People's Democratic Republic of Algeria

Kasdi Merbah University Ouargla

Faculty of New Information and Communication Technologies

Department of Electronics and Telecommunications

ACADEMIC MASTER'S THESIS

Field: Automatic

Specialization: Automatic and Systems

Presented by:

RAHMANI Mohammed
BOUCHAALA Hicham

TITLE

Application of Deep Learning Techniques for Biometric Systems

Defended publicly
on : 24/06/2024

Before the jury:

Mm. KARA Fouzia	**MAA**	**President**	**UKM Ouargla**
Mr. CHLAOUA Rachid	**MCB**	**Thesis Director**	**UKM Ouargla**
Mr. NASRI Nadji	**MAA**	**Examiner**	

Academic Year : 2023/2024

بسم الله الرحمن الرحيم

Acknowledgments

We would like to extend our sincere gratitude to our supervisor, **Dr. CHLAOUA Rachid**, for his invaluable advice and guidance, which were instrumental in the success of this work. Additionally, we express our heartfelt thanks to the jury members, **Dr. KARA Fouzia** and **Dr. NASRI Nadjib** for their willingness to engage in meaningful discussions and provide valuable insights on our thesis.

Dedication

To my Parents, Family, and Friends,

I am deeply grateful for the unwavering support and guidance you have provided me.

My parents and family, your love, and your sacrifices have shaped me into who I am today.

To all my friends, your friendship has added joy and adventure to my journey.

And all my classmates and professors. Thank you all for being integral parts of my life.

With heartfelt appreciation,

Rahmani Mohammed.

To My Beloved Family and Cherished Friends,

I owe a deep debt of gratitude to all of you for your steadfast support and encouragement.

To my parents, your unwavering faith in me and your countless sacrifices have been my greatest motivation.

To my family, your continuous love and backing have been a pillar of strength for me.

To my friends, your companionship and encouragement have enriched my life immensely.

To my classmates and professors, thank you for your guidance and inspiration.

With profound gratitude,

Bouchaala Hicham.

ملخص

تركز هذه الأطروحة على تطوير وتقييم نظام التعرف البيومتري، مع التركيز بشكل خاص على التعرف على بصمة اليد ثنائية الأبعاد باستخدام تكامل بنية الشبكة العصبية العميقة GoogLeNet . يغطي الخلفية النظرية أهمية البيومترية، وأنواع أنظمة التعرف البيومتري، بما في ذلك الأنظمة متعددة الوسائط، ونظرة عامة على التعلم العميق، وأنواعه، وتطبيقاته، وفوائده.
يتضمن نظام التعرف البيومتري المقترح استخدام بنية GoogLeNet للتصنيف واستخراج الميزات. باستخدام قاعدة بيانات PolyU Palmprint ، تشمل التجارب والنتائج اختيار المعايير وتقييم أنظمة التعرف البيومتري الأحادية والمتعددة الوسائط. يتم إجراء دراسة مقارنة لتحليل فعالية النظام المقترح.
في الختام، تقدم هذه الأطروحة رؤى حول تطوير وتنفيذ وتقييم نظام تعرف بيومتري جديد، وتبرز فعاليته المحتملة في التطبيقات الواقعية.

الكلمات المفتاحية: التعرف البيومتري، بصمة اليد ثنائية الأبعاد، GoogLeNet ، التعلم العميق، الأنظمة الأحادية والمتعددة الوسائط، قاعدة بيانات PolyU Palmprint

Abstract

This thesis focuses on developing and evaluating a biometric recognition system, specifically utilizing 2D palmprint recognition integrated with the GoogLeNet deep neural network architecture. The theoretical background encompasses the significance of biometrics, the various types of biometric systems, including multimodal systems, and an overview of deep learning, its types, applications, and benefits.
The proposed biometric recognition system employs the GoogLeNet architecture for both classification and feature extraction. Using the PolyU Palmprint Database, experiments and results include parameter selection and

the evaluation of both unimodal and multimodal biometric systems. A comparative study is conducted to assess the effectiveness of the proposed system.
In conclusion, this thesis provides insights into the development, implementation, and evaluation of a novel biometric recognition system, highlighting its potential effectiveness in real-world applications.

Keywords: Biometric recognition, 2D palmprint, GoogLeNet, Deep learning, Unimodal and multimodal systems, PolyU Palmprint Database.

Résumé

Ce mémoire se concentre sur le développement et l'évaluation d'un système de reconnaissance biométrique, en particulier en utilisant la reconnaissance d'empreintes de paume 2D intégrée à l'architecture du réseau neuronal profond GoogLeNet. Le contexte théorique englobe l'importance de la biométrie, les différents types de systèmes biométriques, y compris les systèmes multimodaux, et une vue d'ensemble de l'apprentissage profond, ses types, applications et avantages.
Le système de reconnaissance biométrique proposé utilise l'architecture GoogLeNet pour la classification et l'extraction de caractéristiques. En utilisant la base de données PolyU Palmprint, les expériences et les résultats incluent la sélection des paramètres et l'évaluation des systèmes biométriques unimodaux et multimodaux. Une étude comparative est menée pour évaluer l'efficacité du système proposé.
En conclusion, ce mémoire fournit des informations sur le développement, la mise en œuvre et l'évaluation d'un nouveau système de reconnaissance biométrique, en soulignant son efficacité potentielle dans des applications réelles.

Mots clés : Reconnaissance biométrique, empreintes de paume 2D, GoogleNet, apprentissage profond, systèmes unimodaux et multimodaux, base de données PolyU Palmprint

Table of Contents

List of Figures

List of Tables

General Introduction

Biometrics holds significant importance in various aspects of our lives, from everyday security measures to advanced technological applications. Biometric systems offer a more reliable and secure way to verify someone's identity compared to traditional methods. Unlike passwords or PINs, biometric identifiers like fingerprints, iris scans, and facial recognition are unique to each individual, making them inherently more difficult to forge or steal. This is because they rely on inherent physical or behavioral traits that are difficult to replicate [3].

Biometric authentication can be significantly faster than traditional methods, allowing for quicker access to devices, applications, or restricted areas. It plays a crucial role in suspect identification, criminal investigations, and border control processes. Biometric authentication can add an extra layer of security for financial transactions, reducing the risk of unauthorized access [3].

Biometric modalities refer to the various unique physiological and behavioral traits that can be used for personal identification. These traits are measured and converted into data that can be compared against a stored template or profile to verify a person's claimed identity. Some of the most common biometric modalities are shown in Fig 0.1.

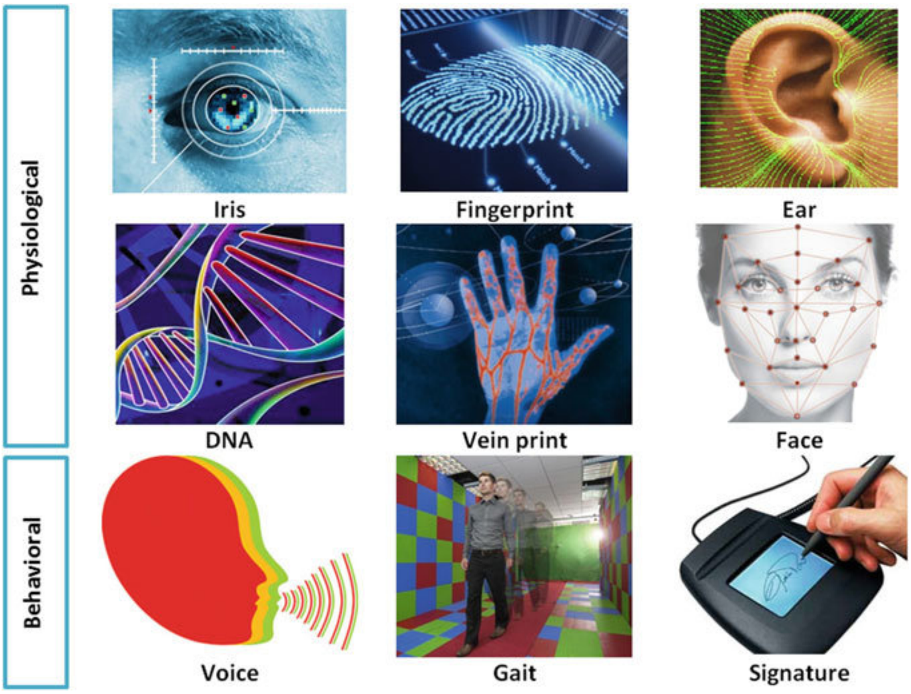


Figure 0.1: Some of the most common biometric modalities.

Physiological Modalities rely on physical characteristics of the body, like as face, fingerprints, iris, retina, hand geometry and DNA, etc. But behavioral Modalities rely on behavioral characteristics, such as the way a person walks, types, or speaks, keystroke dynamics, and signature analysis [4].

The rest of the thesis is organized as follows:

The Second Chapter talks about biometric recognition systems and applications of deep learning techniques in this field. The chapter begins with an introduction to biometric recognition systems, then explores various biometric systems, and then moves on to explain deep learning and its different types and architectures. For deep learning, the types of learning are detailed as supervised, unsupervised, and semi-supervised learning. The chapter also covers the different architectures used in deep learning. Finally, the chapter addresses the practical applications of deep learning techniques in biometric recognition systems, as well as the advantages of using these techniques in this field.

The Third Chapter discusses the technique of 2D palmprint recognition and proposes a recognition model using the GoogleNet deep neural network. It begins with an introduction to palmprint recognition, followed by a description of the proposed recognition, which utilizes GoogleNet as a deep neural network for our recognition. Then, the chapter elucidates the specific contribution of the study, which may involve improvements on previous models or introducing a novel approach to a specific problem. Subsequently, details of the GoogleNet deep neural network are reviewed, including its architectural design, feature extraction processes, and classification. Finally, the unique feature fusion technique used in the proposed model is discussed, where features extracted from palmprints are fused to enhance recognition accuracy.

The Fourth Chapter provides a comprehensive overview of the experimentation process and results related to the studied biometric system. It includes a description of the database used and the process of selecting appropriate parameters. Additionally, it offers an evaluation of the system's performance at both individual and multimodal levels, accompanied by a comparative analysis of the obtained results.

Finally, **The Fifth Chapter** summarizes the thesis and the contribution and presents some concluding remarks. Possible future directions of this research are also discussed in this chapter.

Chapter

1

Theoretical Background

1.1 Introduction

Biometric systems rely on a rich theoretical foundation that draws from various disciplines to achieve reliable and secure identification. This field provides the core concepts for extracting unique features (patterns) from biometric data such as fingerprints, faces, irises, etc. Statistical methods, machine learning algorithms, and image processing techniques all play a crucial role in optimal identification. The Biometric system is defined as "a system that automatically distinguishes and recognizes a person as individual and unique through a combination of hardware and pattern recognition algorithms based on certain physiological or behavioral characteristics that are inherent to that person" [5].

1.2 Proposed Recognition

Biometric systems use unique physiological or behavioral characteristics of a person to reliably recognize. These characteristics are electronically captured, processed by a system, and compared to a stored template or database of known individuals.
The biometric system components are divided into two main parts:

1.2.1 Registration

During registration, the user's biometric traits such as fingerprint, face image, iris pattern, etc... are captured by a sensor. Then, the raw data undergoes feature extraction, where relevant and distinctive features are identified for comparison. Lastly, a unique biometric template (mathematical representation) is created based on the extracted features. This template is securely stored in a database [6].
Fig 1.1 illustrates registration processes in a biometric system.

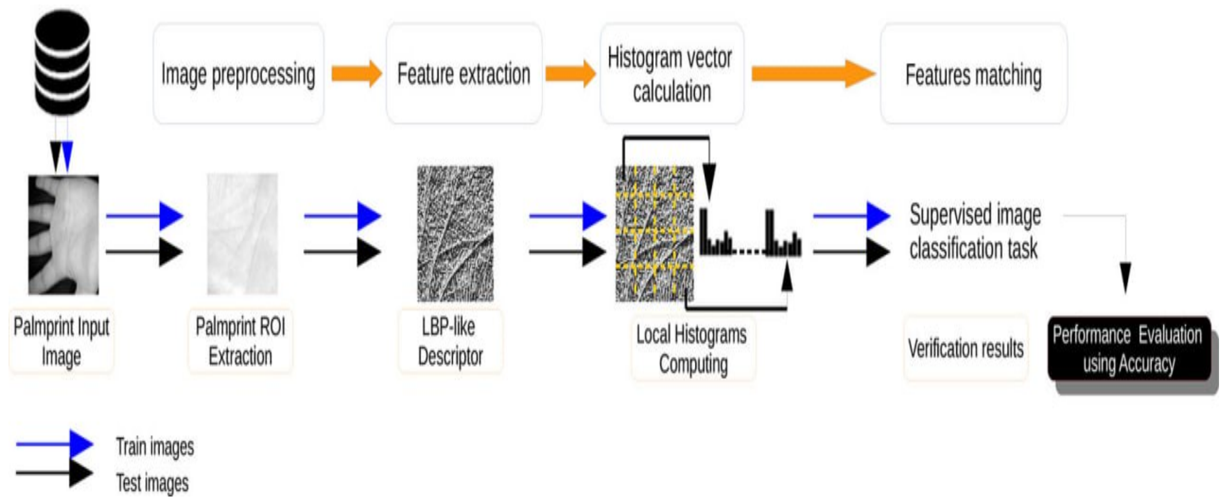


Figure 1.1: Registration processes in a biometric system.

