the help of SVM. The test was made on a sample of 20 men and 20 women with the results having an average accuracy rate of 80% and an indication that the accuracy levels could go up. Based on these test results, it was apparent that blending local and global features could significantly improve the precision of gender recognition and the performance of gender classification tools.

Furthermore, the study presented the significance of different body features on gender classification. The numerical analysis revealed that the head makes a substantial contribution in gender recognition with the buttocks and leg making little contribution towards accurate gender identification.

Conversely, Azzopardi et al. (2016) suggested the use of facial features to classify gender by combining the findings of different SVM classifiers. Typically, three SVM classifiers are used to achieve this goal. The first descriptor, Histogram of Oriented Grardients (HOG), analyzes the features that pertain to shape with the help of histogram intersection kernels. The second visual descriptor, LBP, focuses on the texture of the facial features, also with the help of histogram intersection kernels. The third descriptor analyses the raw pixel values using a linear kernel. The researchers used this approach to examine data that had been mined from FERET. The good performance rate of this test stood at 92.6%, which surpasses the accuracy of some commercial tools such as Face++.

Ardakany et al. (2012) dealt with the challenges of gender recognition through genetic algorithms. The researchers mined relevant data on facial feature records available in FERET database using LBP and PCA. The algorithm allows extracting the features that enhance the capability of the SVM classifiers to accurately recognize whether one is male or female. The genetic algorithm condenses the features being analyzed from 142 to about 71, and thereby increasing level of accuracy to about 98.5, making the results completely reliable.

Collins et al. (2014) sought to identify gender through analyzing the different walking style of individuals. The researchers were inspired by previous studies, which have shown that gait can be used in behavior recognition tests. SVM was used to deduce the respective genders of the sample data in the analysis of the sparse spatio temporal features. The method attained a decent performance rate of 87%.

Another popular proposal in the current literature is the one made by Haitpoglu and Kose (2017), which suggests combining Speed-Up Robust Features (SURF) of bags-of-visual-words (BOW) and SVM. This proposal was tested on different parts of 3560 face samples retrieved from the FERET database to determine the degree of its efficiency and reliability. The results showed that the approach is efficient in classifying gender for records retrieved from FERET database.

From their experiment, inception-v4 network had the highest accuracy level at 98.2% with Adience dataset coming second at 84%. Crime and Pedrini (2017) used a pre-defined face silhouette prototype to create geometric descriptor for classifying gender. The researchers used this method to analyze four sets of face data records and reported better performance than other techniques that use geometric descriptors.

Another study, Liu et al. (2017), focusing on how the unique manner of walking could help in gender recognition shared a more advanced gait energy image (GEI) titled D-GEI.

This was a step in the right direction in terms of building the literature on gait analysis and its place in gender recognition. Most of the existing studies have not extensively dealt with gait analysis, which has resulted in poor performance of gait evaluation in human identification.

The procedure in the D-GEI method begins with creating a dynamic region, followed by establishing the dynamic region of frame, and then computing the weighted average of the dynamic region. These steps help in determining the D-GEI. Once the latter has been established, HOG is used in gradient computation, while SVM classifier helps to determine the respective genders of the sample. Azzopardi et al. (2018) presented another method that could be used in gender recognition based on the facial features of the user. In this proposal, the domain-specific as well as trainable traits are merged to help in gender classification. Fifty-two facial features linked to the eyes, nose, as well as the mouth were mined to act as SURF descriptors.

Likewise, COSFIRE acted as the trainable traits. This method responds strongly to some of the notorious challenges associated with face profiles such as the expression variations, light adjustments, and different poses. It attains a high accuracy rate on some of the top datasets in gender classification: GENDER-FERET and LFW. The technique was also highly reliable when real-world data records were used. The datasets comprised 206 training (144 men and 62 women) images, as well as 200 test (139 men and 61 women) images that were taken when the sample population on the normal walking motion. The purpose of assessing the performance of this algorithm on real scenarios was to predict its accuracy on analyzing images retrieved from videos, that are likely to be more problematic that standard data records. Despite the problematic issues in real scenarios, the algorithm was 91.5% accurate in classifying the genders. The researchers further observe that the COSFIRE is more reliable than the SURF descriptors and

therefore, recommends that it can be used solve several types of visual pattern classification problem.

2.5.2 Neural Approaches Related Work

The respective gender was allocated based on the consolidated evidence from the hand profile with the match score being determined using the score-level fusion and file descriptors. The best performance was a match score of 98%. Notably, the evidence provided by ZMs after utilizing the score-level fusion was almost similar to that of file descriptors. Afifi (2017) shared a logical record of hand images that can be used in biometric detection and gender identification. Authors presented a collection of valuable metadata drawn from Mahmoud's logical record. The researchers also proceeded to offer top-notch approaches for using the data record to in gender detection.

There is also a convolutional neural network (CNN) based on a two-steam data that indicates the gender identification challenge. CNN helps to disclose features that are then conveyed to a support-vector-machine that analyzes data and classifies the biometric recognition issue. There is a strong belief that the suggested data record will significantly contribute towards the creation of better gender detection and biometric systems that rely on hand images.

Gender identification through the use of facial features remains a problematic issue in the field of computer vision. Previous efforts in the face image gender recognition concentrated on enhancing understanding the neural network. Key efforts included the double-layered neural network known as SEXNET that was the brainchild of Gollomb and Lawrence (1991). Another development in face image gender recognition was that of Yen (2011), which sought to eliminate the inaccuracies that are caused by less important data such as changes in face lighting as well as adjustments in facial expressions. The authors developed a novel algorithm that could help to reduce the impact

of the latter disturbances. The 2-D Gabor transform was effective in spotting the facial points, with the SVM discriminative classifier picking out distinctive features to determine the respective gender of the user.

Notably, in another study, Akbulut et al. (2017) sought to determine if face images with deep learning could help in gender classification. Specifically, they used Local Receptive Field-Extreme Learning Machine (LRF-ELM) as well as Convolutional Neural Networks (CNN). The technique was tested using data records that were meant to reveal the age and gender of the sample. The outcome of the experiment revealed an accuracy score of 80% for LRF-ELM, and 87.13% for CNN.

Meanwhile, in the study of Ng et al. (2013), a discriminatively-trained CNN was found to be effective in classifying the gender of pedestrians. The CNN has a complex tier of neural networks that fuse feature mining and grouping into a solitary framework. The researchers recorded an accuracy of 80.4% when they used simple architecture and nominal prepossessing of dataset that provided complete body features of the pedestrians that were sampled. The accuracy of the results is equivalent to other high-tech approaches that do not use hand-manufactured trait miners.

However, it provided more accurate recognition on the frontal face images than those on the left and right side. Meanwhile, Nistor et al. (2017) proposed an automated gender classification technique that relies on CNN. This method begins with network training, which is accomplished through combining numerous face datasets retrieved from different databases including ~70000 facial images from the World Wide Web. Once the data for the networks had been recorded, they were assessed, and compared with various network architectures that displayed better performance.

CHAPTER 3. **Literature Review**

3.1 Introduction

The current chapter emphasizes the opportunities for how to obtain texture information from a palmprint and hand images with different descriptors. How the obtained features can be classified and fused.

Later, we will describe the methodology of the proposed approach for gender recognition.

3.2 Palmprint Identification

In the following parts, we will describe the methods of how to obtain texture information from a palmprint on the basis of such descriptors as Curvelet, Wavelet, Wave Atom, SIFT, Gabor, LBP, and AlexNet.

3.2.1 Wave-Packet Transforms

Wave-packet transforms such as Gabor function, wavelets, wave atoms and curvelet are used in this system. 2D wave packets is denoted as $\varphi_{\mu}(x_1, x_2)$. The main parameters for wave packet architectures which are used as indicator are α and β. If (α = 1), this indicates that multiscale decomposition is used. While (α = 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 (β = 1), it is extended and entirely directional when (β = 0). When (α = 1, β = ½), this match Curvelet. Wavelets are match (α = β = 1), and the Gabor transform is match to (α = β = 0). Wave atoms are described as the point α = β = 1/2.

3.2.2 Gabor Filter

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 as in Figure 3-1.

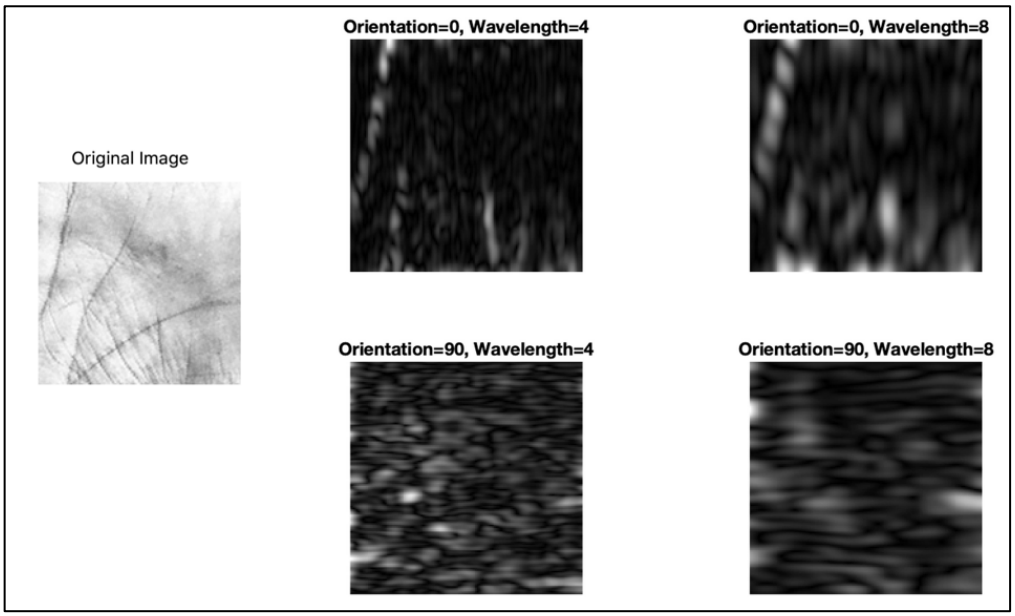

Figure 3-1 Gabor Filter, Orientation [0 90], Wavelength [4 8].

3.2.3 Fast Discrete Curvelet Transform

To extract the curvelet coefficients, Fast Discrete Curvelet Transform via wedge wrap is used for 128x128 input images (Figure 3-2) as in (2) 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 should be evaluated to discrete curvelet factors which are still a fitted structure:

$$c_{j,\ell,b} = \int \hat{f}(\omega) U_{j,\theta_\ell}(\omega) e^{ib.\omega} d\omega, \quad (2)$$

$$where \ f(x) = \sum_{j,\ell,k} \langle f, \varphi_{j,\ell,k} \rangle \varphi_{j,\ell,k}(x), \qquad (conv.\, in\, L^2)$$

Figure 3-2 Curvelet Filter, the original image and all the curvelet coefficients.

3.2.4 Wave Atom Transform

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 (Figure 3-3). As described in (2), the wave atom transform is utilized to extract a cell array comprising the related coefficients from the input image with adjusted size of 64x64:

$$|\varphi_\mu(\omega)| \le C_M . 2^{-j}\left(1 + 2^{-j}\left|\omega - \omega_\mu\right|\right)^{-M} + C_M . 2^{-j}\left(1 + 2^{-j}\left|\omega + \omega_\mu\right|\right)^{-M} \quad (2)$$

$$|\varphi_\mu(x)| \le C_M . 2^{-j}\left(1 + 2^{-j}\left|x - x_\mu\right|\right)^{-M} \; for\; all\; M > 0,$$

$$where\; x_\mu = 2^{-j}n \qquad \omega_\mu = \pi 2^{-j}m \qquad C_1 2^j \le \max_{i=1,2}|m_i| \le C_2 2^j$$

where C_1 and $C_2 > 0$, and will be indirect by the details of the execution.

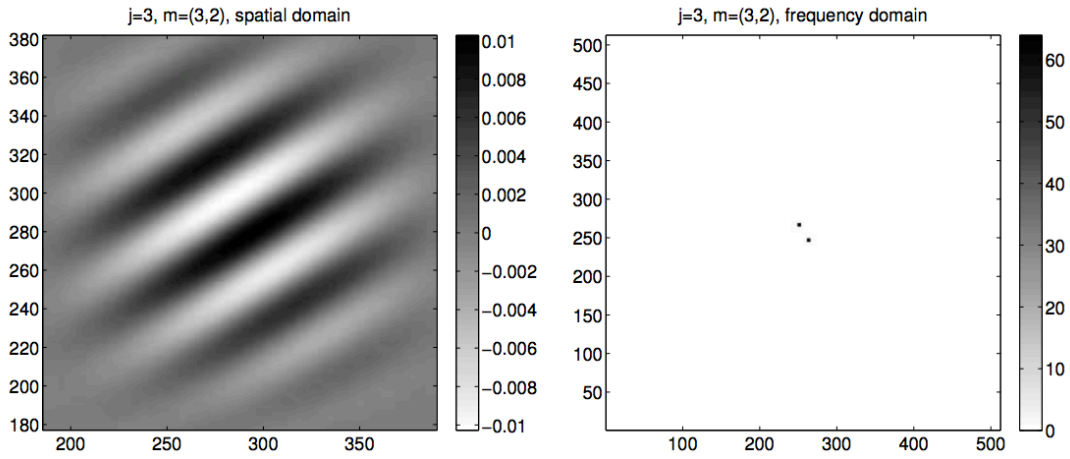

Figure 3-3 Wave Atom Filter, j=3, m=(3,2) for spatial domain and frequency domain.

3.2.5 Discrete Wavelet Transform

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 (Lei et al., 2013). 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 LHj; HLj, and HHj; 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 Figure 3-4.