1.2.2 Recognition

During recognition, the matching engine compares the captured features with the stored templates. Then the decision maker determines whether the match is successful (identification/verification) or not, based on a predetermined threshold. In identification mode, the system compares the newly captured biometric data from an unknown user with all templates in the database to determine his identity. But for verification, the system compares the newly captured biometric data with the pre-registered user's template (claimed identity) to confirm whether the individual is indeed who they claim to be. [6]. Fig 1.2 illustrates recognition processes in a biometric system.

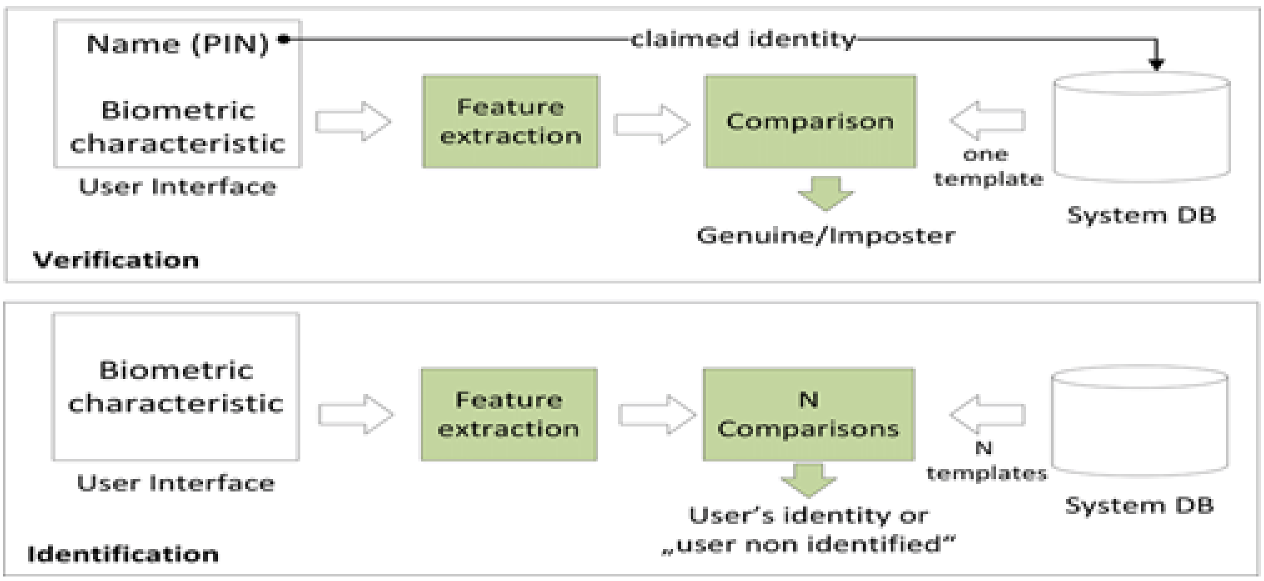


Figure 1.2: Recognition processes in a biometric system.

1.3 Multimodal Biometric Systems

Unimodal biometric system is a type of biometric system that is based on a single biometric modality for recognition. This typefaces limitations such as low security, poor recognition, less accurate results, and spoofing attacks. Also, the performance is greatly affected by environmental factors such as noisy data [7]. These limitations can be overcome by a multimodal biometric system. Multimodal biometrics refers to the use of a combination of two or more biometrics representing an emerging trend. The most compelling reason to combine different modalities is to improve the recognition rate. Multimodal biometric systems take input from single or multiple sensors measuring two or more different modalities of biometric characteristics [7]. According to the sources of information, we have five possibilities to create multibiometric systems as the following Fig 1.3.

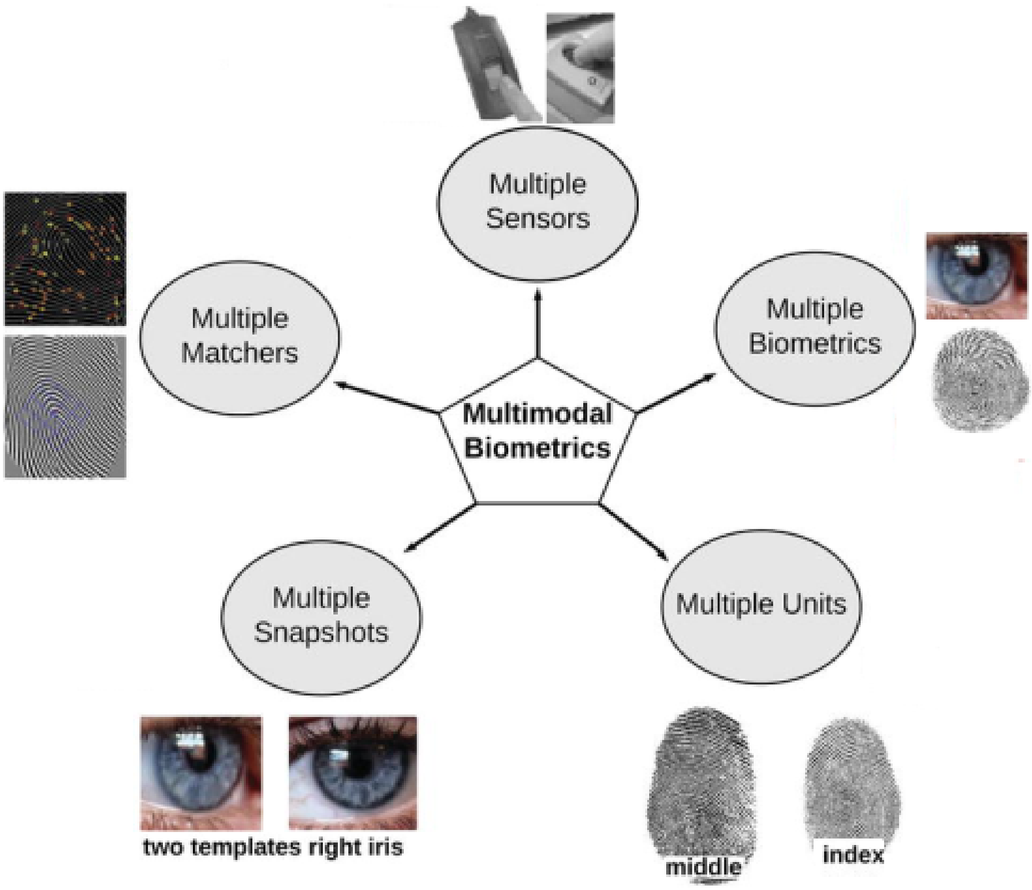


Figure 1.3: The sources of information in the multimodal biometric system.

In contrast to unimodal systems, the advantages of multibiometric systems include:

- Resistance to intra-class similarity: By combining multiple modalities, multimodal biometric systems reduce similarity within the same class, such as facial features.

- Noise resistance: Multimodal systems are more resistant to noise compared to unimodal systems due to their utilization of multiple modalities, which provides additional data for matching.
- Reduced vulnerability to spoofing: It is difficult to spoof multiple modalities simultaneously, making multimodal biometric systems less susceptible to spoofing.

1.4 Deep Learning

Deep learning encompasses a series of algorithms within the realm of machine learning that strive to comprehend information across multiple layers, corresponding to varying levels of abstraction. Typically leveraging artificial neural networks, these algorithms form learned statistical models wherein each layer represents a distinct level of conceptual understanding. Higher-level concepts are derived from lower level ones, and a single lower-level concept may contribute to defining multiple higher-level concepts. Deep learning constitutes a subset of machine learning methods centered on learning representations [8]. The relationship between deep learning algorithms and artificial intelligence types is illustrated in the Fig 1.4.

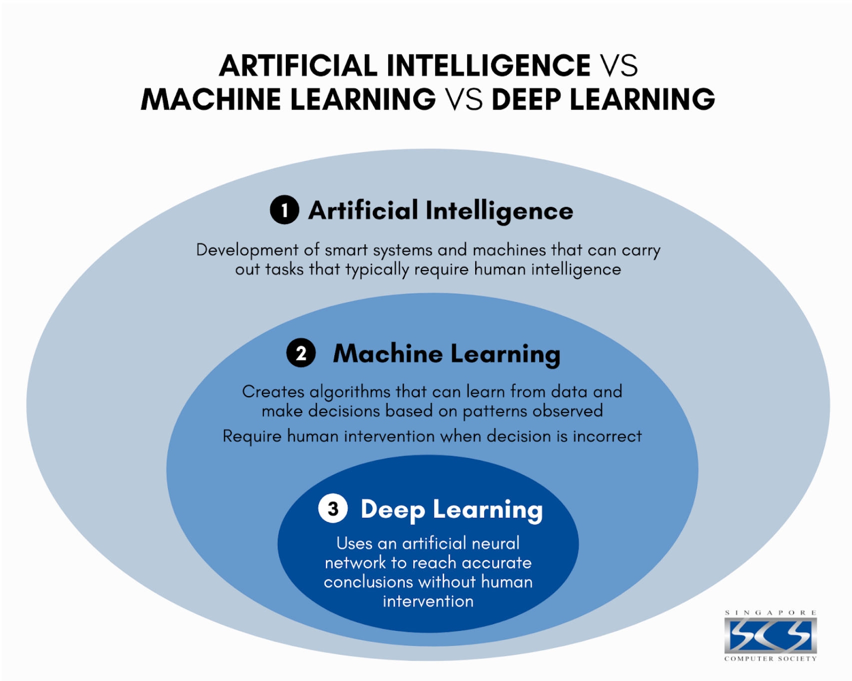


Figure 1.4: Deep learning and artificial neural networks.

1.4.1 Deep Learning Types

Depending on the type of learning process, deep learning algorithms can be categorized into:

a. **Supervised learning:**
   Supervised learning relies on a training dataset containing examples for both inputs and labeled output values. These input-output pairs are used to adjust the parameters of the machine-learning model. Once trained, the model can predict the target variable for new input data. Supervised learning can be categorized into regression problems, where numeric values are predicted, and classification problems, where the prediction is a categorical class (Fig 1.5).

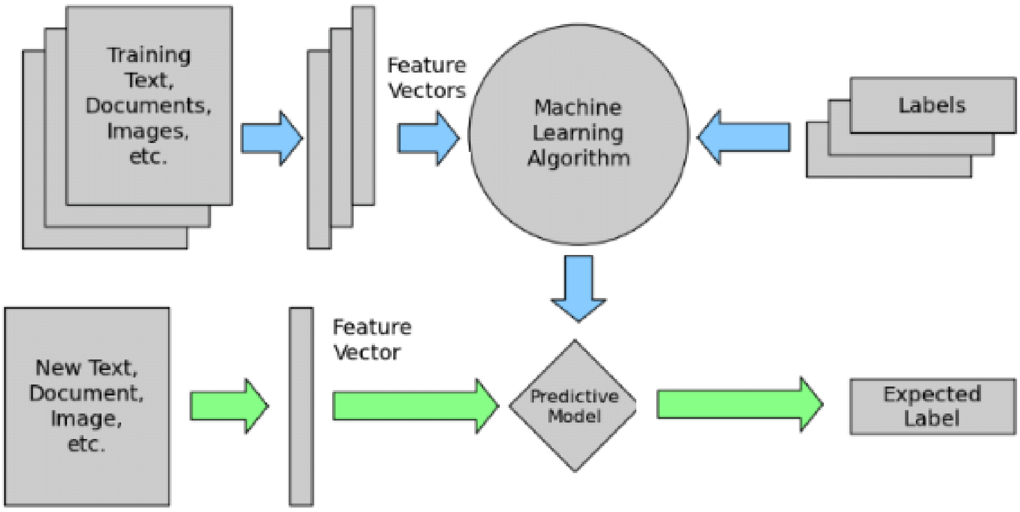


Figure 1.5: Supervised Learning Model.

b. **Unsupervised Learning:**
   Unsupervised learning occurs when a system learns to identify patterns without predefined labels. In this type of learning, the training data consists solely of input variables x, aiming to uncover structural information, like identifying groups of elements with shared characteristics (clustering in Fig 1.6) or reducing the dimensionality of the data by projecting it into a lower-dimensional space (dimensionality reduction).

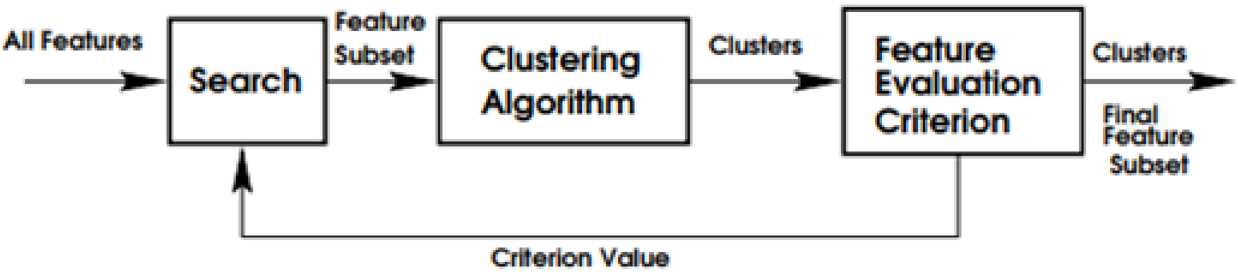


Figure 1.6: Example approach for unsupervised learning.

c. **Semi-supervised learning:**

Semi-supervised learning is a class of supervised learning tasks and techniques that leverage both labeled and unlabeled data for training as shown in Fig 1.7. Typically, only a small portion of the data is labeled, while the majority remains unlabeled. This approach bridges the gap between unsupervised learning, which lacks labeled data altogether, and supervised learning, which relies entirely on labeled data. Researchers have discovered that incorporating unlabeled data alongside a limited amount of labeled data can significantly enhance learning accuracy.

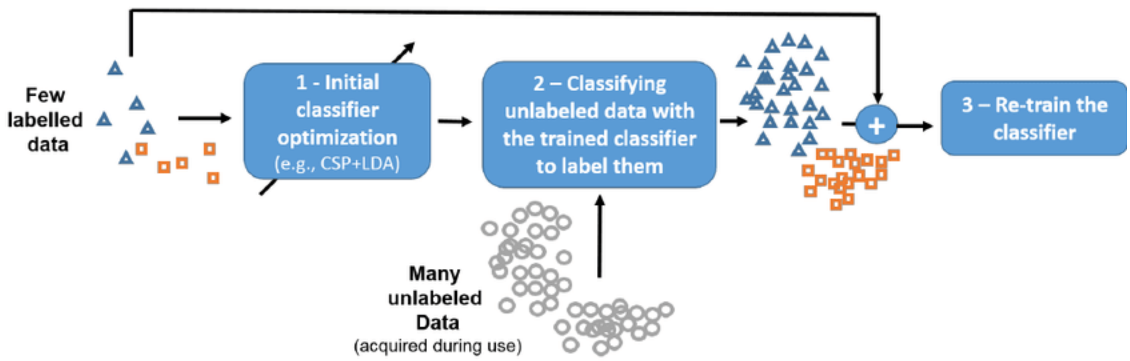


Figure 1.7: Principle of semi-supervised learning.

1.4.2 Applications and Advantages

Nowadays, applications of deep learning include classification, data processing, speech and audio processing, social network analysis, and healthcare, etc... [9]. Other than all the aforementioned applications, deep learning algorithms are also applied to information retrieval, robotics, transportation prediction, autonomous driving, biomedicine, disaster management, and so forth. For that, deep learning has shown its capability to be leveraged in various applications.

The several advantages underpinning deep learning models, including image processing and recognition, speech recognition, self-driving cars, and so on, have sparked such widespread attention. The main benefit of using deep learning models over machine learning (ML) technologies is their capacity to produce new features through a limited range of features in the trained dataset [10]. These models can generate new tasks for solving current ones also covering a variety of human life aspects. A significant amount of time can be saved using deep learning models when dealing with massive datasets, as deep learning algorithms can generate features without the need for human intervention [11].

1.5 Conclusion

In this chapter, the theoretical background of biometric recognition systems is reviewed. Which rely on using individuals' unique characteristics, such as fingerprints and faces to identify them. Then, we discussed multimodal biometric recognition systems, which combine several types of biometric data to improve recognition accuracy. After that, deep learning was reviewed, which is considered one of the innovative techniques in the field of recognition.

Overall, the chapter underscores the intersection of biometrics and deep learning, showcasing how advancements in machine learning technologies have revolutionized the field and opened up new possibilities for robust and efficient biometric authentication systems.

Chapter

2

Proposed Biometric Recognition

2.1 Introduction

Using a reliable recognition method is a critical challenge for both academic and industrial contexts. Given the convenience and accuracy of biometric-based methods, they have huge potential in diverse applications such as e-marketing, security, access control, e-banking, etc [12].

Compared to different types of biometrics, hand-based biometrics have received significant attention in recent years. There has been extensive research in the field of automatic hand biometric recognition, with a variety of techniques proposed in the literature. Based on the information they utilize, hand biometrics recognition techniques can be classified into three main types: texture-based, line-based, and appearance-based [13].

Although the previous methods have shown good performance, experiments are still required to achieve a highly reliable and accurate system capable of operation for high-security applications.

2.2 Proposed Recognition

The use of palmprints is considered a common method in biometric systems. Palmprints are rich in features and can be recognized either through the main lines and wrinkles, which can be obtained even from low-resolution images, or through minutiae details. However, the latter requires high-resolution images of the palmprint [14]. Also, palmprints can provide personal gender information between males and females [15].

In this study, we aim to explore the biometric recognition of two-dimensional palmprints using a deep neural network like GoogLeNet. Fig 2.1 presents the block diagram of the proposed biometric system based on 2D palmprint images.

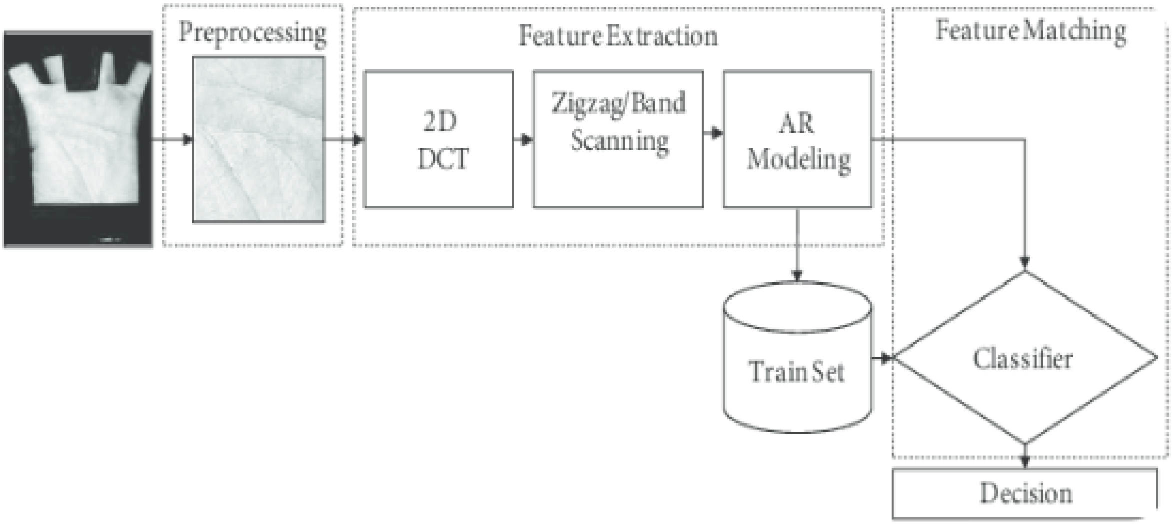


Figure 2.1: Block-diagram of the proposed biometric system[1].

2.3 GoogLeNet Deep Neural Network

Due to their outstanding performance that bypasses even the human ability in object recognition or classification, CNNs are arguably the most popular type of Deep Learning architecture [16]. The convolutional operation is the key component in CNNs. They use filters to extract features of the image, by sliding a filter over the input image, multiplying and accumulating products at every position of the input with this filter [17]. The well-known CNN architectures include AlexNet, VGGNet, GoogLeNet, ResNet, and Dense Net.

GoogLeNet developed by Google team, is the first CNN architecture that replaced the expensive fully connected layers at the end of the model with a simple global average-pooling layer, which averages out the given values of each feature map. This change has dramatically reduced the number of parameters used in the model, which made it faster in the training phase, lighter in size, and higher in performance, compared to its predecessor architectures, such as VGGNet and Alex Net [18]. GoogLeNet handles our input data in a fixed operations stack to give the desired predictions. The image begins from the first layer (input layer) and passes across multiple layers until it reaches the output layer as an output of the model.
The main structure is as follows: Layers: GoogLeNet's structure is built around several layers [19]:

- **Convolutional layer (Conv.)** is the main powerhouse of all CNNs. It performs a dot-product operation for two matrices (image and filter) to extract features from the input patterns.

- **Activation layer (Active.)** is an activation function for adding non-linearity to the Conv. layers.
- **Normalization layer (Norm.)** is an important technique for improving the performance. It is used to normalize the inputs by adjusting and scaling them.
- **Max-pooling layer (Max./ Avg.)** for down-sampling the feature maps, as its name suggests, it takes the maximum from the window. It is used to overcome overfitting problems.
- Finally, **Softmax layer** is the classifier of the model.

2.4 GoogLeNet Architecture

GoogLeNet is a popular pre-trained deep convolutional neural network (CNN) architecture with 22 developed layers to outperform the existing CNN architectures. The credit for the tremendous performance improvement goes to the **IM** which permits multiple convolutions with multiple kernels and max pooling to take place simultaneously within a single layer ensuring that the network trains with optimal weights and selects more useful features [20]. To do this, every inception layer encompasses within itself, variable size convolutional kernels namely 1x1, 3x3, and 5x5, and an additional 3x3 max pooling to capture more discriminative features from the pattern passed from the previous layer. The architecture of **IM** is also shown in Fig 2.2.

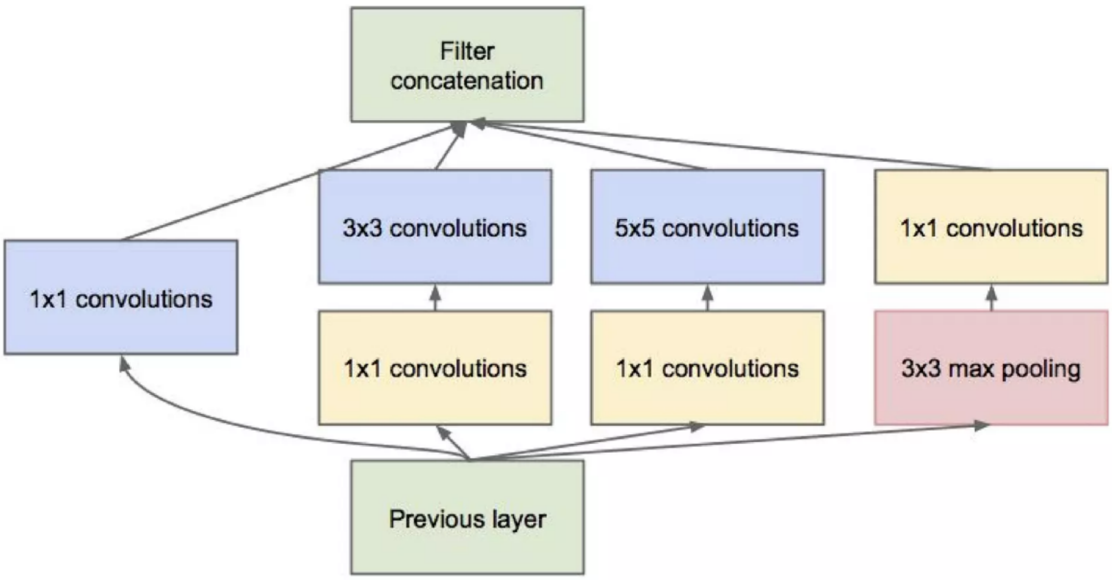


Figure 2.2: GoogLeNet Architecture.

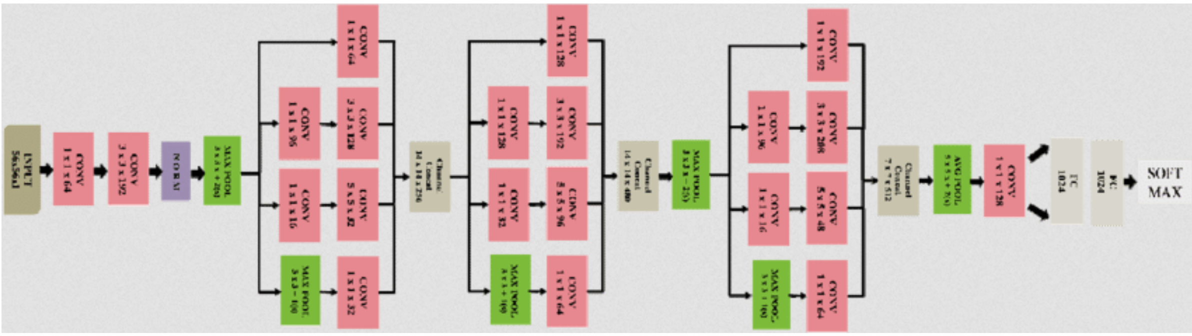


Figure 2.3: GoogLeNet Architecture.

As shown in Fig 2.3, Nine IMs in the GoogLeNet architecture effectively extract the more discriminative features from the source image. Wherein the 1x1 filters along with the max pooling layer perform the dual task of dimensionality reduction and summarizing the contents from the previous layer.
The basic process of how the GoogLeNet model works is this:

1. First, the input image is passed to the first Conv. Layer, where filters are used with different sizes (3x3 and 1x1) to generate feature maps.
2. Then, this feature map will be passed to the Norm. layer and Activ. Layer subsequently.
3. Next, the obtained output is passed to a Max. layer. It is worth mentioning that every Conv. Layer is followed by Norm.
4. And then Activ. Layer, whereas only some Conv. layers are followed by a Max. Layer. This block is repeated several times until the process reaches Avg. Layer.
5. Finally, the SoftMax Layer to solve the multi-class classification problem.