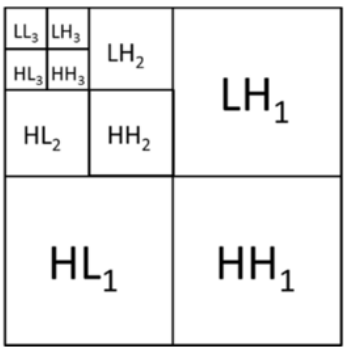

Figure 3-4 Level 3 Wavelet Decomposition (Lei et al., 2013).

Haar wavelet filter is used to perform Single-level wavelet transform as Figure 3-5. The resized input image (64x64 pixels) is used to obtain the approximation coefficients matrix using (4):

$$w(a,b_1,b_2)=\frac{1}{a}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} I(x,y)\psi(\frac{x-b_1}{a},\frac{y-b_2}{a})dxdy \qquad (4)$$

$$where\ \ \psi_H(x,y)=\begin{Bmatrix} 1{,}0\le x<\frac{1}{2},0\le y<1; \\ -1,\frac{1}{2}\le x<1,-1<y\le 0; \\ 0,else \end{Bmatrix}$$

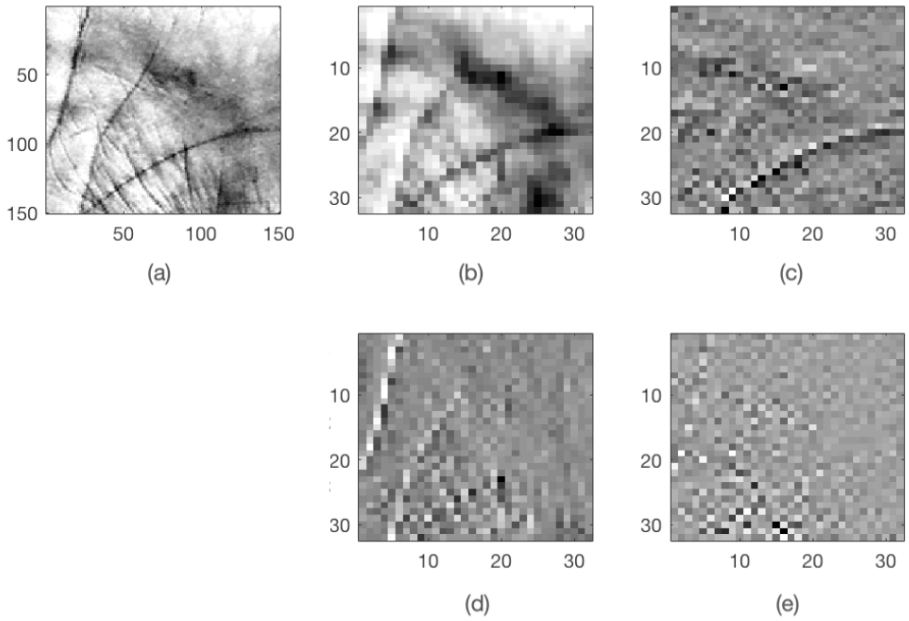

Figure 3-5 Wavelet Filters, (a) Original Image, (b) LL, (c) LH, (d) HL, (e) HH.

3.2.6 Scale Invariant Feature Transform (SIFT)

Constant local feature arguments are extracted using the Scale Invariant Feature Transform (Charfi et al., 2016). As shown in Figure 3-6, 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 (3), (5):

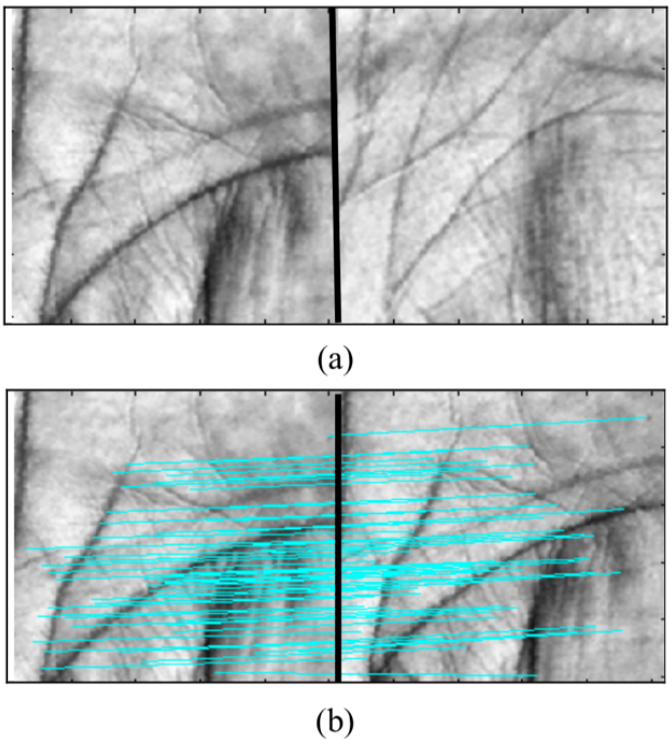

Figure 3-6 Sift Filter, two palm images. (a) Two palm images from different hands, and (b) two palm images from the same hand (Charfi et al., 2016).

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{3}$$

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

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \tag{4}$$

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

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma) * I(x, y)$$

$$= L(x, y, k\sigma) - L(x, y, \sigma) \tag{5}$$

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

As in (6), (7), 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. θ (x, y) is the orientation of the key point.

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \quad (6)$$

$$\theta(x,y) = \tan^{-1}\left(\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)}\right) \quad (7)$$

3.2.7 CNN-AlexNet

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 Figure 3-7. (Krizhevsky et al., 2012), a 1000-way SoftMax delivers distributions beyond 1000 class labels serving as the output of the mentioned connected layers.

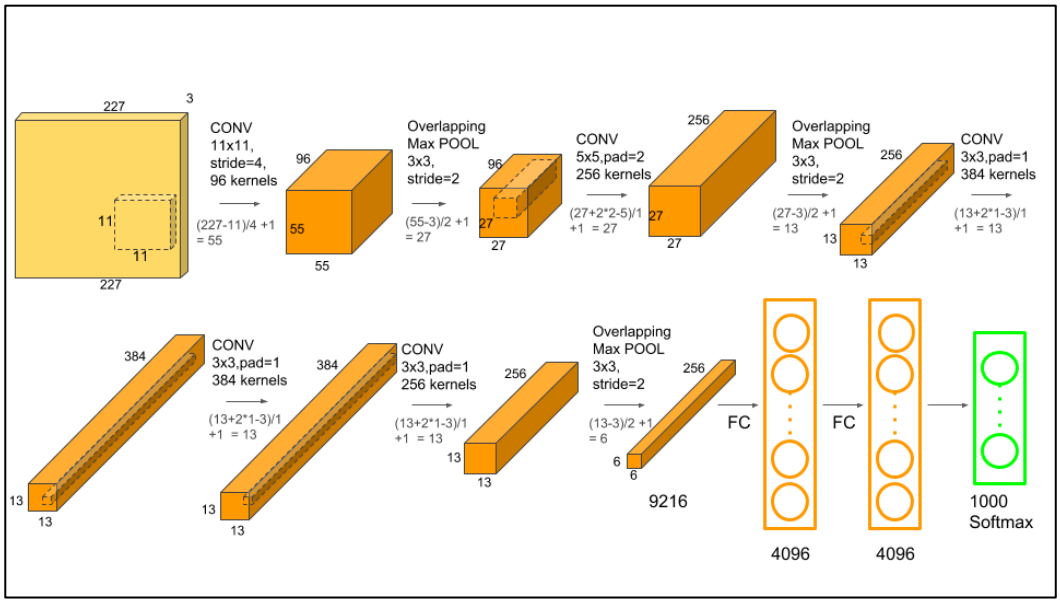

Figure 3-7 AlexNet convolutional neural network architecture (Krizhevsky et al., 2012)

The palmprint identification perspective involves discriminative LBP features (Wang et al., 2006). Upon the detection of the dominant pixel in the representation, the pattern code is matched against neighbors to produce the needed calculation as per (8).

$$LBP_{P,R} = \sum_{p=1}^{p} s(g_p - g_c)2^{p-1} \quad (8)$$

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

3.2.8 CNN-Squeezenet

The highlight of most researchers exploring CNN has been how the field can achieve create precision on machine vision datasets. Several models have been built to achieve different levels of accuracy. Iandola et al. (2016) has reviewed different CNN techniques to establish which model would require simpler parameter but attain a degree of precision that is similar to some other popular techniques. From his analysis, he concluded Squeeze Net has minimal parameters but still achieves a performance level that is comparable to common models. Figure 3-8 and Figure 3-9 shows the organization of convolution filters in the Fire module.

Iandola et al. (2016) used three main approaches when evaluating the different CNN architectures:

Approach 1: Resizing the network 3x3 filters to 1x1 filters

This approach focused on decreasing the number of parameters resizing 3x3 filters to 1x1 filters. At the onset, this approach was confusing as one would assume that replacing 33x3 filters with 1x1 filters would provide less information and thereby, lower the accuracy of the model. Nevertheless, this was not the case. Naturally, 3x3 filters takes spatial data of pixels that are in close proximity. On the other hand, in 1x1 filters solely concentrates on one pixel and extracts the connection of channels within the pixel without focusing the adjacent one.

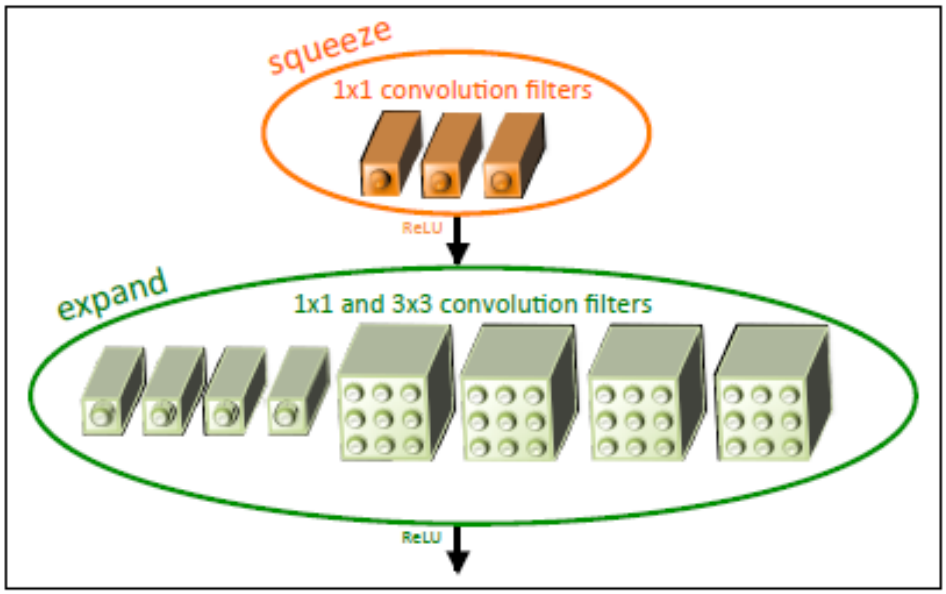

Figure 3-8 Microarchitectural view: Organization of convolution filters in the Fire module. (Iandola et al., 2016)

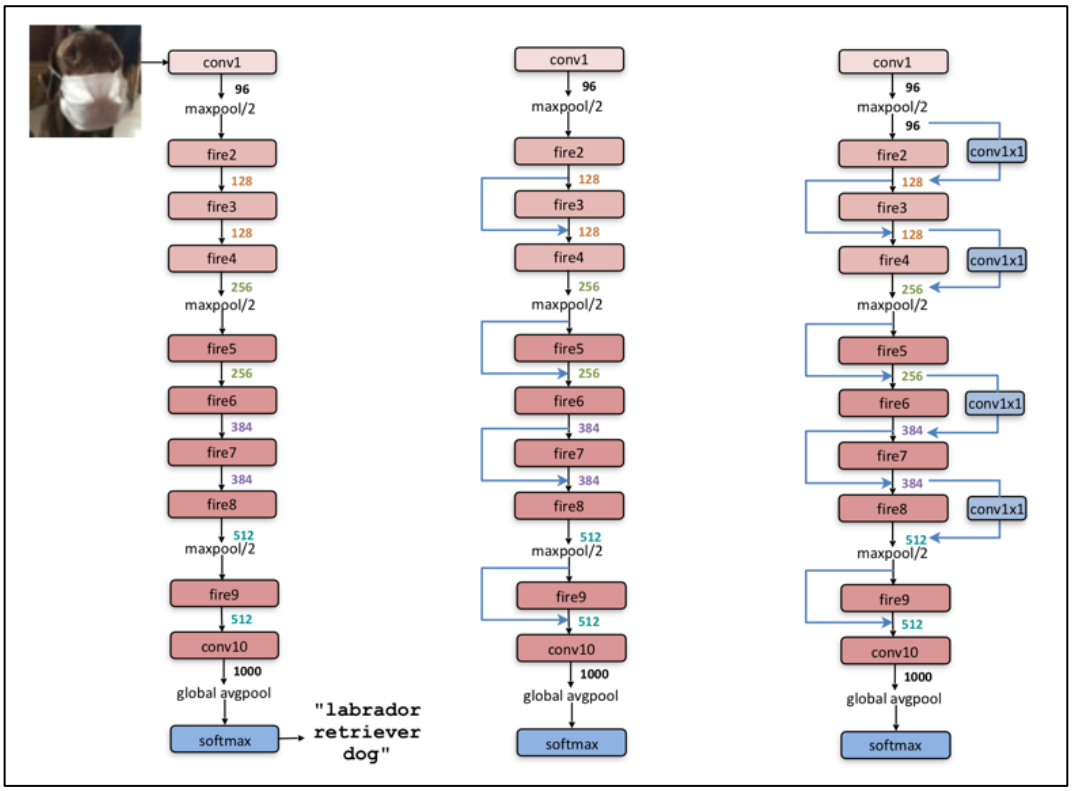

Figure 3-9 Microarchitectural view of Squeezenet architecture (Iandola et al. 2016)

Approach 2: Minimizing the number of inputs in the residual 3x3 filters

The approach focuses decreasing the parameter through cropping filters. This methodical strategy is actualized through feeding "squeeze" stratums into "expand" layers. Squeeze layers consist of 1x1 filters while "expand" layers have both 3x3 filters and 1x1 filters. Therefore, aggregate number of parameters are minimized through the lowering the "squeeze" stratum filters. This process is referred to as the "fire module" is used as the foundational block of creating the Squeezenet design.