2.5 GoogLeNet Feature Extraction

GoogLeNet utilizes a variety of filters to extract features from images. These filters include Edge Filters, Corner Filters, Texture Filters, Effect Filters, Repetition Filters, Color Filters, and Shape Filters. These filters are used synchronously in GoogLeNet to extract a comprehensive set of features from images, allowing for more efficient representation and usage in tasks such as classification or object detection [18].

- **Convolution Operation:** This refers to the process of passing filters over an image gradually, where the values of the filters are multiplied by the pixel values in specific regions of the image. The products are calculated and summed up at a single point.
- **Filtering:** Filters are used to detect various patterns in the image. For example, a filter may represent an edge in the image, and this filter is used to detect edges in

the image.

- **Transformations:** Filters can transform the image in different ways. For example, approximation filters can perform transformations to ensure the reduction of the image size.
- **Learned Features:** In deep neural networks, the image is processed through several layers of convolutional filters. Each layer applies a specific set of filters to extract certain features. These filters are learned during the model training based on the data.

In summary, convolutional filters are essential tools for feature extraction and analysis from images. They are used in filtering and transformation operations to achieve a comprehensive representation of the image [8] (show Fig 2.4).

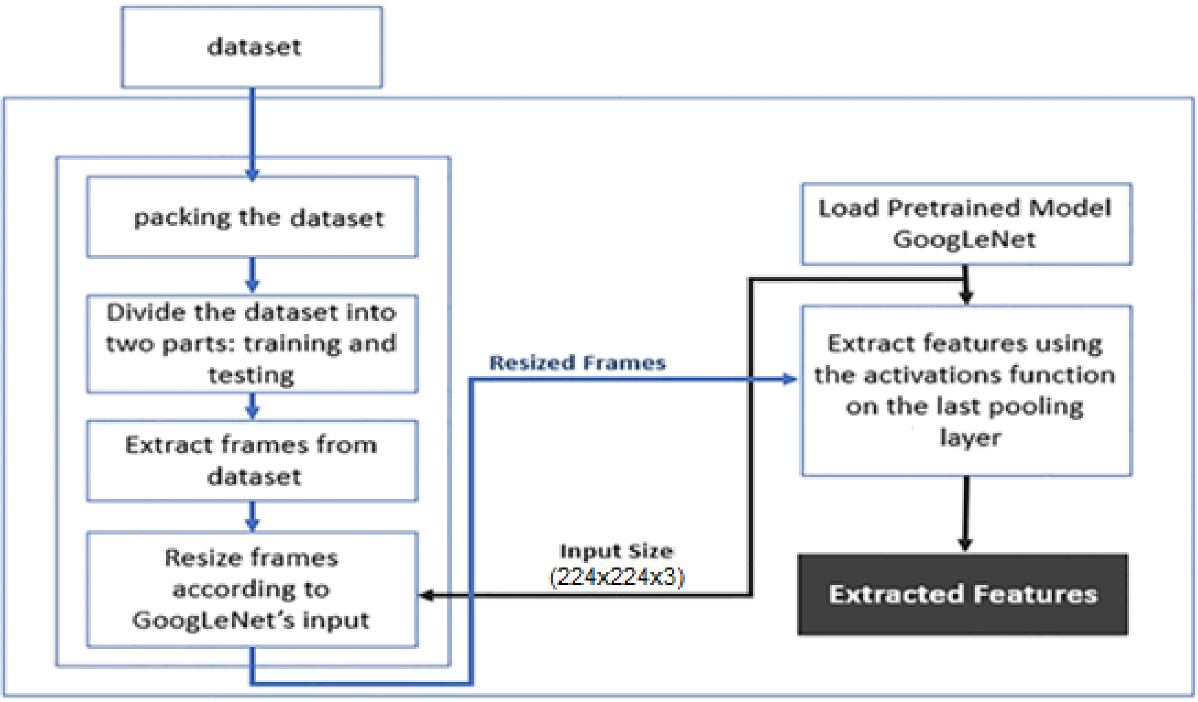


Figure 2.4: GoogLeNet features extraction steps [2].

2.6 GoogLeNet classification

The classification process using GoogLeNet involves inputting images into the model, where the images are passed through the layers of the neural network to extract distinctive features. The model is trained on a large dataset of images to enable it to effectively recognize patterns and information within the images. After the model is trained, it can be used to classify new images into appropriate categories. This is done by analyzing the features extracted from the images and comparing them to the patterns present in the training data. The classification process of GoogLeNet is presented in Fig 2.5.

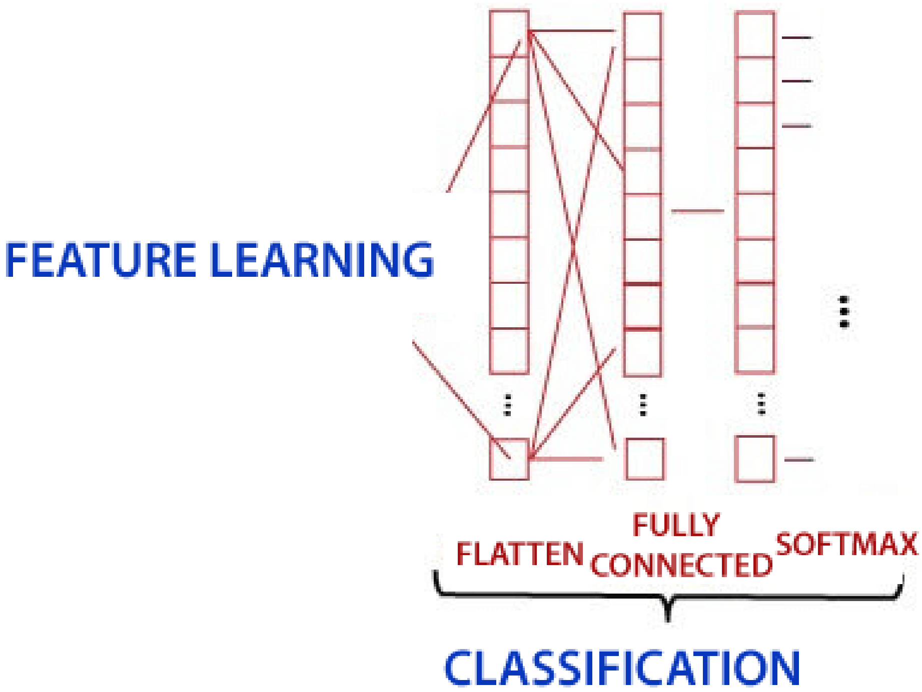


Figure 2.5: GoogLeNet features classification process.

Based on GoogLeNet classification, high accuracy in image classification can be achieved, each image is associated with a ground truth category, and performance is measured based on the classifier predictions that receive the highest scores. Typically, two numbers are reported: the accuracy rate, which compares the ground truth against the first predicted class, and the error rate, which compares the ground truth against the first predicted classes. An image is considered correctly classified if the ground truth is among the top, regardless of its rank within them [18].

2.7 Feature fusion

Depending on the type of data fusion can be classified into two major groups: fusion before matching and fusion after matching. Fusion before matching (fusion pre-classification) can take place either at the image level or at the feature level. Fusion after matching (fusion post-classification) can be divided into two: score levels or the decision level.

The raw biometric data represented the richest source of information although it is expected to be contaminated by noise. For image level fusion, we used Discrete Wavelet

Transform (DWT) [21]. Where the idea of this transform is to decompose a signal into different resolutions. The decomposition of the image is performed by two filters into sub-bands: high-pass and HA low-pass LA. The reconstruction of the image also called synthesis requires the use of two filters (HS, LS). Image Fusion based on DWT decomposition (Approximation, Horizontal detail, Vertical detail, Diagonal detail) is presented in Fig 2.6. Suppose M1 and M2 are two images, the fusion steps are [21]:

1. DWT (M1) = [A1, V1, H1, D1] and DWT (M2) = [A2, V2, H2, D2]
2. The averages of all sub-bands is: [A, G, H, D] = ([A1, V1, H1, D1] + [A2, V2, H2, D2]) / 2
3. In the end, apply the inverse DWT on the results obtained: M = IDWT ([A, V, H, D])
4. The image M is the fusion of two pictures M1 and M2.

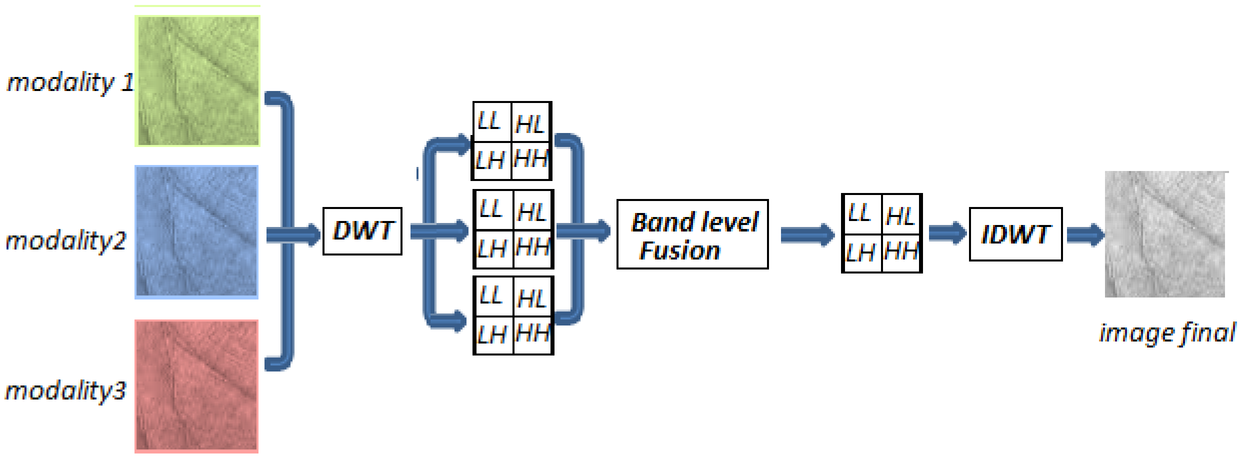


Figure 2.6: Image level fusion based on DWT.

2.8 Conclusion

In this chapter, we introduce a system for two-dimensional palmprint recognition using a deep neural network of the GoogLeNet type. The basic structure of GoogLeNet and how features are extracted and classified using it were explained. Additionally, the advanced feature fusion technique that was used to merge the extracted features was discussed. We hope this work represents a significant step towards improving the accuracy of palmprint recognition. In the following chapters, we will present the results of the experiments to evaluate the proposed system's performance and compare it with other available systems.

Chapter

3

Expermentations and Resaults

3.1 Introduction

In this chapter, we delve into the experimental phase of our study, where we rigorously test and evaluate the performance of the proposed biometric system. The primary objective is to assess the effectiveness and reliability of the system in accurately identifying individuals based on their unique biological traits.

Biometric systems have gained substantial attention in recent years due to their potential to provide secure and convenient authentication solutions across various domains. To evaluation of robustness, we conduct a series of experiments designed to analyze different aspects of the biometric system. We begin by providing an overview of the used database, followed by the selection of pertinent parameters crucial for system optimization. Subsequently, we present the evaluation results of both unimodal and multimodal biometric systems. It is noteworthy that: The hardware that we used to conduct this experiment was very limiting. The specification is listed in the following:

- Computer: HP ProBook 450 g5.
- CPU: Intel(R) Core(TM) i3-5300U CPU @ 2.30GHz 2.40 GHz.
- RAM: 12,00 Go DDR3.

The software environment used to conduct this experiment is as follows:

- Integrated Development Environment (IDE): Math Works MATLAB 2021a.
- Operating system: Windows 10 Pro 64 bits, processor x64.

3.2 Database Description

To carry out the image processing experiment, we require a digital image. The digital image can be inputted, related experiments can be carried out, and related results can be formed. Such an image database is provided to the researcher to explore the related scientific research. The Hong Kong Polytechnic University (PolyU) 2D_Palmprint Database [22] is an existing palm print image database made available worldwide to advance research in the area of biometric palm print recognition systems. The Biometric Research Centre (UGC/CRC) at Hong Kong Polytechnic University [22] has developed a palm print database to advance research and to provide researchers working in the area of palm print recognition with a platform to compare the effectiveness of palm print recognition algorithms.

Although 2D palm print recognition can achieve high accuracy [22], the PolyU 2D Palmprint Database contains 7680 samples collected from 384 different palms. Twenty samples from each of these palms were collected in two separate sessions, where 10 samples were captured in each session, respectively. The average time interval between the two sessions is one month. All palm print images are of the same size and dimension, such as 384 X 284 [22]. Each 2D palm print image was recorded in a BMP format image file. The palm print images have a name sequence and can be interpreted as follows: e.g., a palm print image named "`PolyU_001_F.bmp`" can be interpreted as the initial word PolyU is the copyright for the Polytechnic University of Hong Kong. Then, "`001`" indicates the subject enrollment number as it varies from person to person. The following "`F`" indicates the session enrollment, whether it is the first or second. And finally, the format BMP of the image file.

3.3 Parameters Selection

The performance of the GoogLeNet algorithm depends on the good tuning of deep learning parameters.GoogLeNet has generic parameters, such as the number of layers, number of filters and block sizes, etc. To select GoogLeNet algorithm parameters, we need to do a series of experiments that provide the required parameters for the best performance (minimum error rates).

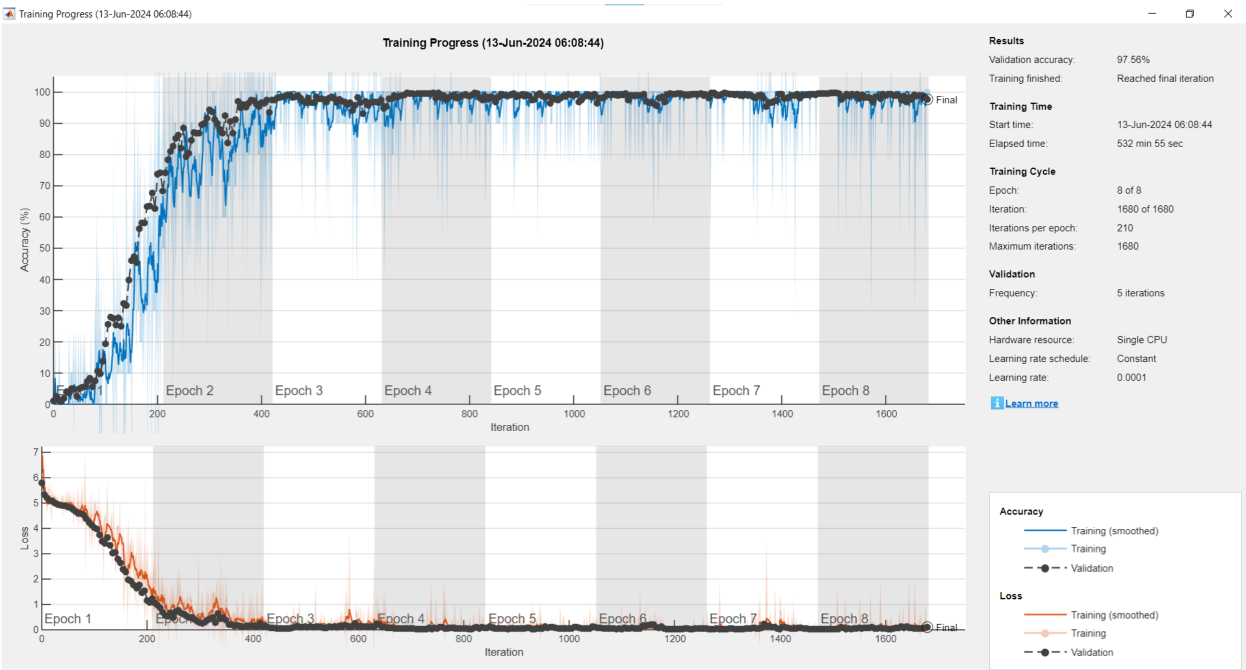
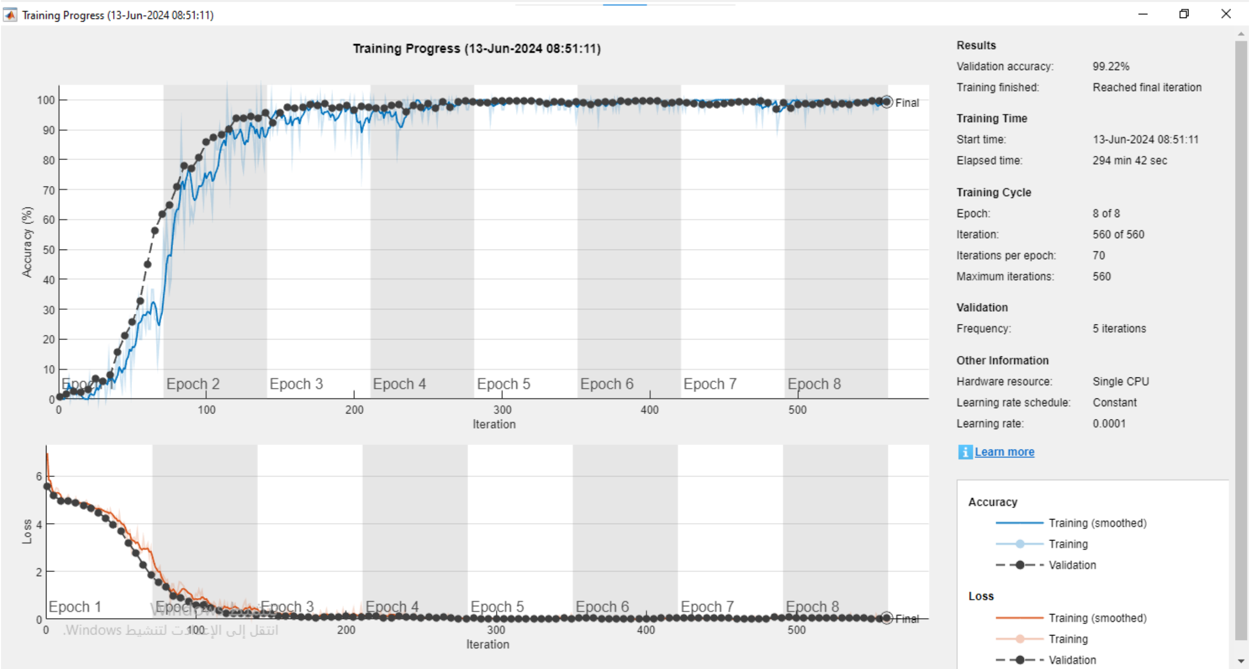
3.3.1 Selection of Mini batch size

Mini batch size is used to create, pre-process, and manage mini-batches of data for training using custom training loops, convert data to a different precision, or apply a custom function to pre-process your data. Firstly, to determine the 'mini-batch size' configurations in our approach, we describe the sub-results related to the proposed Mini Batch Size configuration parameter. When using different numbers of configurations such as 10, 30, 40, and 50 for each test. With save of default parameters: 'Epochs equal 8' and 'Initial Learn Rate equal 10^{-4}'. In Table 3.1, we present the test results of our palmprint recognition systems.

Number	GAR(%)	Time	EER(%)
10	97.55	532min 55sec	2.45
30	99.22	294min 42sec	0.78
40	99.89	235min 53sec	0.11
50	99.56	236min 5sec	0.44

Table 3.1: Results for training different sizes of GoogLeNet.

From Table 3.1, it is evident that the set of four configurations for mini-batch size provides better results in terms of GAR. In this case, the 10-size mini-batch size for GoogLeNet can achieve a GAR of 97.55% at a time T= 532min 55s. Additionally, from this table, we can observe that the configuration of 30 offers a GAR of 99.22% at a time T = 294min 42s. Furthermore, using the configuration of 40 that offers a GAR of 99.89% at a time T =235min 53s. Finally, using the configuration of 50 for GoogLeNet that yields a GAR of 99.56% at a time T = 236min 5s. The curves for the four GoogLeNet configurations of mini-batch size are shown in Figures 3.1, where the Accuracy (GAR) is plotted against the Iteration (epochs).


Training Progress (13-Jun-2024 06:08:44)
Results
Validation accuracy: 97.56%
Training finished: Reached final iteration
Training Time
Start time: 13-Jun-2024 06:08:44
Elapsed time: 532 min 55 sec
Training Cycle
Epoch: 8 of 8
Iteration: 1680 of 1680
Iterations per epoch: 210
Maximum iterations: 1680
Validation
Frequency: 5 iterations
Other Information
Hardware resource: Single CPU
Learning rate schedule: Constant
Learning rate: 0.0001
Learn more
Accuracy
Training (smoothed)
Training
Validation
Loss
Training (smoothed)
Training
Validation
Accuracy (%)
Loss
Iteration
Final

Training Progress (13-Jun-2024 08:51:11)
Results
Validation accuracy: 99.22%
Training finished: Reached final iteration
Training Time
Start time: 13-Jun-2024 08:51:11
Elapsed time: 294 min 42 sec
Training Cycle
Epoch: 8 of 8
Iteration: 560 of 560
Iterations per epoch: 70
Maximum iterations: 560
Validation
Frequency: 5 iterations
Other Information
Hardware resource: Single CPU
Learning rate schedule: Constant
Learning rate: 0.0001
Learn more
Accuracy
Training (smoothed)
Training
Validation
Loss
Training (smoothed)
Training
Validation
Accuracy (%)
Loss
Iteration
Final

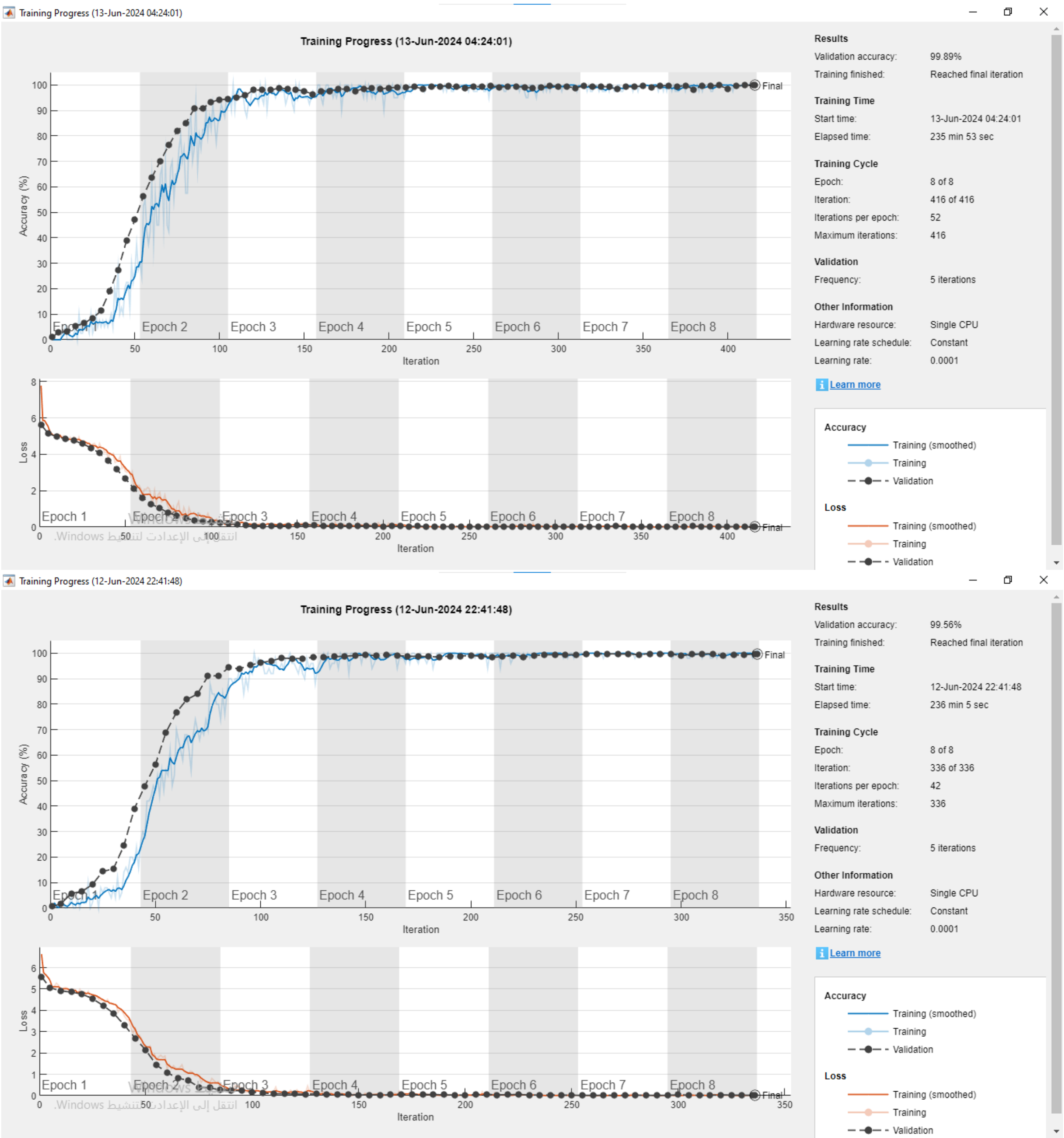


Figure 3.1: The results of GoogLeNet for different mini-batch size parameters.