Approach 3: Downsample the end part of the network to create large feature maps in the convolution strata

Iandola et al. (2016) propose that reducing the stride with the later convolution strata results in larger feature map, which enhances the accuracy of recognition. The manner in which the activation maps occur at the end of the network differentiates this design from well-known networks such as the VGG which is characterized with feature maps that shrink as one gets to end parts of the network.

The building block of Squeezenet is referred to as fire module. The fire module has two strata: squeeze layer and an expand layer. A Squeezenet has a pile of fire modules that are accompanied with minute pooling layers. Notably, both squeeze layer and the expand layer parallel activation map size. Nevertheless, the depth decreases in squeeze layer, while it surges the expand layer. Meanwhile, the neural models have a bottleneck layer and a trait of increasing. Besides, they have a tendency to generate a pattern of a surging depth, while minimizing activation map size with the intention of obtaining an elevated level abstract.

As illustrated in the chart above, the squeeze segment only has 1x1 filters. The latter implies that it operates as completely-attached layer dealing with feature points on the

same area. The spatial abstract cannot be derived from the system. Notably, it helps in decreasing the depth of feature map, which simplifies and increases the speed of computation in the expand layer.

Reducing depth enables to achieve an effect when the number of computations performed by 3x3 filters in the expand layer decreases. The speed increases due to the fact that 3x3 filters required only as much computation as 1x1 filters need. Logically speaking, too much squeezing will decrease the exchange of data, while limited number of 3x3 filters is bound to inevitable limit space resolution. As in Table 3-1, the SqueezeNet architecture enabled a 50X decrease in the size of the model when compared to AlexNet; it is worth noting that top-1 as well as top-5 accuracy of AlexNet was exceeded.

Table 3-1 Comparing SqueezeNet to model compression approaches

CNN architecture	Compression Approach	Data Type	Original → Compressed Model Size	Reduction in Model Size vs. AlexNet	Top-1 ImageNet Accuracy	Top-5 ImageNet Accuracy
AlexNet	None (baseline)	32 bit	240MB	1x	57.2%	80.3%
AlexNet	SVD (Denton et al., 2014)	32 bit	240MB → 48MB	5x	56.0%	79.4%
AlexNet	Network Pruning (Han et al., 2015b)	32 bit	240MB → 27MB	9x	57.2%	80.3%
AlexNet	Deep Compression (Han et al., 2015a)	5-8 bit	240MB → 6.9MB	35x	57.2%	80.3%
SqueezeNet (ours)	None	32 bit	4.8MB	**50x**	57.5%	80.3%
SqueezeNet (ours)	Deep Compression	8 bit	4.8MB → 0.66MB	**363x**	57.5%	80.3%
SqueezeNet (ours)	Deep Compression	6 bit	4.8MB → 0.47MB	**510x**	57.5%	80.3%

3.2.9 Cross Validation

Originally introduced by Vapnik (1995), 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 (Charfi et al., 2016).

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

$$\min_{w,\xi_n} \quad = \frac{1}{2} w^T w + C \sum_{n=1}^{N} \xi_n \tag{9}$$

$$where \quad w^T x_n \, t_n \geq 1 - \xi_n \qquad \forall_n$$

$$and \quad \xi_n \geq 0 \ \forall_n$$

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

The unconstrained optimization problem in (9) that is recognized as the primal form problem of L1-SVM:

$$min_w \quad = \frac{1}{2} \, w^T \, w + C \sum_{n=1}^{N} \max(1 - w^T \, x_n \, t_n \, , 0) \tag{10}$$

Meanwhile L1-SVM is not differentiable, the L2-SVM is used to minimizes the squared hinge loss as in (11):

$$min_w \quad = \frac{1}{2} \, w^T \, w + C \sum_{n=1}^{N} \max(1 - w^T \, x_n \, t_n \, , 0)^2 \tag{11}$$

The class label of a test data x is:

$$arg_t \max(w^T x) \, t \tag{12}$$

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

$$a_k(x) = w^T x \tag{13}$$

the forecast class is

$$arg_k \max \ a_k(x) \tag{14}$$

3.2.10 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 discusses settings involve the match score level fusion as representative of the principal fusion level in biometrics (Ross et al., 2006).

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 (Ross et al., 2006). 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).

3.3 Gender Recognition

This part focuses on the methods used in gender recognition in biometric identification from feature extraction to score fusion in the last level.

3.3.1 Features Extraction

Feature extraction discloses the nature of shape in a given pattern and thereby, simplifying the process of sorting the pattern using a formal method. It normally entails minimizing the number of random variables being analyzed until one is left with the main variables. In pattern recognition and in image processing, feature extraction is a particular type of reducing dimensionality. Its key objective is to mine the pertinent material from the original sample and present in a way that makes it easy for image processing and pattern classification (Kumar and Bhatia, 2014).

3.3.2 Discrete Wavelet Transform (DWT)

The Continuous Wavelet Transform (CWT) changes an uninterrupted signal into one that has two uninterrupted variables: translation and scale. The subsequent signal after the modifications made by CWT has a less complicated interpretation and beneficial when it comes to time-frequency analysis. The continuous wavelet transform of continuous function, x(t) relative to real-valued wavelet, ψ(t) is described by:

$$W(a,b) = \frac{1}{\sqrt{a}} \int \psi\left(\frac{t-b}{a}\right) s(t)dt \tag{15}$$

where ψ is the analyzing wavelet, a represents a time dilation, b a time translation, and the bar stands for complex conjugate.

DWT is a robust signal processing tool. Although DWT has some similarities with CWT, the difference is more apparent when it comes to scale and position values: the former scales and position values apply powers of two.

The values of s and t are: $s=2^j$, $\tau= k*2^j$ and $(j, k) \in Z^2$ as shown in (16):

$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{t - k * 2^j}{2^j}\right) \tag{16}$$

The focus in DWT as well as inverse DWT is disintegration and restoration. The disintegration and restoration are achieved through LPF and HPF. The effect of wavelet disintegration is a structured decomposition that occurs in tiers. The level of disintegration is selected as per the preferred cutoff frequency.

Figure 3-10 illustrates DWT undergoing a process of LPF and HPF (Haddadi et al., 2014).

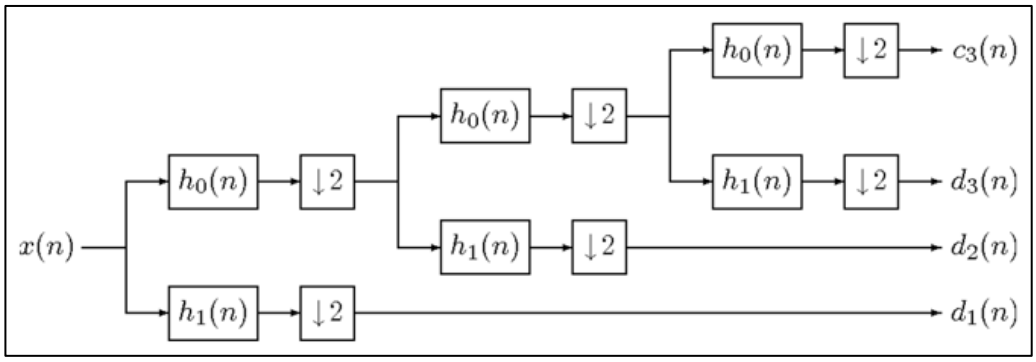

Figure 3-10 A three-level forward DWT via a two-channel iterative filter bank (Haddadi et al., 2014)

3.3.3The Discrete Haar Wavelet Transform

The original Haar definition is as follows:

$$haar(0,t) = 1, for\ t \in [0,1);$$

$$and \quad haar(1,t) = \begin{cases} 1, for\ t \in \left[0, \frac{1}{2}\right), \\ -1, for\ t \in \left[\frac{1}{2}, 1\right) \end{cases} \tag{17}$$

One of the key characteristics of the Haar functions other than the Haar (0, t), the i^{th} Haar function can occur through the restriction of the $(j-1)^{th}$ function to be fifty percent of the interval where it is different from zero, through multiplying it with p2 and scaling over the interval [0, 1]. These traits make Haar function makes share a degree of link with the wavelet concept and can be used to define wavelet. In this context, two Haar functions represent universal functions, while the remaining functions represent the local functions. As such, the odd rectangular pulse pair, a Haar function in this setting, denotes the oldest and the most basic wavelet. The purpose of using DWT is to get data that is highly differentiated by offering distinct solution at separate segments of the time-frequency domain. Through wavelet transforms, one can create unequal segments of time-frequency

plane in line with the time-spectral aspects of the signal. The wavelet technique has robust links to the traditional standard basis of the Haar functions. Scaling as well as shrinking a standard wavelet produces classical Haar functions.

Let ψ: R → R, the Harr wavelet function is defined by the formula:

$$\psi(t) = \begin{cases} 1, for\ t\ \in \left[0, \frac{1}{2}\right), \\ -1, for\ t\ \in \left[\frac{1}{2}, 1\right), \\ 0, otherwise. \end{cases} \tag{18}$$

for any Haar function (except function haar (0, t))

From basis (18) may be generated by means of the formulas:

$$\psi_i^j(t) = \sqrt{2^j}\,\psi\left(2^j t - i\right),$$

$$i = 0,1,\dots,2^j - 1.\ and\ j = 0,1,\dots,log_2\ N - 1 \tag{19}$$

The constant $\sqrt{2^j}$ is chosen so that:

$$\text{the scalar produc} < \psi_i^j, \psi_i^j \geq 1, \psi_i^j(t)\ \in\ L^2(R)$$

Let Φ: R → R, the Harr scaling function is defined by the formula:

$$\Phi(t) = \begin{cases} 1, & for\ t\ \in [0,1), \\ 0, & for\ t\ \notin [0,1). \end{cases} \tag{20}$$

Similarly, to the properties of the wavelet function, for scaling function one can define the family of functions:

$$\Phi_i^j(t) = \sqrt{2^j}\ \Phi\left(2^j t - i\right)$$

$$i = 0,1,\dots,2^j - 1.\ and\ j = 0,1,\dots,log_2\ N - 1 \tag{21}$$

The constant $\sqrt{2^j}$ is chosen so that:

$$\text{the scalar product}\ <\ \Phi_i^j, \Phi_i^j \geq 1, \Phi_i^j(t)\ \in\ L^2(R)$$

In the two-dimensional case, three sets of detail coefficients residing in the horizontal, vertical, and diagonal directions. The subbands LH_j, HL_j, and HH_j, j = 1, 2... J are the detail coefficients, as noted above, where j is the scale and J denotes the largest or coarsest scale in the decomposition. Figure 3-11 shows a representation of the multilevel wavelet decomposition at level 3.

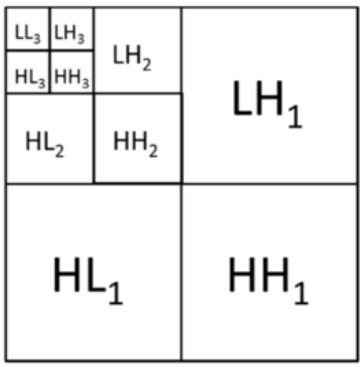

Figure 3-11 Level 3 Wavelet Decomposition (Williams and Li, 2016)

3.3.4 Squeeze Net

Squeezenet and discrete wavelet transform employ such feature extraction instruments. 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. As mentioned in Section 3.2.8, there are three key techniques used in Squeezenet to design CNN systems (Iandola et al., 2016). 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.

3.3.5 Cross Validation

The suggested system uses SVM for classifying the different genders on the basis of train-test split approach that is one of techniques for cross-validation. The predictive patterns are computed through the division of the basic model to create a training set to train the system, and set test for evaluation.

SVM was previously proposed by Vapnik (1995). The support vector machine is among the maximum margin classifiers and based on the Structural Risk Minimization. SVM plan input silhouette to a top dimensional domain where the uppermost separating hyperplane is gathered. Linear support vector machine is primarily pronounced for binary classification.

Take for instance that the training data set as well as the labels (x_n, y_n), n=1, ..., N, $x_n \in$ $\mathbb{R}^p$, $t_n \in \{-1, +1\}$, SVMs learning has the future controlled optimization:

$$\begin{aligned} \min_{w,\xi_n} \quad & = \frac{1}{2} w^T w + C \sum_{n=1}^{N} \xi_n \\ s.t. \quad & w^T x_n \, t_n \geq 1 - \xi_n \qquad \forall_n \\ & \xi_n \geq 0 \qquad \forall_n \end{aligned} \tag{22}$$

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

The unconstrained optimization problem as in (23) which is recognized as the primal form problem of L1-SVM:

$$min_w \quad = \frac{1}{2} w^T w + C \sum_{n=1}^{N} \max(1 - w^T x_n \, t_n \, , 0) \tag{23}$$

Meanwhile L1-SVM is not differentiable, the L2-SVM is used which minimizes the squared hinge loss as in (24):

$$min_w \ = \frac{1}{2} w^T w + C \sum_{n=1}^{N} \max(1 - w^T x_n \ t_n \ , 0)^2 \tag{24}$$

To expect the class label of a test data x:

$$arg_t \max(w^T x) \, t \tag{25}$$

To extend SVMs for multiclass problem, one-vs-rest approach is used (Tank, 2013). Representing the output of the *k-th* SVM as in (26)

$$a_k(x) = \ W^T x \tag{26}$$

the forecast class is

$$arg_k \max a_k \ (x) \tag{27}$$

3.3.6 Score Fusion

A key question that should be addressed when handling information on combining systems is the nature of data that needs to be merged using the fusion module. There are several approaches that have been devised to help in the latter at different levels including the sensor level, feature level, rank level, as well as the decision level. In line with our recommended system, we will analyze combining systems considering it is the principal level of fusion at the match score level (Ross et al., 2006).