Therefore, the system can achieve the best accuracy with the configuration of **40 mini-batch sizes** for GoogLeNet compared to the other configurations, which produces a GAR of 99.89% with an EER of 0.11

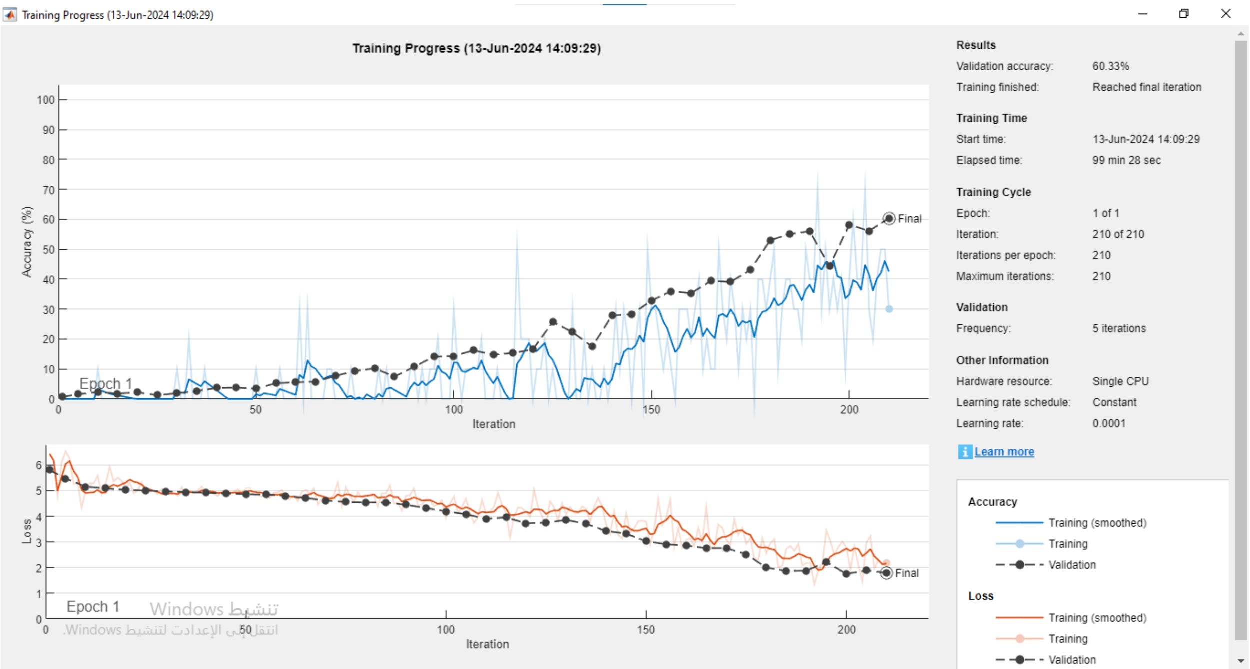
3.3.2 Selection of Max eboch

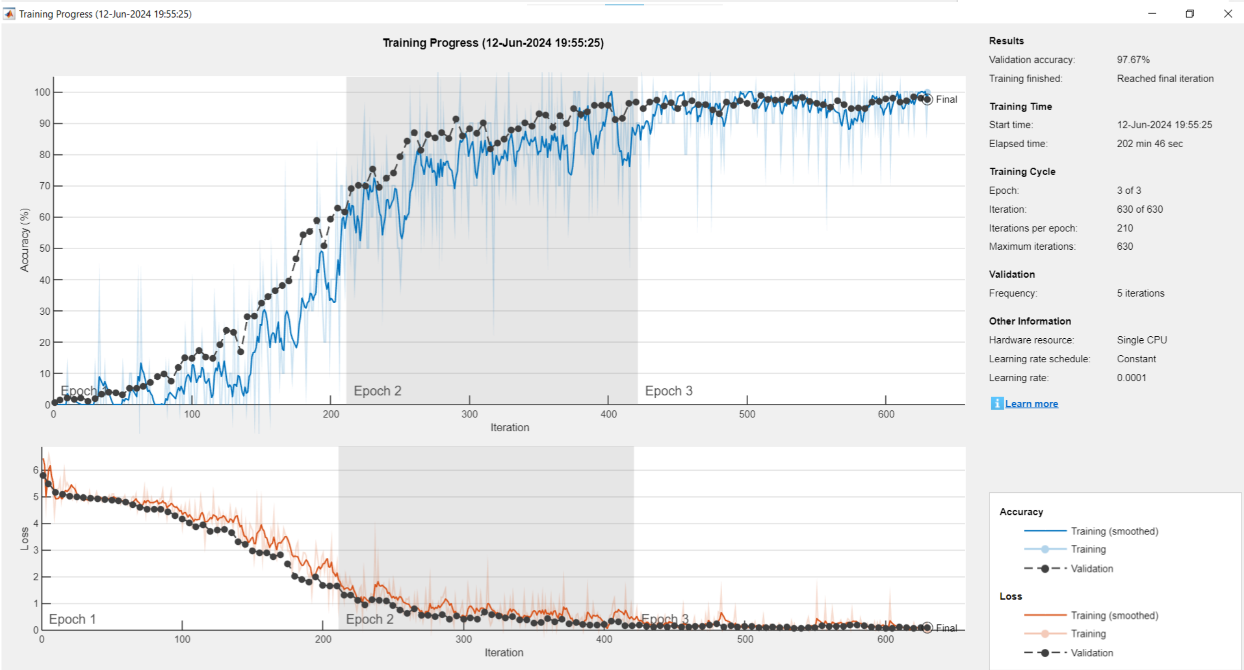
An epoch is defined as the number of times an algorithm visits the data set. In other words, an epoch is one backward and one forward pass for all the training. Secondly, to select the number of epochs, this subsection describes the results of the proposed epochs number parameter. When, we used a different probability of numbers such as 1, 3, 6, and 8 of each test, with saved other default parameters: 'mini-batch size equal to 10' and 'Initial Learn Rate equal to 10^{-4}'. Thus, Table 3.2 presents the test results of the epochs parameter for our systems.

Number	GAR(%)	Time	EER(%)
1	60.33	99min 28sec	39.67
3	97,67	202min 46sec	2.33
6	99,00	393min50sec	1.00
8	97.55	532min 55sec	2,45

Table 3.2: Results of different epochs number parameters.

From this Table 3.2, it is evident that the set of four configurations for epochs provides better results in terms of GAR. In this case, the 1 number of epochs for GoogLeNet can achieve a GAR of 60.33% at a time T = 99min 28s. Additionally, from this table, we can observe that the configuration of 3 offers a GAR of 97.67% at a time T = 202min 46s. Furthermore, using the configuration of 6 that offers a GAR of 99.00% at a time T = 393min 50s. Finally, using the configuration of 8 that yields a GAR equal to 97.55% at a time T = 532min 55s. The curves of the four cases of GoogLeNet epoch configurations are shown in Figure 3.2, where the Accuracy (GAR) is plotted against the Iteration (epochs).


Training Progress (13-Jun-2024 14:09:29)
Results
Validation accuracy: 60.33%
Training finished: Reached final iteration
Training Time
Start time: 13-Jun-2024 14:09:29
Elapsed time: 99 min 28 sec
Training Cycle
Epoch: 1 of 1
Iteration: 210 of 210
Iterations per epoch: 210
Maximum iterations: 210
Validation
Frequency: 5 iterations
Other Information
Hardware resource: Single CPU
Learning rate schedule: Constant
Learning rate: 0.0001
Learn more
Accuracy
Training (smoothed)
Training
Validation
Loss
Training (smoothed)
Training
Validation
Accuracy (%)
Loss
Iteration
Epoch 1
Final


Training Progress (12-Jun-2024 19:55:25)
Results
Validation accuracy: 97.67%
Training finished: Reached final iteration
Training Time
Start time: 12-Jun-2024 19:55:25
Elapsed time: 202 min 46 sec
Training Cycle
Epoch: 3 of 3
Iteration: 630 of 630
Iterations per epoch: 210
Maximum iterations: 630
Validation
Frequency: 5 iterations
Other Information
Hardware resource: Single CPU
Learning rate schedule: Constant
Learning rate: 0.0001
Learn more
Accuracy
Training (smoothed)
Training
Validation
Loss
Training (smoothed)
Training
Validation
Accuracy (%)
Loss
Iteration
Epoch 1
Epoch 2
Epoch 3
Final

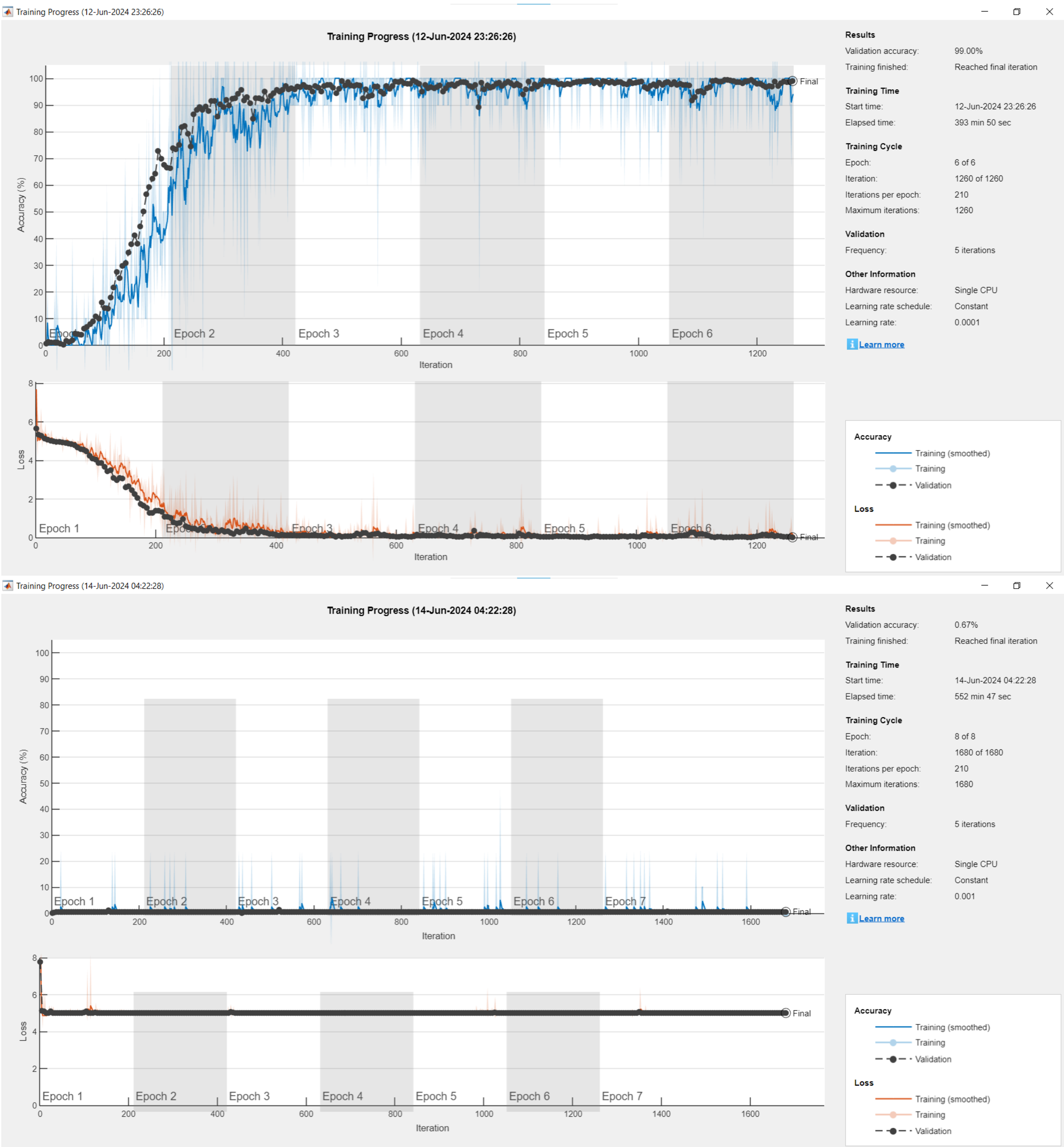


Figure 3.2: The results of GoogLeNet for different Max Eboch parameters.

Therefore, the system can achieve the best accuracy with the configuration of 6 epochs for GoogLeNet compared to the other configurations, which produce a GAR of 99.00% with an EER of 1.00%.

3.3.3 Selection of initial Learn Rate

The initial learning rate used for training is specified as a positive scalar. If the learning rate is too low, then training can take a long time. If the learning rate is too high, then training might reach a sub-optimal result or diverge. Thirdly, to select the last parameter (the learning rate), this subsection describes the results of the proposed parameter. When, we use a different learning time of rates as $10^{-3}, 10^{-4}, 10^{-5} and 10^{-6}$, with also save other default parameters: 'Epochs equal to 6' and 'mini-batch size equal to 10'. Thus, Table 3.3 presents the test results of the learning rate parameter for our recognition systems.

Number	GAR(%)	Time	EER(%)
10^{-3}	67,00	552min 47sec	33,00
10^{-4}	97,55	532min 55sec	2,45
10^{-5}	99,00	804min 6sec	1,00
10^{-6}	15,56	526sec41min	84,44

Table 3.3: Results of different Initial learn rate number parameters.

From Table 3.3, it is evident that the set of four configurations for the Initial learning rate provides better results in terms of GAR. In this case, the rate of 10^{-4} of the Initial learning rate for GoogLeNet can achieve a GAR of 97.55% at a time T = 532min 55s. Additionally, from this table, we can observe that the configuration of rate 10^{-3} offers a GAR of 67.00% at a time T = 552min 47s. Furthermore, using the configuration of rate 10^{-5} that offers a GAR of 99.00% at a time T = 804min 6s. Finally, using the configuration of rate 10^{-6} that yields a GAR equal to 15.56% at a time T = 526min 41s. The curves of the four cases of GoogLeNet epochs configurations are shown in Figure 3.3, where the Accuracy (GAR) is plotted against the Iteration (epochs).