At the match score level, the relationship between input and prototype biometric feature silhouettes. Synthesis at the core level entails the successful combination of various match scores output to obtain a biometric recognition decision (Ross et al., 2006). Some of the strategies used in consolidating the different biometric matchers include Majority Voting, and Weighted Majority Voting. Our proposed system recommends the use of Mode Voting Technique (MVT) when consolidating information at the decision level.

The method utilizes the standard class label values that are retrieved from the predicted label array obtained through the SVM discriminate classifier. MVT is used to establish the common non-repeated values in the predict label array X for the purpose of biometric identification.

$$Z = \text{mode}\ (X_{k.i}) \tag{28}$$

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.

CHAPTER 4. **Dense Hand-CNN Palm Identification System**

4.1 Introduction

The current chapter 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 experiments on IITD and CASIA databases have presented the efficiency of the proposed approach.

4.2 Block Diagrams

Two systems are used in this approach, the block diagram of the first system is shown in Figure 4-1. Whereas the second system shown in Figure 4-2. The segmented palmprint (ROI) palmprint of CASIA (2010) and IITD databases (2007) are used for the two systems.

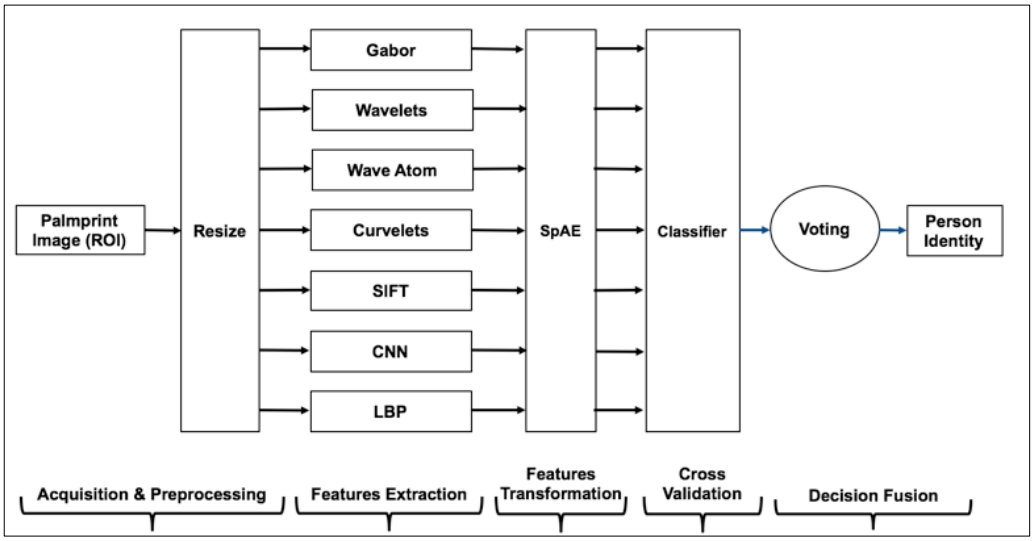

Figure 4-1 System 1: Human Identification using texture-based descriptors for Palmprint Images (Elgallad et al., 2017)

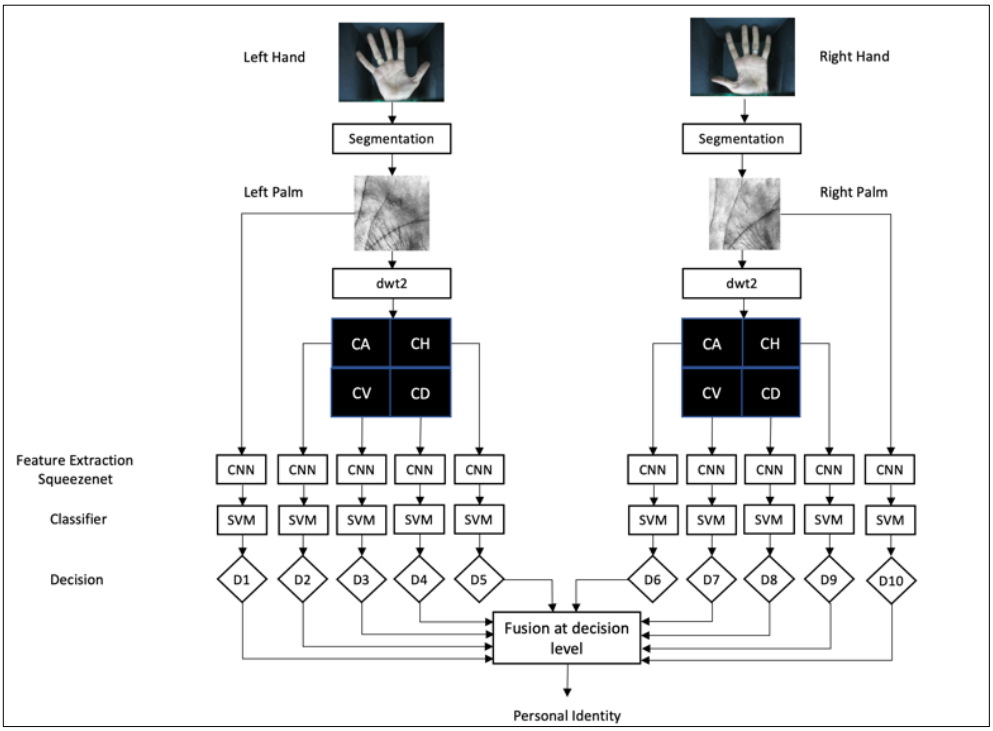

Figure 4-2 System 2: Dense Hand-CNN: CNN architecture based on later fusion of neural and wavelet features.

4.3 Databases

The IIT Delhi palmprint image database consists of the hand images collected from the students and staff at IIT Delhi, New Delhi, India. This database has been acquired in the IIT Delhi campus during July 2006 - Jun 2007 using a simple and touchless imaging setup. All the images are collected in the indoor environment and employ circular fluorescent illumination around the camera lens. The currently available database is from 230 users, about 10 images for each, left and right with size 150×150 pixels in grayscale. All the subjects in the database are in the age group 12-57 years. About ten images from each subject, from each of the left and right hand for the region of interest (ROI) of the images.

CASIA database contains 240 subjects, around 10 images for each subject classified left and right with size 192×192 pixels. All palmprint images are 8 bit gray-level JPEG

files, taken by self-developed palmprint recognition device in Chinese Academy of Sciences, Institute of Automation (CASIA).

The images are resized to 227×227 pixel and converted to RGB to extract features for CNN.

4.4 Preprocessing

In MATLAB 2018a and later, the function augmentedImageDatastore is used for efficient preprocessing of images for deep learning including image resizing. It generates batches of training, validation, test, and prediction data, with optional preprocessing such as resizing, rotation, and reflection. Resize images to make them compatible with the input size of your deep learning network.

Augment training image data with randomized preprocessing operations to help prevent the network from overfitting and memorizing the exact details of the training images.

The output of the function is the training and testing set stored in datastore image.

4.5 Features Extraction

In the initial setting, the system 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.

The number of Gabor features is 120 features which in turn pass through sparse autoencoder to increase the number of features up to five multiples and output of 600 features is gained which achieved better results.

The coefficients of Wavelets matrix are arranged to a one-dimensional array of 1024 elements. This array in turn passes through sparse autoencoder with a hidden layer of two-times of its input, and output of 2048 features is gained.

Regarding waveatom transform, the output array is extracted from the input image which is resized to 64×64 at the scale 3 and arranged to get the features array of 1120 elements. This array in turn passes through sparse autoencoder with a hidden layer of two-times of its input, and output of 2240 features is gained.

Fast Discrete Curvelet Transform is used to obtain the cell array of 121 elements which in turn passes through sparse autoencoder with a hidden layer of eight-times of its input, and output of 968 features is gained.

On the other hand, Scale Invariant Feature Transform is used to extract constant local feature points (1024 features) from 150×150 images. The sparse auto-encoder is used with a hidden layer of two-times of its input, and output of 2048 features is gained.

AlexNet convolutional neural network is also used to extract the features of the training and test images (4096 features) from the last fully connected layer. The features pass through sparse autoencoder with a hidden layer of two-times of its input, and output of 8192 features is gained.

Local Binary Pattern (LBP) returns the histogram of 64×64 input images. The resulting features array of 256 is passed through sparse autoencoder with a hidden layer of four-times of its input, and output of 1024 features is gained which achieved the best result.

The number of extracted features from each descriptor are illustrated in Table 4-1.

Table 4-1 Number of features for the first system

Descriptor	Features	SpAE	Features with SpAE
Gabor	120	5	600
Wavelets	1024	2	2048
Waveatoms	1120	2	2240
Curvelet	121	8	968
SIFT	1024	2	2048
AlexNet	4096	2	8192
LBP	256	4	1024
All Features	7761		17120

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.

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 number of extracted features in second system are illustrated in Table 4-2.

Table 4-2 Number of features for the second system

Database	IITD	CASIA
Input image size	150x150 grayscale	192x192 grayscale
dwt2 output image size	75x75 grayscale	96x96 grayscale
Image preprocessing	227x227 RGB	227x227 RGB
Squeezenet fire9-concat features	**100352**	**100352**

4.6 Cross Validation

SVM and Softmax are used as classifiers in the first system, while SVM used only in the second system. Both of them are based on cross-validation technique. In order to calculate predictive models by dividing the original model into a training set to train the structure, and a test set for evaluation.

4.7 Score Fusion (MVT)

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. Figure 4-3 explain 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.

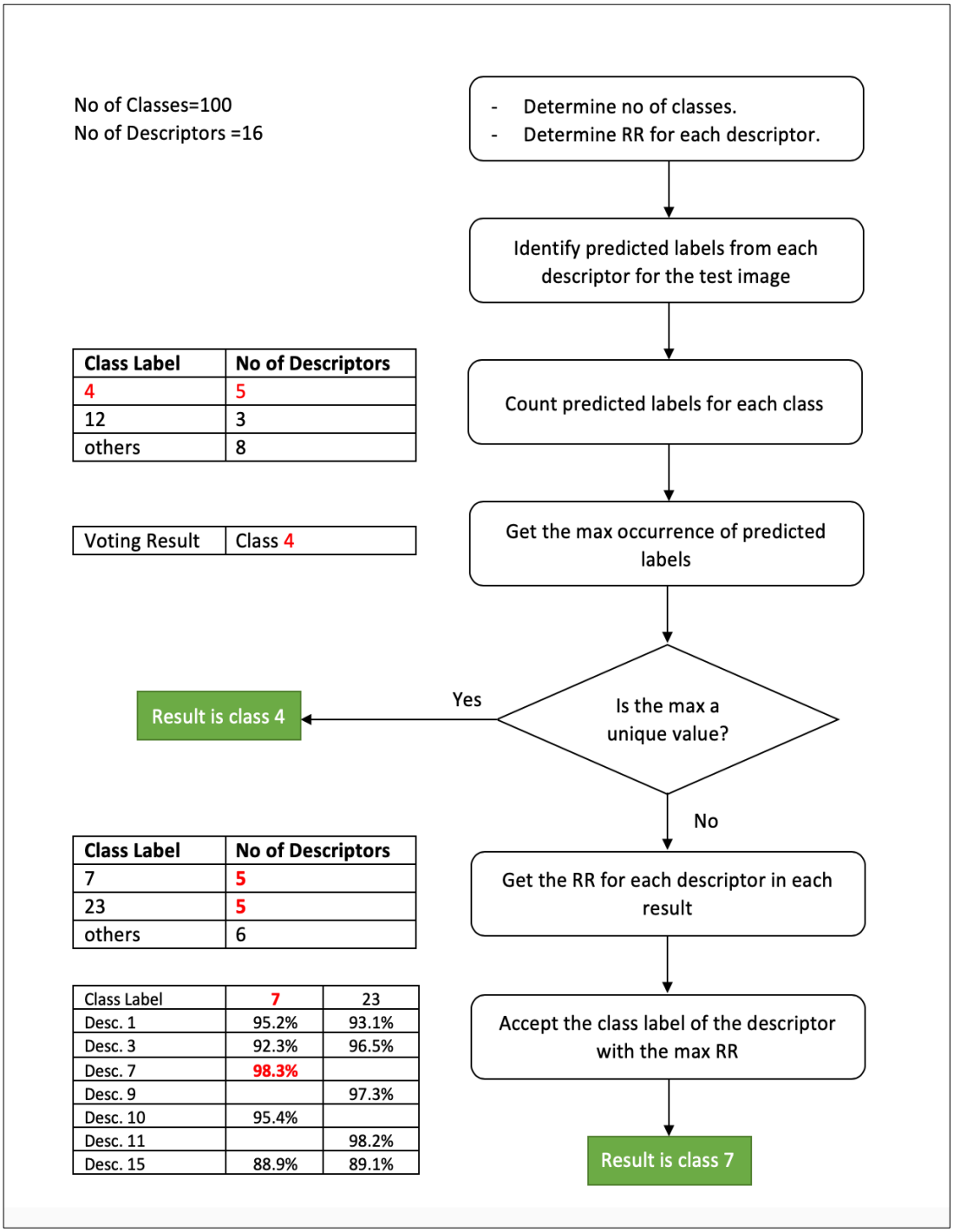

Figure 4-3 Flowchart of Mode Voting Technique (MVT)

CHAPTER 5. **CWNN-Net Gender Recognition system**

5.1 Introduction

This chapter focuses on gender recognition in biometric identification and explains how it can be used to enhance the accuracy of biometric identification processes.

The key contribution is the application of mode voting method for accurate identification of a person gender at the fusion decision level. The experiments on 11k database (2017) , and CASIA database (2010) databases have presented the efficiency of the proposed approach.