But in this case, we will put the curves related to the best values.

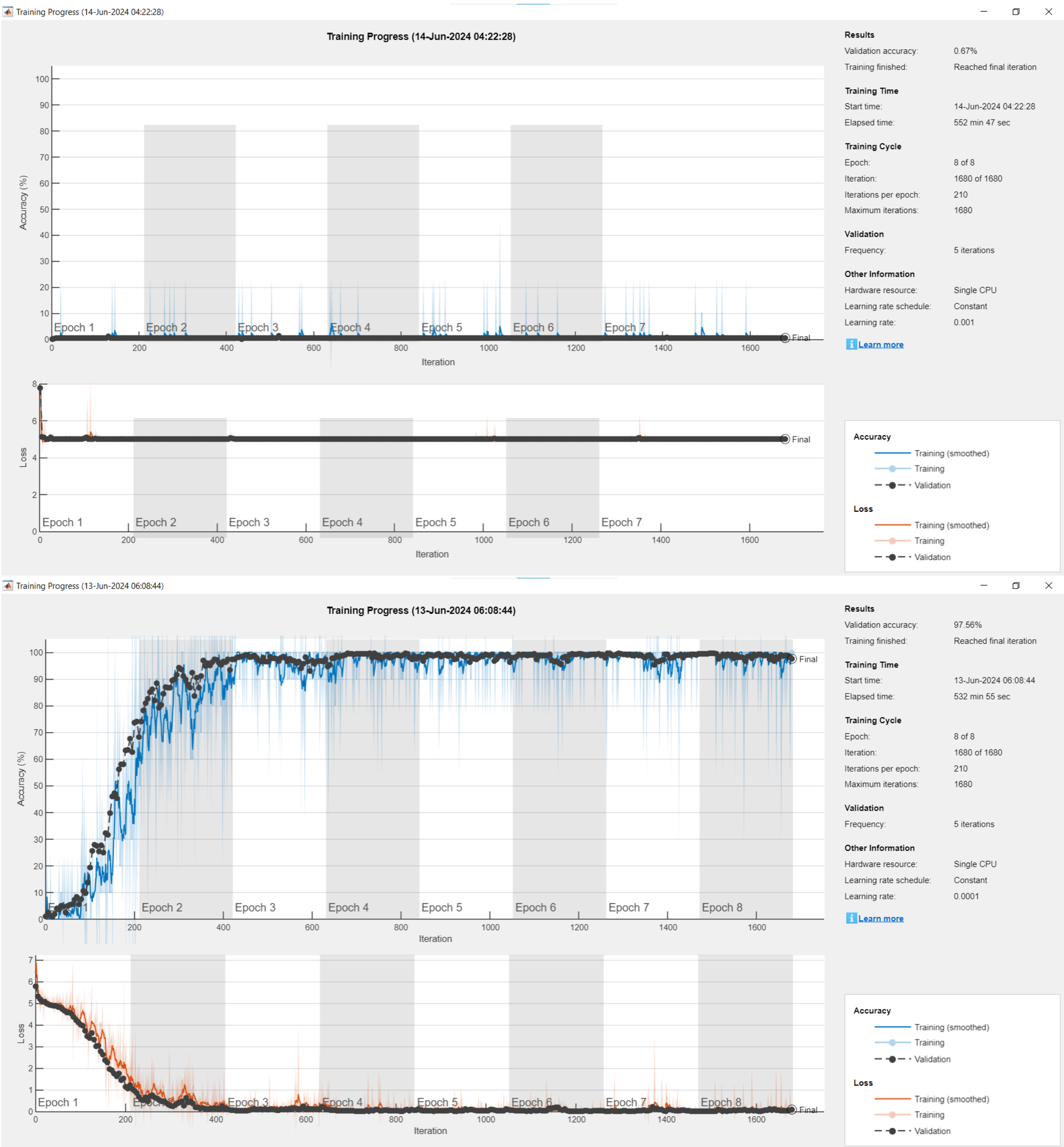


Figure 3.3: The results of GoogLeNet for different initialLearnRate parameters.

Therefore, the system can achieve the best accuracy with the configuration of 1e-5 Initial learning rate of GoogLeNet compared to the other configurations, which produces a GAR equal to 99.00% with an EER equal to 1.00%.

3.4 Biometric System Evaluation

Depending on the preceding results, the GoogLeNet algorithm can be set to the following parameters: the mini-batch size equal to 40, the epochs number is 6, also the initial learning rate is 10^{-5}. Therefore, we have decided to choose these parameters for the rest of the test study.

3.4.1 Obtained Results of Unimodal Systems

In unimodal biometrics, theoretically, it might be very proficient but in reality, it has various numbers of challenges when enrolling a large number of people.
For an unimodal system, we will use the parameters we have already determined for the accuracy of its results, which is a GAR equal to 92.00% with an EER equal to 8.00% at a Time equal to 126min 43s. The accuracy results of unimodal systems based on GoogLeNet are described in Figure 3.4.

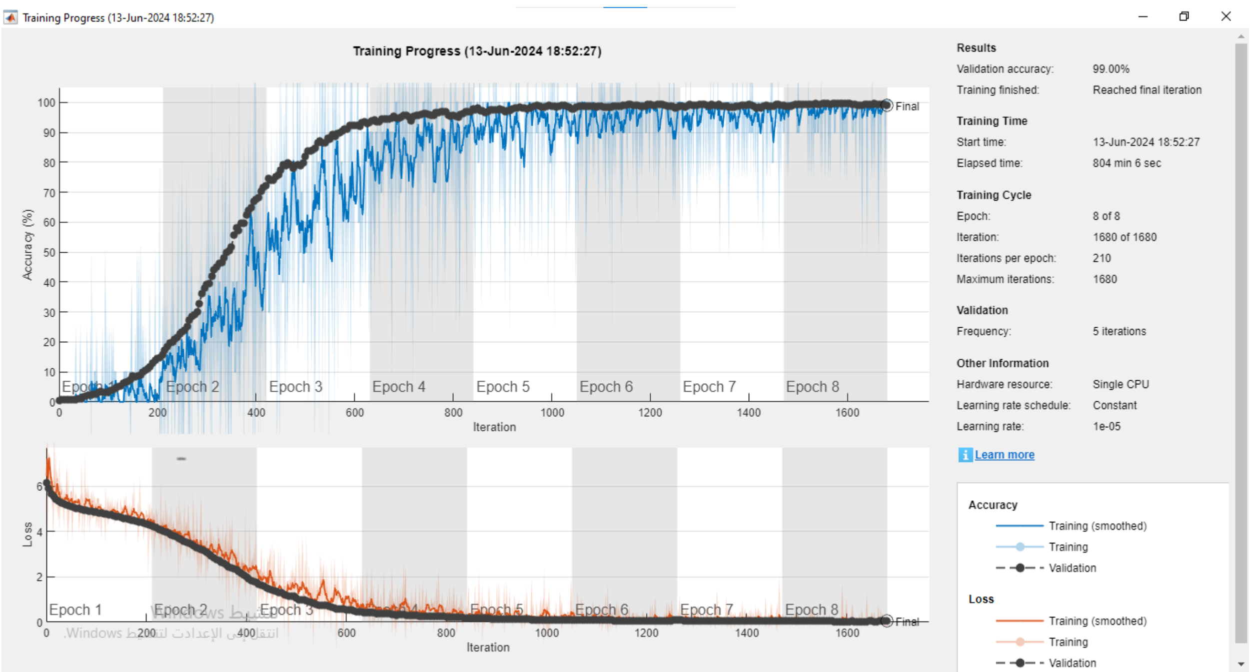


Figure 3.4: The Unimodal results of GoogleNet algorithm.

Also, for unimodal tests, we will test the GoogLeNet algorithm by predicting a person's palmprint to see if the prediction is correct. The performances of some random tests are presented in Figure 3.5. The results show that the prediction is correct for our tests.

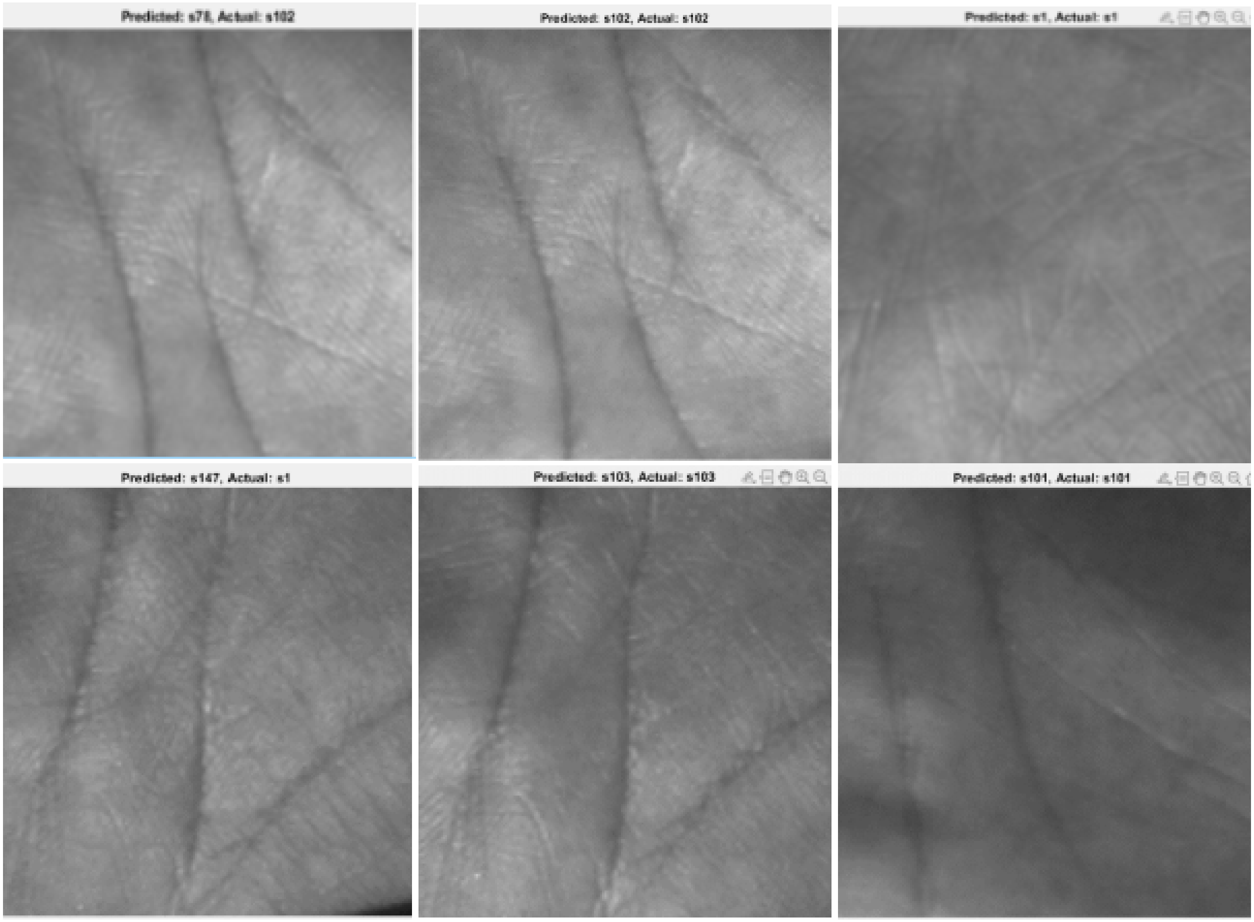


Figure 3.5: The unimodal tests prediction results of GoogleNet algorithm.

3.4.2 Obtained Results of Multimodal Systems

In multimodal systems, biometric information can be fused at several different levels. In our work, we only chose the image level fusion, and we used an efficient image fusion algorithm, with the help of 2D-DWT techniques. For multimodal, we will use the same configuration parameters we have already determined, from the results which are a GAR equal to 91.00% with an EER equal to 9.00% at a Time equal to 128min 32s. The accuracy results of multimodal systems based on the AlexNet algorithm are described in Figure 3.6.