5.2 Block Diagram

Figure 5-1 indicates the CWNN-Net system which this system. The system works together with Squeezenet to extract features while SVM remains as the discriminative classifier. The final interpretation is made using the mode voting method.

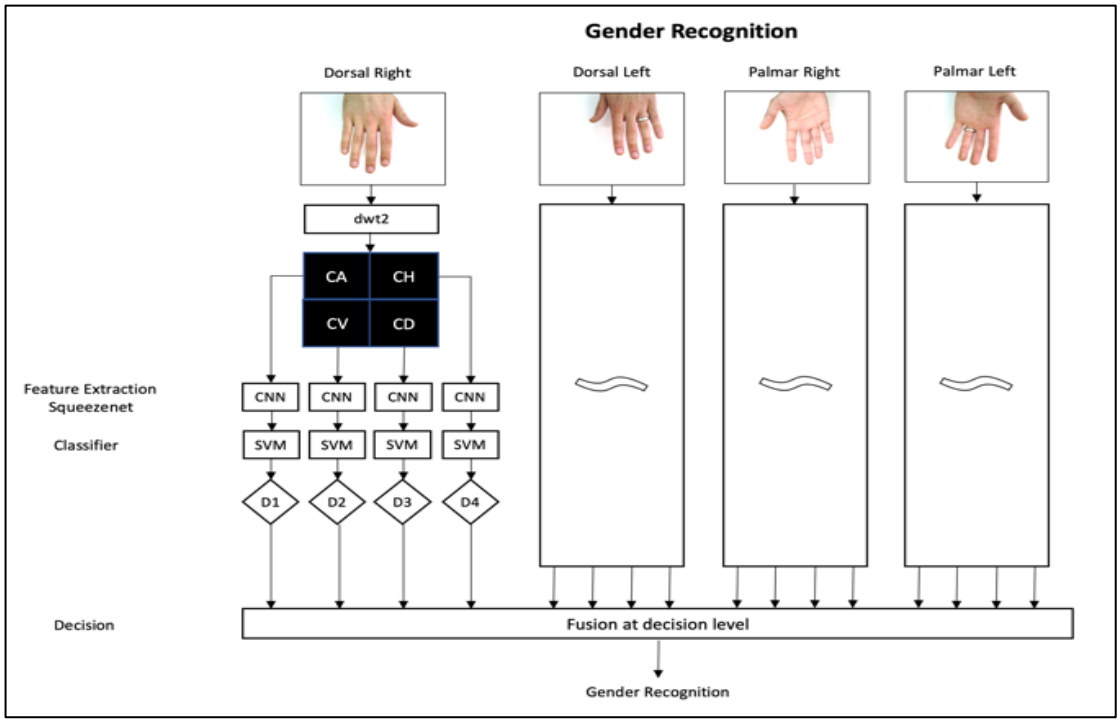

Figure 5-1 CWNN-Net: A New Convolution Wavelet Neural Network for Gender Classification using Palm print

5.3 Database

There are two databases that are relied upon in this system are 11k database (2017) , and CASIA database (2010), each containing sufficient hand images and palmprint images respectively. The images were changed to RGB and minimized to 227×227 pixels to enhance CNN feature extraction.

The 11k database had hand images taken from individuals whose photos were stored in jpg formats. The hand images displayed the palm/dorsal, and left and right images of the individuals. In total, images of 190 subjects, aged between 18 and 75, were retrieved from the database.

As for the CASIA database, it had 5,503 palmprint images belonging to 312 persons and representing both the left and right palms. The images of JPEG format each 8 bit gray-level. Slight adjustments were made to the images to facilitate CNN feature extraction. Specifically, they were modified to RGB and minimized to 227×227 pixels.

5.4 Preprocessing

As stated before, in MATLAB 2018a and later, the function augmentedImageDatastore is used for efficient preprocessing of images for deep learning including image resizing. It generates batches of training, validation, test, and prediction data, with optional preprocessing such as resizing, rotation, and reflection. Resize images to make them compatible with the input size of your deep learning network.

Augment training image data with randomized preprocessing operations to help prevent the network from overfitting and memorizing the exact details of the training images.

The output of the function is the training and testing set stored in datastore image.

5.5 Features Extraction

The main objective of the system is to gain high recognition rate in a speedy processing time for each image.

The using of Squeezenet because it has 50x fewer parameters compared with AlexNet, and has a model compression technique which can compress Squeezenet to less than 0.5MB.

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 number of extracted features in the system are illustrated in Table 5-1.

Table 5-1 Number of Features for the system

Database	11k	CASIA
Input image size	1600x1200 RGB	192x192 grayscale
dwt2 output image size	800x600 RGB	96x96 grayscale
Image preprocessing	227x227 RGB	227x227 RGB
Squeezenet fire9-concat features	**100352**	**100352**

5.6 Cross Validation

SVM is used as classifier in the system. It is based on cross-validation technique. In order to calculate predictive models by dividing the original model into a training set to train the structure, and a test set for evaluation.

FITCECOC function can fit a multiclass model for Support Vector Machine or other classifiers.

5.7 Score Fusion (MVT)

Our proposed system recommends the use of Mode Voting Technique (MVT) when consolidating information at the decision level.

The method utilizes the standard class label values that are retrieved from the predicted label array obtained through the SVM discriminate classifier. MVT is used to establish the common non-repeated values in the predict label array for the purpose of biometric identification.

CHAPTER 6. **Experimental Results**

6.1 Introduction

In this chapter the results of all systems are presented and discussed. The first system of Dense Hand-CNN Palm Identification followed by CWNN-Net Gender Recognition system, and to compare their results with the related work to view how these systems reach its objectives.

6.2 Experimental Results of Dense Hand-CNN Palm Identification System

In the two systems of palmprint, two sets of databases are which are CASIA Palmprint Database (2010) and Delhi Touchless Palmprint Database IIT version 1.0 (2007). CASIA Palmprint database contains 2400 palmprint images for 240 subjects from left and right palms in size 192×192 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 150×150 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 227×227 pixels. In both systems, SVM classifier is used.

The following results answer the first question of the research which stated that "*How can the representation of data through the descriptors in the feature extraction step in building the classification model affect the predictive and computationally of the performance of palm identification system?*".

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 150×150 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 64×64 input images. The output array has 256 features. The recognition rate is 66.81% with IITD, and 82.54% with CASIA.

Table 6-1 illustrates the results of the used seven descriptors for both databases. These results answer the second question of the research which stated that "*How can the fusion at decision level achieve an interesting classification accuracy and improve the recognition rate of the identification system?*".

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% in SVM and 96.09 with Softmax in IITD database, with processing time for each image 1.88 and 0.99 sec for both classifiers. With CASIA database, the recognition rate is 100%, with 2.36 sec processing time for each image.

Table 6-1 Results of the first system without SpAE

Descriptor	Features	IITD		CASIA
		SVM	Softmax	SVM
Gabor	120	70.29	71.74	87.46
Wavelets	1024	80.94	83.04	92.46
Waveatoms	1120	77.32	76.30	92.39
Curvelet	121	79.57	84.75	91.16
SIFT	1024	96.96	91.96	98.91
AlexNet	4096	93.91	68.26	98.55
LBP	256	66.81	70.22	82.54
Voting		99.57	96.09	100
Time / Sec		1.88	0.99	2.36

Table 6-2 illustrates the system with SpAE, using SVM classifier, SpAE increases the results slightly for each descriptor in most cases and the feature fusion with SpAE increases the recognition rate by more than 50%. The voting result is the same for SVM.

Using Softmax classifier, the results values vary between high and low. The feature fusion of all descriptors increases slightly and the voting result changes by around 1%.

Table 6-2 Results of the first system with SpAE

Descriptor	New Features	IITD	
		SVM	Sofmax
Gabor	600	75.07	76.52
Wavelets	2048	87.32	82.39
Waveatoms	2240	76.30	68.04
Curvelet	968	89.49	85.00
SIFT	2048	91.16	74.78
AlexNet	8192	84.28	64.35
LBP	1024	71.23	69.35
Voting		100	97.39
Time / Sec		18.39	22.46

For the result in the second system, 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 6-3 illustrates the detailed result of the second system.

Table 6-3 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			

6.3 Experimental Results of CWNN-Net Gender Recognition system

As apparent in the methodology section, two databases (11k Database and CASIA Database) were used to obtain the relevant data records.

SVM was applied to discriminately sort the features. The features were retrieved from single-level 2D DWT through the application of Haar wavelet filter. Through this process, DWT provided an approximation coefficient matrix cA as well as detail coefficients matrices cH, cV, and cD (horizontal, vertical, and diagonal, respectively).

Two experiments are carried out over 11k database, one on all dataset, and the other on the dataset after extracting hand images with accessories. The predicted label arrays obtained from SVM as a classifier for the obtained features are fused using mode voting technique.

Table 6-4 and Table 6-5 summarize the results obtained for both experiments. These results answer the third question of the research which stated that "*What is the role of CNN in the improvement of the performance of the identification system?*".

The recognition rate for all dorsal dataset ranged from 97.57% to 99.71%, with mode voting technique on the resulting 8 predicted label arrays, the recognition rate reached 99.86% with processing time 0.27 sec for each image.

For all palmer dataset ranged from 96.86% to 98.57%, with mode voting technique on the resulting 8 predicted label arrays, the recognition rate reached 100% with processing time 0.28 sec for each image.

By Applying mode voting technique on the resulting 16 predicted label arrays for previous both dorsal and palmar images, the recognition rate reached 100% with processing time 0.55 sec for each image.

After excluding images with accessories, the recognition rate for dorsal dataset ranged from 97.27% to 99.64%, and for palmer dataset ranged from 97.27% to 99.27%. Also, after applying mode voting technique on the resulting 16 predicted label arrays, the recognition rate reached 100% with processing time 0.57 sec for each image.

For CASIA database, the experiment is carried out over the database which contains only palmprint images. Table 6-6 summarize the obtained results.

The recognition rate for palmar dataset ranged from 81% to 97.75%. After applying mode voting technique on the resulting 8 predicted label arrays (as the database only contain palmar side), the recognition rate reached 98% with processing time 0.25 sec for each image as in Table 6-6.

It is noted that, the high recognition rate is nearly obtained from the approximation coefficient matrix cA in both databases.

Table 6-4 Results of 11k database for all images

Descriptor	dwt2	dorsal L	dorsal R	palmar L	palmar R
squeezenet fire9-concat	cA	99.43	99.71	97.29	98.57
	cH	98.71	97.57	97.71	98.29
	cV	98.75	98.29	97.43	98.00
	cD	98.14	97.86	96.86	98.43
Voting		99.86		100	
		100.00			
Sys. Time/image in sec.		0.27		0.28	
		0.55			

Table 6-5 Results of 11k database excluding acc. images

Descriptor	dwt2	dorsal L	dorsal R	palmar L	palmar R
squeezenet fire9-concat	cA	98.73	99.64	99.27	98.73
	cH	98.00	97.82	98.00	97.27
	cV	98.18	98.00	98.00	98.00
	cD	97.45	97.27	97.27	97.82
Voting		100		100	
		100.00			
Sys. Time/image in sec.		0.57			

Table 6-6 Results of CASIA database

Descriptor	dwt2	palmar L	palmar R
squeezenet fire9-concat	cA	97.75	95.25
	cH	92.25	82.50
	cV	87.75	88.50
	cD	82.00	81.00
Voting		98.00	
Sys. Time/image in sec.		0.25	

CHAPTER 7. **Conclusion and Future Work**

7.1 Result Summary of Dense Hand-CNN Palm Identification System

The results of palmprint identification system 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 7-1 displays the performance of the proposed palmprint recognition systems vs. current systems in the related work that use IITD and CASIA databases.

Table 7-1 Performance comparison of proposed palmprint recognition system vs. existing systems using IITD and CASIA databases

Author	Features Extractor	Features Classifier	Database	RR (%)
Zhao et al., 2013	**SIFT**	**competitive code algorithm**	**IITD**	**Equal Error Rate = 0.49**
Varshney et al., 2014	**DWT - DCT**	**Euclidean Distance**	**IITD, PolyU**	**94.44, 95.65**
Charfi et al., 2014	**SIFT**	**Matching Score**	**IITD**	**Palmprint = 94.05** **Hand shape + Palmprint = 97.82**
Jaswal et al., 2015	**2D Gabor filter**	**Euclidean Distance**	**CASIA** **IIT Delhi**	**90.76** **91.4**
Misar and Gharpure, 2015	**Discrete Wavelet Transform**	**Neural Network**	**IITD**	**75.6**
Charfi et al., 2015	**SIFT and Gabor**	**Matching score**	**IITD**	**Palmprint = 91.08** **Hand shape + Fingers + Palmprint = 98.04**
Charfi et al., 2016	**SIFT sparse representation**	**SVM**	**IITD** **Bosphorus**	**IITD: Palmprint = 96.73** **Hand shape + Palmprint = 99.57** **Bosphorus: Palmprint = 94.95** **Hand shape + Palmprint = 97.61**
Afifi, 2017	**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**

A significant issue in this comparison table is the processing time that is explained in Table 7-2, 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 7-2 Processing time comparison

	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

As stated before, the second system achieved the main objective of it that is the reduction the processing time for each image and maintaining the high recognition rate obtained from the first system.

This reduction in processing time is due to the using of the Squeezenet that has 50x fewer parameters compared with AlexNet. Due to these advantages, it's used as a feature extractor in the second system.

The Cumulative Matching Characteristic identify how many times the system can identify the correct identity from the first round, second round, and so on, and cumulate the values to reach the estimated max value of recognition rate.

Figure 7-1 shows CMC Curves of system1, probability of detecting the correct identity within the top K ranks for the descriptors.