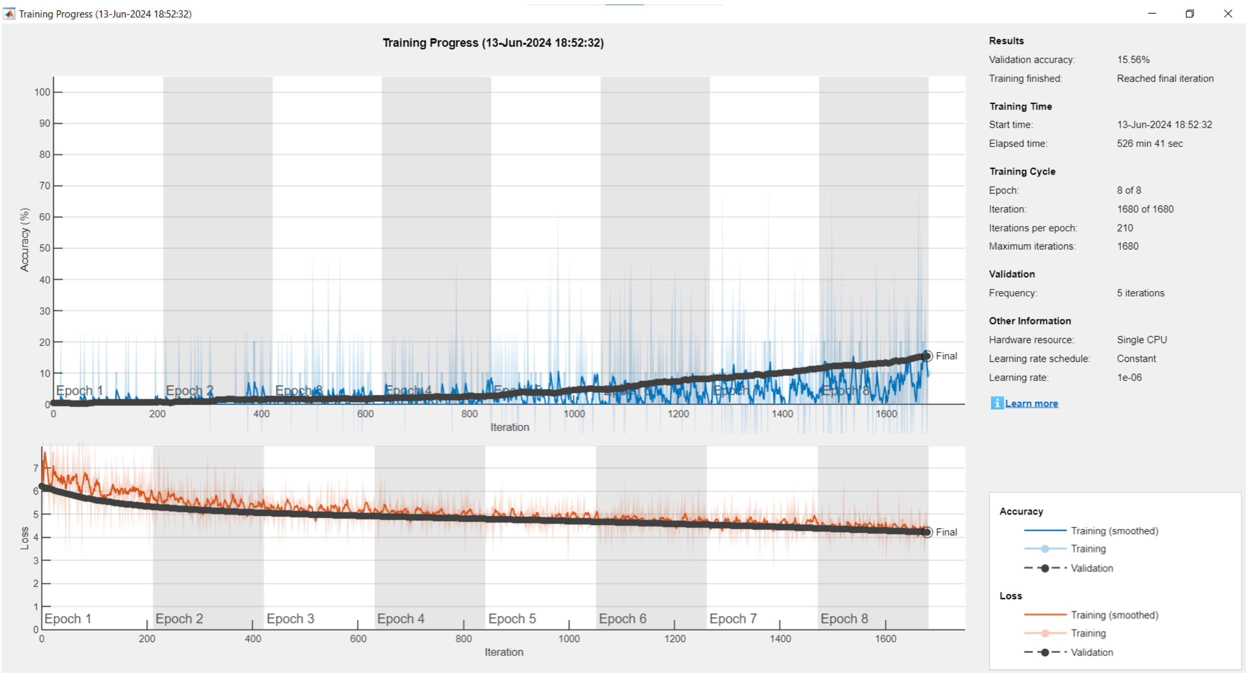


Figure 3.6: The multimodal test results of GoogLeNet algorithm.

Also, for multimodal tests, we will test the GoogLeNet algorithm with an efficient image fusion algorithm by predicting a person's palmprint images. The results show that the prediction is correct for our tests, the performances of some random tests are presented in Figure 3.7.

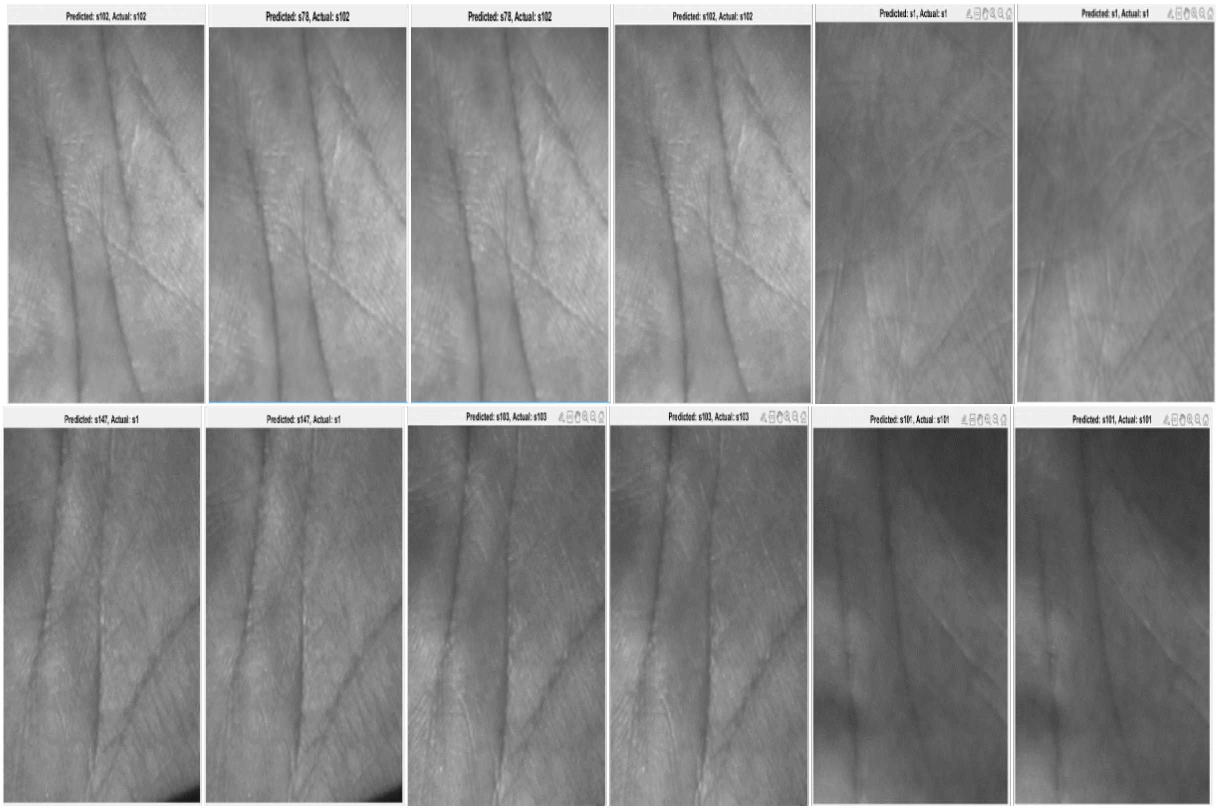


Figure 3.7: The multimodal tests prediction results of GoogLeNet algorithm.

The purpose of this study is to improve the performance and effectiveness of recognition and classification systems and provide biometric identification based on palmprint recognition. The obtained results showed an accuracy rate equal to 92.00% for the unimodal and an accuracy rate equal to 91.00% for multimodal systems, respectively.

3.5 Conclusion

In conclusion, this chapter presents a comprehensive exploration of four key experiments and their results. It begins with an introduction, followed by a detailed description of the database utilized. The process of parameter selection is then outlined, leading to the evaluation of the biometric performance. Results from both unimodal and multimodal systems are presented and analyzed. Finally, the chapter concludes with insights drawn from the experiments, paving the way for subsequent research and advancements in biometric systems.

General Conclusion

Deep learning can play a vital role in developing and enhancing biometric systems. In biometric systems, it is based on using individual biological or behavioral data, such as fingerprints, facial recognition, and iris scans, to verify a person's identity. Deep learning, as an advanced analytical tool, is one of the key tools that can be used to improve the accuracy and security of these systems.

In this research, we focused on the topic of palmprint recognition using the Google Net network, also known as InceptionNet, which is one of the famous models in the field. Using the GoogLeNet technique, a precise model has been developed that can identify the unique patterns and features of users based on the provided biometric palmprint images. It is also important that these models are developed in a way that takes into account privacy and security issues, ensuring that sensitive data is handled securely and that systems are not vulnerable to hacking or exploitation.

To this study, the related results were achieved especially when we changed the batch sizes. The best value turned out to be 40 for the batch size. For Max epochs and Initial Learn Rate, it was 6 and it was 10^{-5}, respectively. In unimodal biometrics, it may be theoretically very proficient but in reality, it faces a different number of challenges when registering a large number of people. The obtained results showed an accuracy rate equal to 92% for unimodal systems and an accuracy rate equal to 91% for multi-modal systems. The technology of handprint recognition using GoogLeNet or other deep neural networks faces challenges and potential future developments:

Challenges

- **Privacy:** There is a need to enhance security measures to protect sensitive user data.

- **Recognition Accuracy:** Continuous improvement is required to ensure correct and rapid recognition of users' handprints.
- **Compatibility and Integration:** Seamless integration with access control systems and other applications is essential.

Potential Developments

- **Broader Applications:** Handprint recognition could find diverse applications in sectors such as security, healthcare, e-commerce, and beyond.
- **Artificial Intelligence and Machine Learning:** Ongoing advancements in deep learning models and the use of AI can enhance performance.
- **Decentralized Control:** Evolution towards self-recognition systems and secure decentralized access control is anticipated.

Bibliography

[1] Block diagram of the proposed palmprint-based recognition system. URL https://www.researchgate.net/figure/Block-diagram-of-the-proposed-palmprint-based-recognition-system_fig2_299593831. Accessed: 18 May, 2024.

[2] Deep learning and handcrafted features for one-class anomaly detection in uav video - scientific figure on researchgate, 2024. URL https://www.researchgate.net/figure/GoogLeNet-features-extraction-Steps_fig7_344389627. Accessed: 27 May, 2024.

[3] Anil K. Jain, Arun Ross, and Sharath Pankanti. Biometrics: A tool for information security. *IEEE Transactions on Information Forensics and Security*, 1(2):125–143, 2006.

[4] Anil K. Jain, Arun Ross, and Salil Prabhakar. An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Image- and Video-Based Biometrics*, 14(1):4–20, 2004.

[5] Tony Dunstone and Neil Yager. *Biometric System and Data Analysis: Design, Evaluation, and Data Mining*. Springer, 2006.

[6] R. Chlaoua. *Combination of Multiple Biometrics for Recognition of Persons*. PhD thesis, University of Kasdi Merbah Ouargla, 2019.

[7] Y. Chen, S. C. Dass, and A. K. Jain. Fingerprint quality indices for predicting authentication performance. In *Proceedings of the Fifth International Conference on Audio and Video-Based Person Authentication (AVBPA'05)*, pages 160–170, 2005.

[8] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. The MIT Press, Cambridge, MA, USA, 2016. ISBN 978-0262035613.

[9] Johannes Abel and Tim Fingscheidt. A dnn regression approach to speech enhancement by artificial bandwidth extension. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pages 219–223. IEEE, 2017.

[10] Theodoros Kotsiopoulos, Panagiotis Sarigiannidis, Dimitrios Ioannidis, and Dimitrios Tzovaras. Machine learning and deep learning in smart manufacturing: the smart grid paradigm. *Computer Science Review*, 40:100341, 2021.

[11] Abhinav Gupta, Alagan Anpalagan, Ling Guan, and Ahmad S. Khwaja. Deep learning for object detection and scene perception in self-driving cars: survey, challenges, and open issues. *Array*, page 100057, 2021.

[12] S. and M. Kumar. A comprehensive survey on the biometric recognition systems based on physiological and behavioral modalities. *Expert Systems with Applications*, 143, 2020. doi: 10.1016/j.eswa.2019.113114.

[13] A. Kumar and D. Zhang. Personal authentication using multiple palmprint representation. *Pattern Recognition*, 38(10):1695–1704, Oct. 2005.

[14] Mohamed Himeur. Palmprint recognition using deep learning, 2021.

[15] Z. Xie, Z. Guo, and C. Qian. Palmprint gender classification by convolutional neural network. *IET Computer Vision*, 12(4):476–483, 2018. doi: 10.1049/iet-cvi.2017.0475.

[16] Vivienne Sze, Yu-Hsin Chen, Tien-Ju Yang, and Joel S. Emer. Efficient processing of deep neural networks: A tutorial and survey. *Proceedings of the IEEE*, 105(12): 2295–2329, 2017.

[17] F. F. D. Santos et al. Analyzing and increasing the reliability of convolutional neural networks on gpus. *IEEE T Reliab*, 68(2):663–677, 2019.

[18] Christian Szegedy et al. Going deeper with convolutions. In *2015 IEEE Conference on (CVPR)*, pages 1–9, 2015.

[19] Mustafa Al-Qizwini, Iman Barjasteh, Hayder Al-Qassab, and Hayder Radha. Deep learning algorithm for autonomous driving using googlenet. In *2017 IEEE Intelligent Vehicles Symposium (IV)*, pages 89–96, 2017.

[20] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2818–2826, Las Vegas, NV, 2016.

[21] L. Huang, H. Zhuang, and S. Morgera. A method towards biometric feature fusion. *International Journal of Biometrics*, 1(4):479–494, 2009.

[22] Harvey Motulsky. *Statistics Guide: Statistical analyses for laboratory and clinical researchers*. GraphPad Software, Inc., 1999.

Acronyms

2D Two-Dimensional

GoogLeNet Google Neural Network

CNN Convolutional Neural Network

PolyU The Hong Kong Polytechnic University

PIN Personal Identification Number

DNA Deoxyribonucleic Acid

Biometric Biometric

API Application Programming Interface

NN Neural Network

ML Machine Learning

e-banking Electronic Banking

VGGNet Visual Geometry Group Network

ResNet Residual Network

DenseNet Densely Connected Convolutional Network

ReLU Rectified Linear Unit

Max-Pooling Max Pooling

Avg-Pooling Average Pooling

Conv. Convolutional

Norm. Normalization

Activ. Activation

Softmax Softmax

IM Inception Module

DWT Discrete Wavelet Transform

IDWT Inverse Discrete Wavelet Transform

IDE Integrated Development Environment

Edge Filters Edge Detection Filters

Corner Filters Corner Detection Filters

Texture Filters Texture Analysis Filters

Effect Filters Effect Application Filters

Repetition Filters Repetition Detection Filters

Color Filters Color Analysis Filters

Shape Filters Shape Detection Filters

HS High-pass Synthesis

LS Low-pass Synthesis

GAR Genuine Acceptance Rate

EER Equal Error Rate

BMP Bitmap

CPU Central Processing Unit

RAM Random Access Memory

UGC/CRC University Grants Committee / Central Research Committee