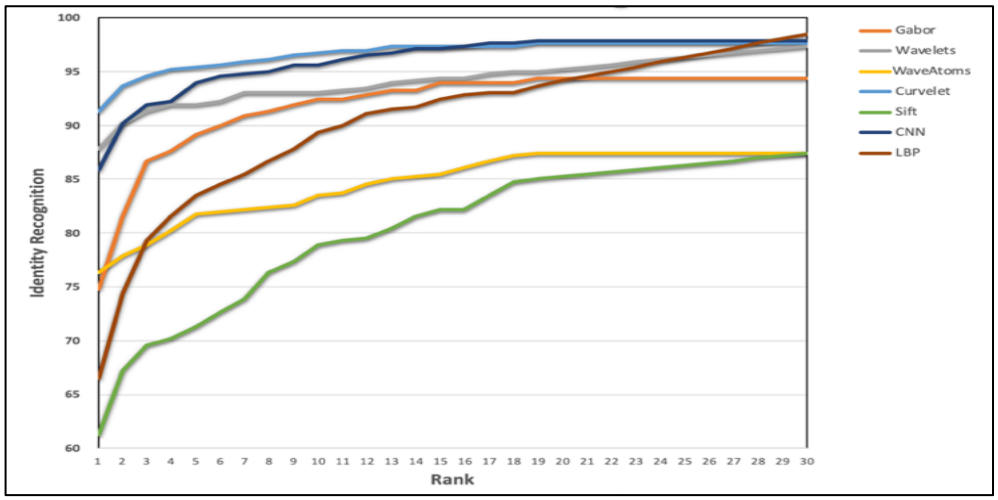

Figure 7-1 CMC Curves: Probability of detecting the correct identity within the top K ranks

7.2 Result Summary of CWNN-Net Gender Recognition system

From the results, it is clear that the fusion at decision level using the mode voting technique guarantees an excellent recognition rate regardless of low recognition rate of some datasets. The mode voting technique ranks top of the list of SVM classifiers used for each database.

Table 7-3 shows the performance comparison of proposed system CWNN using 11k and CASIA databases. It is clear that the performance is stable for both databases, the dorsal and palmar results in 11k database are nearly the same, so one of them can be used alone to reduce the system processing time. Also, the processing time for two databases is nearly the same for each image.

Table 7-3 Performance comparison of proposed system CWNN using 11k and CASIA databases

Database	11k			CASIA
Hand	Dorsal L, R	Palmar L, R	All	Palmar L, R
No of Images	5680	5396	11076	5502
Feature extraction	Squeezenet & dwt2			
Classifier	SVM			
No of predicted arrays	8	8	16	8
Result by voting	99.86	100	100	98
Sys. Time/image in sec.	0.27	0.28	0.55	0.25

Table 7-4 shows the performance comparison of the proposed systems CWNN vs. existing systems in the previous work that use 11k database.

Table 7-4 Performance comparison of proposed system CWNN vs. existing systems using 11k

11k	Afifi 2017		CWNN	
	Dorsal	Palmar	Dorsal	Palmar
All Images			99.86	100
no accessories	97.30	94.20	100	100

7.3 Conclusion

7.3.1 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. MVT is used to establish the common non-repeated values in the predict label array for the purpose of biometric identification and the Inference is made using mode voting approach.

7.3.2 Dense Hand-CNN

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 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

The selection of Squeezenet and DWT in the second system depend 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.

7.3.3 CWNN-Net

This work reports a CWNN-Net: a new convolution wavelet neural network for gender classification using palm print system based on the mode voting technique, and compares the performance of the system using two datasets. The novelty comes from using mode voting technique at decision level. Our experimental results show the efficiency of the suggested system. From the results, it is clear that the fusion at decision level using the mode voting technique guarantees an excellent recognition rate regardless of low recognition rate of some datasets. The mode voting technique ranks top of the list of SVM classifiers used for each database.

The performance comparison of proposed system CWNN using 11k and CASIA databases show that the performance is stable for both databases, the dorsal and palmar results in 11k database are nearly the same, so one of them can be used alone to reduce the system processing time. Also, the processing time for two databases is nearly the same for each image.

Using Squeezenet and DWT in this system shows a promising result due to the advantage of Squeezenet and wavelet decomposition.

7.4 Future Work

We believe that the mode voting technique can serve as a step towards the construction of more accurate gender recognition and biometric identification systems. In the future, we will adopt this deep learning method in real-time palmprint recognition system and develop a more intelligent machine learning algorithm for feature extraction in palmprint recognition.

References

[1] Afifi, M., (2017). 11K Hands: Gender recognition and biometric identification using a large dataset of hand images." arXiv preprint arXiv:1711.04322.

[2] Akbulut, Y., Şengür, A., & Ekici, S., (2017). Gender recognition from face images with deep learning. International Artificial Intelligence and Data Processing Symposium (IDAP), Malatya, pp. 1-4.

[3] Amayeh, G., Bebis, G., & Nicolescu, M., (2008). Gender classification from hand shape. In: Computer Vision and Pattern Recognition Workshops. CVPRW'08. IEEE Computer Society Conference on, IEEE, pp. 1-7.

[4] Jain, A., Dass, S., & Nandakumar, K., (2004). Soft Biometric Traits for Personal Recognition Systems, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 731–738.

[5] Ardakany, A.R., & Louis, S.J., (2012). Improving Gender Recognition Using Genetic Algorithms. In: Bui L.T., Ong Y.S., Hoai N.X., Ishibuchi H., Suganthan P.N. (eds.) Simulated Evolution and Learning. Lecture Notes in Computer Science, vol. 7673. Springer, Berlin, Heidelberg.

[6] Azzopardi, G., Greco, A., & Vento, M., (2016). Gender Recognition from Face Images Using a Fusion of SVM Classifiers. In: Campilho A., Karray F. (eds.) Image Analysis and Recognition. Lecture Notes in Computer Science, vol. 9730. Springer, Cham.

[7] Azzopardi, G., Greco, A., Saggese, A., & Vento, M., (2018). Fusion of domain-specific and trainable features for gender recognition from face images. IEEE Access, vol. PP, no. 99, pp. 1-1.

[8] CASIA Palmprint Database, http://biometrics.idealtest.org/

[9] Charfi, N., Trichili, H., Alimi, A.M., & Solaiman, B., (2014). Bimodal biometric system based on SIFT descriptors of hand images. IEEE International Conference on Systems, Man, and Cybernetics (SMC), San Diego, CA, 2014, pp. 4141-4145.

[10] Charfi, N., Trichili, H., Alimi, A.M., & Solaiman, B., (2015). Personal recognition system using hand modality based on local features. 11th International Conference on Information Assurance and Security (IAS), Marrakech, pp. 13-18.

[11] Charfi, N., Trichili, H., Alimi, A.M., & Solaiman, B., (2016). Bimodal biometric system for hand shape and palmprint recognition based on SIFT sparse representation. Multimedia Tools and Applications, pp.1-26.

[12] Chen, G.Y., Bui, T.D., & Krzyak, A., (2006). Palmprint classification using dual-tree complex wavelets, in: Proceeding of International Conference on Image Processing, pp. 2645–2648.

[13] Chen, J., & Moon, Y., (2008). Using SIFT features in palmprint authentication, 19th International Conference on Pattern Recognition, pp. 1–4.

[14] Cirne, M.V.M., & Pedrini, H., (2017). Gender recognition from face images using a geometric descriptor. IEEE International Conference on Systems, Man, and Cybernetics (SMC), Banff, AB, pp. 2006-2011.

[15] Collins, M., Miller, P., & Zhang, J., (2014). Gait Based Gender Recognition Using Sparse Spatio Temporal Features. In: Gurrin, C., Hopfgartner, F., Hurst, W., Johansen, H., Lee, H., & O'Connor, N., (eds.) Multimedia Modeling. MMM 2014. Lecture Notes in Computer Science, vol. 8326. Springer, Cham.

[16] Daugman, J. (2000). Combining Multiple Biometrics. Available: http://www.cl.cam.ac.uk/users/jgdlOOO/combine/combine.html.

[17] Demanet & Ying, (2007). Wave atoms and sparsity of oscillatory patterns," Elsevier, pp. 368-387.

[18] Dong, K., Feng, G., & Hu, D., (2005). Digital curvelet transform for palmprint recognition, Advances in Biometric Person Authentication, Springer Berlin Heidelberg, pp. 639-645.

[19] Elgallad, E.A., Charfi, N.A., Alimi, M., & Ouarda, W., (2017). Human identity recognition using sparse auto encoder for texture information representation in palmprint images based on voting technique. Sudan Conference on Computer Science and Information Technology (SCCSIT), Elnihood, Sudan, pp. 1-8.

[20] Fogel, I., & Sagi, D., (1989). Gabor filters as texture discriminator. Biol. Cybern. 61, 103–113.

[21] Font-Aragones, X., & Faundez-Zanuy, M., (2013). Hand-Based Gender Recognition Using Biometric Dispersion Matcher. In: Apolloni, B., Bassis, S., Esposito, A., & Morabito, F., (eds.) Neural Nets and Surroundings. Smart Innovation, Systems and Technologies, vol. 19. Springer, Berlin, Heidelberg.

[22] Golomb, B.A., & Lawrence, D.T., (1991). SEXNET: A neural network identifies sex from human faces. Advances in Neural Information Processing Systems, pp. 572–577.

[23] Haddad, Z., Beghdadi, A., Serir, A., & Mokraoui, A., (2009). A new fingerprint image compression based on wave atoms transform. IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), Ajman, pp. 89-94.

[24] Haddadi, R., Abdelmounim, E., El Hanine, M., & Belaguid, A., (2014). Discrete Wavelet Transform Based Algorithm for Recognition of QRS Complexes. World of Computer Science & Information Technology Journal, 4(9).

[25] Han, C.C., (2004). A hand-based personal authentication using a coarse-to-fine strategy, Image and Vision Computing 22 (11) 909–918.

[26] Hatipoglu, B., & Kose, C., (2017). A gender recognition system from facial images using SURF based BoW method. International Conference on Computer Science and Engineering (UBMK), Antalya, pp. 989-993.

[27] He, K., Zhang, X., Ren, S., & Sun, J., (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[28] Hu, J., Shen, L., & Sun, G., (2017). Squeeze-and-excitation networks. arXiv preprint arXiv:1709.01507, 7.

[29] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K.Q., (2017). Densely Connected Convolutional Networks. In CVPR (Vol. 1, No. 2, p. 3).

[30] Huang, Y.S., & Suen, C.Y., (1995). Method of Combining Multiple Experts for the Recognition of Unconstrained Handwritten Numerals. IEEE Transactions on Pattern Analysis and Machine Intelligence, 17(l):90-94.

[31] Iandola, F.N., Han, S., Moskewicz, M.W., Ashraf, K., Dally, W.J., & Keutzer, K., (2016). Squeezenet: Alexnet-level accuracy with 50x fewer parameters and< 0.5 mb model size. arXiv preprint arXiv:1602.07360.

[32] Iglesias, F.S., Buemi, M.E., Acevedo, D., & Jacobo-Berlles, J., (2014). Evaluation of Keypoint Descriptors for Gender Recognition. In: Bayro-Corrochano, E., & Hancock, E., (eds.) Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications. Lecture Notes in Computer Science, vol. 8827. Springer, Cham.

[33] IIT Delhi Touchless Palmprint Database version 1.0, http://web.iitd.ac.in/~ajaykr/Database_Palm.htm

[34] Jain, A. K., Prabhakar, S., & Chen, S. (1999c). Combining Multiple Matchers for a High Security Fingerprint Verification System. Pattern Recognition Letters, 20(11 -13): 1371-1379.

[35] Jaswal, G., Nath, R., & Kaul, A., (2015). Texture based palm Print recognition using 2-D Gabor filter and sub space approaches. International Conference on Signal Processing, Computing and Control (ISPCC), Waknaghat. pp. 344-349.

[36] Kekre, H., Tanuja, B., Sarode, K., & Tirodkar, A.A., (2012). 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.

[37] Krizhevsky, A., Sutskever, I., & Hinton, G.E., (2012). ImageNet classification with deep convolutional neural networks. In Advances in neural information processing systems, pp. 1097-1105.

[38] Kumar, G., & Bhatia, P.K., (2014). A detailed review of feature extraction in image processing systems. In Advanced Computing & Communication Technologies (ACCT), Fourth International Conference, pp. 5-12.

[39] Kuncheva, L. I. (2004). Combining Pattern Classifiers - Methods and Algorithms. Wiley.

[40] Lam, L., & Suen, C.Y., (1997). 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, 27(5):553-568.

[41] LeCun, Y., Boser, B., & Denker, JS., (1989). Backpropagation applied to handwritten zip code recognition. Neural Computation, 1(4):541–551

[42] LeCun, Y., Kavukcuoglu, K., & Farabet, C., (2010). Convolutional networks and applications in vision. In Proc. of ISCAS, Paris, pp. 253–256.

[43] Lei, L., Wang, C., & Liu, X., (2013). Discrete Wavelet Transform Decomposition level determination exploiting sparseness measurement. International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering, 7,1182-1185.

[44] Liu, F., Zhou, L., Lu, Z.M., & Nie, T., (2015). Palmprint Feature Extraction Based on Curvelet Transform. Journal of Information Hiding and Multimedia Signal Processing, 6(1), pp.131-139.

[45] Liu, T., Sun, B., Chi, M., & Zeng, X., (2017). Gender recognition using dynamic gait energy image. IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Chengdu, pp. 1078-1081.

[46] Michael, G.K.O., Connie, T., & Teoh, A.B.J., (2008). Touch-less palm print biometrics: Novel design and implementation. Image and Vision Computing, 26(12), pp.1551-1560.

[47] Minaee, S., & Wang, Y., (2016). Palmprint recognition using deep scattering convolutional network. arXiv preprint arXiv:1603.09027.

[48] Ming, W., & Yubo, Y., (2014). Gender Classification Based on Geometry Features of Palm Image. The Scientific World Journal, vol. 2014, Article ID 734564.

[49] Misar, M., & Gharpure, D., (2015). 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. doi: 10.1109/ISPTS.2015.7220094.

[50] Mohammed A.A., Jonathan Wu Q.M., & Sid-Ahmed M.A. (2010) Application of Wave Atoms Decomposition and Extreme Learning Machine for Fingerprint Classification. In: Campilho A., Kamel M. (eds.) Image Analysis and Recognition.

ICIAR 2010. Lecture Notes in Computer Science, vol. 6112. Springer, Berlin, Heidelberg.

[51] Mu, M., Ruan, Q., & Shen, Y., (2010). Palmprint Recognition Based on Discriminative Local Binary Patterns Statistic Feature. International Conference on Signal Acquisition and Processing, Bangalore, pp. 193-197.

[52] NEC Automatic Palmprint Identification System (2003). Available: http://www.nectech.com/afis/download/PalmprintDtsht.q.pdf

[53] Ng, C., Tay, Y., & Goi, B., (2013). A Convolutional Neural Network for Pedestrian Gender Recognition. In: Guo, C., Hou, ZG., & Zeng, Z., (eds.) Advances in Neural Networks. Lecture Notes in Computer Science, vol. 7951. Springer, Berlin, Heidelberg.

[54] Nistor, C., Marina, A.C., Darabant, A.S., & Borza, D., (2017). Automatic gender recognition for "in the wild" facial images using convolutional neural networks. 13th IEEE International Conference on Intelligent Computer Communication and Processing (ICCP), Cluj-Napoca, pp. 287-291.

[55] Ojala, T., Pietikainen, M., & Harwook, D., (1996). A comparative study of texture measures with classification based on feature distribution. Pattern Recognition, 29:51–59.

[56] Ouarda, W., Trichili, H., Alimi, A.M., & Solaiman, B., (2013). Combined local features selection for face recognition based on Nave Bayesian classification. Hybrid Intelligent Systems (HIS), 13th International Conference on, Gammarth, pp. 240-245.

[57] Ouarda, W., Trichili, H., Alimi, A.M., & Solaiman, B., (2014). Face recognition based on geometric features using Support Vector Machines. Soft Computing and Pattern Recognition (SoCPaR), 6th International Conference of, Tunis, pp. 89- 95.

[58] Ouarda, W., Trichili, H., Alimi, A.M., & Solaiman, B., (2014). MLP Neural Network for face recognition based on Gabor Features and Dimensionality Reduction techniques. Multimedia Computing and Systems (ICMCS), International Conference on, Marrakech, pp. 127-134.

[59] Ouarda, W., Trichili, H., Alimi, A.M., & Solaiman, B., (2016). Bag of Face Recognition Systems Based on Holistic Approaches. 15th International Conference on Intelligent Systems Design and Applications (ISDA), pp. 430-436.

[60] Promila, & V. Laxmi, (2012). Palmprint Matching Using LBP. International Conference on Computing Sciences, Phagwara, pp. 110-115. doi: 10.1109/ICCS.2012.55

[61] Ross, A.A., Nandakumar, K., & Jain, A., (2006). Handbook of multibiometrics (Vol. 6). Springer Science & Business Media, pp.73-82.

[62] Simonyan, K., & Zisserman, A., (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[63] Sumana, I.J., Islam, M.M., Zhang, D., & Lu, G., (2008). Content based image retrieval using curvelet transform. In Proceedings of 2008 International Workshop on Multimedia Signal Processing, pp. 11–16.

[64] Sun, Q., Zhang, J., Yang, A., & Zhang, Q., (2017). Palmprint recognition with deep convolutional features. In Chinese Conference on Image and Graphics Technologies (pp. 12-19). Springer, Singapore.

[65] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A., (2015). Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1-9.

[66] Tang, Y., (2013). Deep learning using linear support vector machines. arXiv preprint arXiv: 1306.0239.

[67] Teuner, A., Pichler, O., & Hosticka, B.J., (1995). Unsupervised texture segmentation of images using tuned matched Gabor filters. IEEE Trans. Image Process. 4 (6), 863–870.

[68] Vapnik, V., (1995). The Nature of Statistical Learning Theory. NY: Springer-Verlag.

[69] Varshney, V., Gupta, R., & Singh, P., (2014). Hybrid DWT-DCT based method for palm-print recognition. IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), Noida, pp. 000007-000012.

[70] Wang, J., Hu, W., Wang, Z., & Chen, Z., (2011). Human Identification and Gender Recognition from Boxing. In: Sun, Z., Lai, J., Chen, X., & Tan, T., (eds.) Biometric Recognition. Lecture Notes in Computer Science, vol. 7098. Springer, Berlin, Heidelberg.

[71] Wang, X., Gong, H., Zhang, H., & Zhuang, Z., (2006) Palmprint Identification using Boosting Local Binary Pattern. Pattern Recognition, 18th International Conference on Pattern Recognition (ICPR'06), Hong Kong, 2006, pp. 503-506.

[72] Williams, T., & Li, R., (2016). Advanced image classification using wavelets and convolutional neural networks. In Machine Learning and Applications (ICMLA), 15th IEEE International Conference, pp. 233-239.

[73] Xu, L., Krzyzak, A., & Suen, C.Y., (1992). Methods for Combining Multiple Classifiers and their Applications to Handwriting Recognition. IEEE Transactions on Systems, Man, and Cybernetics, 22(3): pp. 418-435.

[74] Xu, X., Zhang, D., Zhang, X., & Cao, Y., (2009). Palmprint Recognition Based on Discrete Curvelet Transform and Support Vector Machine, Journal of Infrared and Millimeter Waves, vol. 28, no. 6, pp. 456-460.

[75] Yan, C., (2011). Face Image Gender Recognition Based on Gabor Transform and SVM. In: Shen, G., & Huang, X., Advanced Research on Electronic Commerce, Web Application, and Communication. Communications in Computer and Information Science, vol. 144. Springer, Berlin, Heidelberg.

[76] Yang, W., Chen, C., Ricanek, K., Sun, C., (2011). Gender Classification via Global-Local Features Fusion. In: Sun, Z., Lai, J., Chen, X., & Tan, T. (eds.) Biometric Recognition. Lecture Notes in Computer Science, vol. 7098. Springer, Berlin, Heidelberg.

[77] Zeng, Z., & Huang, P., (2011). 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, pp. 1917-1921.

[78] Zhao, D., Pan, X., Luo, X., & Gao, X., (2015). Palmprint recognition based on deep learning, In: International Conference on Wireless, Mobile and Multi-Media, pp. 214-217.

[79] Zhao, Q., Bu, W., & Wu, X., (2013) SIFT-based image alignment for contactless palmprint verification, in: Proceedings of International Conference on Biometrics, pp. 16.

List of Publications

[1] Elgallad, E.A., Charfi, N.A., Alimi, M., & Ouarda, W., (2017). Human identity recognition using sparse auto encoder for texture information representation in palmprint images based on voting technique. Sudan Conference on Computer Science and Information Technology (SCCSIT), Elnihood, Sudan, pp. 1-8.

[2] Elgallad, E.A., Ouarda, W., & Alimi, M., (2019). CWNN-Net: A New Convolution Wavelet Neural Network for Gender Classification using Palm Print. International journal of advanced computer science and applications (IJACSA). 2019; 10(5): pp. 29-36, DOI: 10.14569/IJACSA.2019.0100516.

[3] Elgallad, E.A., Ouarda, W., & Alimi, M., (2019). Dense Hand-CNN: A Novel CNN Architecture based on Later Fusion of Neural and Wavelet Features for Identity Recognition. International journal of advanced computer science and applications (IJACSA). 2019; 10(6): pp. 368-378, DOI: 10.14569/IJACSA.2019.0100647.

Appendix A - Research MATLAB Codes

A.1 MVT

```
clc
data='dorsal';
left=1;
right=0;

squeezenet=1;

test=[];
train=[];

if left == 1 && right == 1
    mfile=['testImages ',data,'.mat'];
    load(mfile);
    test=cellstr(testImages.Labels);
    load(['trainingImages ',data,'.mat']);
    train=cellstr(trainingImages.Files);
else
    if left == 1
        data=[data,' left'];
        mfile=['testImages ',data,'.mat'];
        load(mfile);
        test=[test;cellstr(testImages.Labels)];

        load(['trainingImages ',data,'.mat']);
        train=[train;cellstr(trainingImages.Files)];
    end
    if right == 1
        data=[data,' right'];
        mfile=['testImages ',data,'.mat'];
        load(mfile);
        test=[test;cellstr(testImages.Labels)];

        load(['trainingImages ',data,'.mat']);
        train=[train;cellstr(trainingImages.Files)];
    end
end


arr=[];
mydir='/Users/araby/Google Drive/ElGallad.Araby/My
Software/Gender 11K/CNN/CV result/predict ';

if alex_6 == 1
    desc='alexnet ';
    mat_file=[mydir,desc,data,' fc6.mat'];
    load(mat_file);
    arr=[arr;predictedLabels(:,1)'];
end

if alex_7 == 1
    desc='alexnet ';
    mat_file=[mydir,desc,data,' fc7.mat'];
    load(mat_file);
```

```
    arr=[arr;predictedLabels(:,1)'];
end

if alex_8 == 1
    desc='alexnet ';
    mat_file=[mydir,desc,data,' fc8.mat'];
    load(mat_file);
    arr=[arr;predictedLabels(:,1)'];
end

if vgg16_6 == 1
    desc='vgg16 ';
    mat_file=[mydir,desc,data,' fc6.mat'];
    load(mat_file);
    arr=[arr;predictedLabels(:,1)'];
end

if vgg16_7 == 1
    desc='vgg16 ';
    mat_file=[mydir,desc,data,' fc7.mat'];
    load(mat_file);
    arr=[arr;predictedLabels(:,1)'];
end

if vgg16_8 == 1
    desc='vgg16 ';
    mat_file=[mydir,desc,data,' fc8.mat'];
    load(mat_file);
    arr=[arr;predictedLabels(:,1)'];
end

if vgg19_6 == 1
    desc='vgg19 ';
    mat_file=[mydir,desc,data,' fc6.mat'];
    load(mat_file);
    arr=[arr;predictedLabels(:,1)'];
end

if vgg19_7 == 1
    desc='vgg19 ';
    mat_file=[mydir,desc,data,' fc7.mat'];
    load(mat_file);
    arr=[arr;predictedLabels(:,1)'];
end

if vgg19_8 == 1
    desc='vgg19 ';
    mat_file=[mydir,desc,data,' fc8.mat'];
    load(mat_file);
    arr=[arr;predictedLabels(:,1)'];
end

if resnet50 == 1
    desc='resnet50 ';
    mat_file=[mydir,desc,data,'.mat'];
    load(mat_file);
    arr=[arr;predictedLabels(:,1)'];
end

if resnet101 == 1
    desc='resnet101 ';
```

```
    mat_file=[mydir,desc,data,'.mat'];
    load(mat_file);
    arr=[arr;predictedLabels(:,1)'];
end

fprintf(' No of Desc.  : %1.0f\n ', size(arr,1));

no_of_persons=size(arr,2);
test_arr_size=size(test,1);
step=test_arr_size/no_of_persons;

ones=0;
for i= 1 : no_of_persons
    f = sum(arr(:,i) == 'F');
    m = sum(arr(:,i) == 'M');

    if f > m &&  test{i,1} == 'F'
        ones=ones+1;
    elseif m > f && test{i,1} == 'M'
            ones=ones+1;
    else
        fprintf('%100s \n, ', trainingImages.Files{i,1});
    end

end

acc= ones/no_of_persons*100;
fprintf('\n Mean accuracy: %5.2f\n ', acc);

% desc='gabor';
% mat_file=[data,'_',desc,'_features.mat'];
% load(mat_file);
% arr=[arr;outputclassSVM.desc];

% desc='wave';
% mat_file=[data,'_',desc,'_features.mat'];
% load(mat_file);
% arr=[arr;outputclassSVM.desc];

% desc='atom';
% mat_file=[data,'_',desc,'_features.mat'];
% load(mat_file);
% arr=[arr;outputclassSVM.desc];

% desc='curvelet';
% mat_file=[data,'_',desc,'_features.mat'];
% load(mat_file);
% arr=[arr;outputclassSVM.desc];

% desc='lbp';
% mat_file=[data,'_',desc,'_features.mat'];
% load(mat_file);
% arr=[arr;outputclassSVM.desc];

% desc='sift';
% mat_file=[data,'_',desc,'_features.mat'];
% load(mat_file);
% arr=[arr;outputclassSVM.desc];
```

A.2 Dense Hand-CNN Code

A.2.1 System 1

```
clear;
clc;
rng('default')
tic;
warning('off','all');

global no_of_persons;
global no_of_left;
global no_of_ltrain;
global no_of_ltest;
global no_of_right;
global no_of_rtrain;
global no_of_rtest;
global classifier;

no_of_persons = 230;
no_of_left = 5;
no_of_ltrain = 3;
no_of_ltest = no_of_left-no_of_ltrain;
no_of_right = 0;
no_of_rtrain = 0;
no_of_rtest = no_of_right-no_of_rtrain;

% gabor, wave, atom, curvelets, SIFT, CNN, LBP
descriptors_arr=[1 1 1 1 1 1 1];

% 1: softmax(AE)1, 2:softmax2, 3:SVM1(AE), 4:SVM2, 5:ED
classifier = 2;

%% Get Features for all descriptors and its labels

[dataset,datasetLabels,datasetLabels1d]=getFeaS3(descriptors_a
rr);
%load('Fea_SpA_No.mat');
%load('Fea_SpA_Yes.mat');

%%
fprintf('dataset           :   %4d x %4d\n', size(dataset));
fprintf('dataset lables    :   %4d x %4d\n',
size(datasetLabels));
fprintf('dataset lables 1d :   %4d x %4d\n',
size(datasetLabels1d));
fprintf('\n');


%% Extract TRAIN TEST %%

[trainData,trainLabels,trainLabels1d,testData,testLabels,testL
abels1d]=train_test(dataset,datasetLabels1d);
```

```
xTrain=trainData;
tTrain=trainLabels;      % for softmax
tTrain1d=trainLabels1d;  % for SVM

xTest=testData;
tTest=testLabels;        % for softmax
tTest1d=testLabels1d;    % for SVM

fprintf('train dataset     :   %4d x %4d\n', size(xTrain));
fprintf('train lables      :   %4d x %4d\n', size(tTrain));
fprintf('train lables 1d   :   %4d x %4d\n', size(tTrain1d));
fprintf('\n');

fprintf('test dataset      :   %4d x %4d\n', size(xTest));
fprintf('test lables       :   %4d x %4d\n', size(tTest));
fprintf('test lables 1d    :   %4d x %4d\n', size(tTest1d));
fprintf('\n');

%% Classifiers

if classifier == 1
    % Softmax & autoencoder

[autoenc1,autoenc2,feat2]=sparse_ae(xTrain,2,0.5,1000,0.19);

    softnet = trainSoftmaxLayer(feat2,tTrain,'MaxEpochs',300);
    deepnet = stack(autoenc1,autoenc2,softnet);
    view(deepnet)

    deepnet = train(deepnet,xTrain,tTrain);
    y = deepnet(xTest);
    [c,cm,ind,per]=confusion(tTest,y);
    disp(100-c*100);

    tst= descriptors_arr(1) + descriptors_arr(2) +
descriptors_arr(3) + descriptors_arr(4) + descriptors_arr(5);
    if tst == 1
        if descriptors_arr(1) > 0
            outputclass.gabor=y;
        elseif descriptors_arr(2) > 0
            outputclass.wave=y;
        elseif descriptors_arr(3) > 0
            outputclass.atom=y;
        elseif descriptors_arr(4) > 0
            outputclass.curvelet=y;
        elseif descriptors_arr(5) > 0
            outputclass.sift=y;
        end

        save('outputClass.mat', 'outputclass','-append')
    end
end

%%
if classifier == 2
    % Softmax Classifier 2

    tst= descriptors_arr(1) + descriptors_arr(2) +
descriptors_arr(3) + descriptors_arr(4) + descriptors_arr(5) +
descriptors_arr(6) + descriptors_arr(7) ;
    if tst == 7
```

```
        X=dataset;
        T=datasetLabels1d;
        Classifier2_Softmax(X,T,0);
    end
end

%%
if classifier == 3
    % SVMae

    tst= descriptors_arr(1) + descriptors_arr(2) +
descriptors_arr(3) + descriptors_arr(4) + descriptors_arr(5);
    if tst == 5
        Classifier3_SVM(xTrain,tTrain1d,xTest,tTest1d,0);
    end
end

%%
if classifier == 4
    % linear SVM

    tst= descriptors_arr(1) + descriptors_arr(2) +
descriptors_arr(3) + descriptors_arr(4) + descriptors_arr(5) +
descriptors_arr(6) + descriptors_arr(7) ;
    if tst == 7
        sc_fea=dataset;
        sc_label=datasetLabels1d;
        Classifier4_SVM(sc_fea,sc_label,0);
    end
end

%%
if classifier == 5
    % ED
    m=xTrain';
    v=xTest';
    for i=1:size(xTest,2)
        for j=1:size(xTest,1)
            y(j)=v(i,j);
        end
        for k=1:size(xTrain,2)
            for jj=1:size(xTrain,1)
                x(jj)=m(k,jj);
            end

            d=y-x;
            d=power(d,2);
            sumd=sum(d);
            fit(k)= power(sumd,.5);
        end
        reg=min(fit);
        for u=1:size(xTrain,2)
            if(reg==fit(u))
                index(i)=u;
            end
        end
    end

    File9=fopen('r7.dat','w');
    for i=1:size(xTest,2)
        fprintf(File9,' %d \n ',index(i));
    end
```

```
    nim2=size(xTest,2);
    train=no_of_ltrain;
    test=no_of_ltest;
    a1=1;
    File1=fopen('r7.dat','r');
    [v1]=fscanf(File1,'%f',[1,nim2]);
    fclose(File1);
    v1=transpose(v1);

    ones=0;
    for i=1:2:nim2
        for j=0 : train -2
            if v1(i+j) == a1 || v1(i+j) == a1+1  || v1(i+j)==
a1+2 % || v1(i+j)== a1+3 || v1(i+j)== a1+4 || v1(i+j)== a1+5
                ones=ones+1;
            end
        end
        a1=a1+train;
    end
    sprintf('average is %5.2f % \n', (ones)/nim2*100)
end

%%
toc;

**************************************************



function
[trainData,trainLabels,trainLabels1d,testData,testLabels,testL
abels1d] = train_test(dataset,datasetLabels1d)

global no_of_persons;
global no_of_left;
global no_of_ltrain;
global no_of_ltest;
global no_of_right;
global no_of_rtrain;
global no_of_rtest;

i=1;
j=1;
dex=1;
trainLabels=zeros(no_of_persons,no_of_persons*(no_of_ltrain+no
_of_rtrain));
trainLabels1d=[];
testLabels =zeros(no_of_persons,no_of_persons*(no_of_ltest
+no_of_rtest));
testLabels1d=[];

for index = 1 : no_of_left + no_of_right : no_of_persons *
(no_of_left + no_of_right)

    for ii=1 : no_of_ltrain
        trainData(:,i)=dataset(:,index+ii-1);
        trainLabels(dex,i)=1;
        trainLabels1d(i,1)=datasetLabels1d(index+ii-1);
        i=i+1;
    end
```

```
    for jj= no_of_ltrain+1 : no_of_left
        testData(:,j)=dataset(:,index+jj-1);
        testLabels(dex,j)=1;
        testLabels1d(j,1)=datasetLabels1d(index+ii-1);
        j=j+1;
    end

    for kk= no_of_left+1 : no_of_left + no_of_rtrain
        trainData(:,i)=dataset(:,index+kk-1);
        trainLabels(dex,i)=1;
        trainLabels1d(i,1)=datasetLabels1d(index+ii-1);
        i=i+1;
    end

    for ll= no_of_left + no_of_rtrain+1 : no_of_left+
no_of_right
        testData(:,j)=dataset(:,index+ll-1);
        testLabels(dex,j)=1;
        testLabels1d(j,1)=datasetLabels1d(index+ii-1);
        j=j+1;
    end
    dex=dex+1;
end


end


function
[autoenc1,autoenc2,sparseArr]=sparse_ae(feaArr,no_of_ae,factor
,Epoch,SP)

% rng('default')
autoenc1=[];
autoenc2=[];

if size(feaArr,1) > 0
    if no_of_ae == 0
        sparseArr = feaArr;
    end

    if no_of_ae > 0
        hiddenSize1 = fix(size(feaArr,1) * factor);
        autoenc1 = trainAutoencoder(feaArr,hiddenSize1 , ...
             'MaxEpochs',Epoch, ...
             'UseGPU',false, ...
             'L2WeightRegularization',0.004, ...
             'SparsityRegularization',4, ...
             'SparsityProportion',SP, ...
             'ScaleData', true);
% 'SparsityProportion' ó Desired proportion of training
examples a neuron reacts to
% 0.05 (default) | positive scalar value in the range from 0
to 1
% Desired proportion of training examples a neuron reacts to,
specified as the comma-separated pair
% consisting of 'SparsityProportion' and a positive scalar
value.
% Sparsity proportion is a parameter of the sparsity
regularizer.
```

```
% It controls the sparsity of the output from the hidden
layer.
% A low value for SparsityProportion usually leads to each
neuron in the hidden layer "specializing"
% by only giving a high output for a small number of training
examples.
% Hence, a low sparsity proportion encourages higher degree of
sparsity. See Sparse Autoencoders.
% Example: 'SparsityProportion',0.01 is equivalent to saying
that each neuron in the hidden layer
% should have an average output of 0.1 over the training
examples.

        feat1 = encode(autoenc1,feaArr);
        sparseArr = feat1;
        if no_of_ae > 1
            hiddenSize2 = fix(hiddenSize1 * factor);
            autoenc2 = trainAutoencoder(feat1,hiddenSize2 ,
...
                'MaxEpochs',Epoch, ...
                'UseGPU',false, ...
                'L2WeightRegularization',0.004, ...
                'SparsityRegularization',4, ...
                'SparsityProportion',SP, ...
                'ScaleData', true);
            feat2 = encode(autoenc2,feat1);
            sparseArr = feat2;
        end
    end


end

end

**********************************************
```

A.2.2 System 2

```
clc;
tic
rng('default');

folder='/Users/araby/Documents/Database/Palmprint_DB/IITD
Palmprint V1/Segmented/Left_cat';
imds = imageDatastore(folder, 'IncludeSubfolders',true,
'LabelSource','foldernames');

[imdsTrain,imdsTest] = splitEachLabel(imds,0.6,'randomized');

net = squeezenet();

layer = 'fire9-concat';
net.Layers

%%
```

```
inputSize = net.Layers(1).InputSize;

augmentedTrainingSet = augmentedImageDatastore(inputSize(1:2),
imdsTrain, 'ColorPreprocessing', 'gray2rgb');
augmentedTestSet = augmentedImageDatastore(inputSize(1:2),
imdsTest, 'ColorPreprocessing', 'gray2rgb');

%%

trainingFeatures = activations(net, augmentedTrainingSet,
layer, ...
    'MiniBatchSize', 32, 'OutputAs', 'columns');
trainingLabels = imdsTrain.Labels;


testFeatures = activations(net, augmentedTestSet, layer, ...
    'MiniBatchSize', 32, 'OutputAs', 'columns');
testLabels = imdsTest.Labels;

classifier = fitcecoc(trainingFeatures, trainingLabels, ...
    'Learners', 'Linear', 'Coding', 'onevsall',
'ObservationsIn', 'columns');


predictedLabels = predict(classifier, testFeatures,
'ObservationsIn', 'columns');

confMat = confusionmat(testLabels, predictedLabels);

% Convert confusion matrix into percentage form
confMat = bsxfun(@rdivide,confMat,sum(confMat,2));

% Display the mean accuracy
mean(diag(confMat))


toc

**********************************************
```

A.3 CWNN-Net

```
clc;
tic;
rng('default');
load('HandInfo_sorted.mat');
%%
dorsal=1;
palmar=0;
```

```
tstImgNo=350;

% 1:cA 2:cH
% 3:cV 4:cD
filter = 1;

left=1;
right=0;

remove_dorsal_nail_polish=0;
remove_accessories=0;

net = squeezenet();
net1='squeezenet';

featureLayer = 'fire9-concat';

%%

if dorsal==1
    tmp = 'dorsal ';
else
    tmp = 'palmar ';
end

if left==1
    tmp = [tmp,'left'];
else
    tmp = [tmp,'right'];
end

wantedelements = HandInfo(strcmp({HandInfo.aspectOfHand}, tmp));

if remove_dorsal_nail_polish == 1
    tst = arrayfun(@(x) x.nailPolish == 0, wantedelements);
    wantedelements=wantedelements(tst);
end

if remove_accessories == 1
    tst = arrayfun(@(x) x.accessories == 0, wantedelements);
    wantedelements=wantedelements(tst);
end

rootFolder = ['/Users/araby/Documents/Database/11k
Hands/HandsWave',num2str(filter)];

net.Layers(1);
net.Layers(end);
net.Layers

disp(tmp)

%%
names = extractfield(wantedelements, 'imageName');
categories=extractfield(wantedelements, 'gender');
categories=categorical(categories);
imageDS = imageDatastore(fullfile(rootFolder, names),
'Labels',categories);

T = countEachLabel(imageDS)
```

```
%%


%[trainingSet, testSet] = splitEachLabel(imageDS, 0.7,
'randomize');
[testSet,trainingSet] = splitEachLabel(imageDS, tstImgNo,
'randomize');
Tr = countEachLabel(trainingSet)
Ts = countEachLabel(testSet)

imageSize = net.Layers(1).InputSize;
augmentedTrainingSet = augmentedImageDatastore(imageSize,
trainingSet, 'ColorPreprocessing', 'gray2rgb');
augmentedTestSet = augmentedImageDatastore(imageSize, testSet,
'ColorPreprocessing', 'gray2rgb');

trainingFeatures = activations(net, augmentedTrainingSet,
featureLayer, ...
    'MiniBatchSize', 32, 'OutputAs', 'columns');

trainingLabels = trainingSet.Labels;

classifier = fitcecoc(trainingFeatures, trainingLabels, ...
    'Learners', 'Linear', 'Coding', 'onevsall', 'ObservationsIn',
'columns');

testFeatures = activations(net, augmentedTestSet, featureLayer,
...
    'MiniBatchSize', 32, 'OutputAs', 'columns');

predictedLabels = predict(classifier, testFeatures,
'ObservationsIn', 'columns');

testLabels = testSet.Labels;


X = ['no of features: ',num2str(length(trainingFeatures))];
disp(X)

% Tabulate the results using a confusion matrix.
confMat = confusionmat(testLabels, predictedLabels);

% Convert confusion matrix into percentage form
confMat = bsxfun(@rdivide,confMat,sum(confMat,2));

savename=['predict ',net1,' ',tmp,' ',num2str(filter),'.mat'];
save(savename,'predictedLabels')

% Display the mean accuracy
X=round(mean(diag(confMat))*10000)/10000;
X = ['Accuracy: ',num2str(X*100),' %'];
disp(X)
toc

%% test one image (to be in mobile app)
i=1;
newImage = imread(fullfile(rootFolder, char(names(i))));
ds = augmentedImageDatastore(imageSize, newImage,
'ColorPreprocessing', 'gray2rgb');
```

```
imageFeatures = activations(net, ds, featureLayer, 'OutputAs',
'columns');
label = predict(classifier, imageFeatures, 'ObservationsIn',
'columns');

X=[char(names(i)),'is ',categories(i),' and recognized as ',
label(i)];
disp(X)
```