

ANKARA YILDIRIM BEYAZIT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED
SCIENCES



HUMAN IDENTIFICATION VERIFICATION FROM
BIOMETRIC DORSAL HAND VEIN IMAGES BASED
ON DEEP LEARNING GENERATIVE ADVERSARIAL
NETWORKS

Ph.D. Thesis by

Khaled Mohamed AB ALASHIK

Department of Computer Engineering

July 2021

ANKARA

**HUMAN IDENTIFICATION VERIFICATION FROM
BIOMETRIC DORSAL HAND VEIN IMAGES BASED
ON DEEP LEARNING GENERATIVE ADVERSARIAL
NETWORKS**

A Thesis Submitted to

The Graduate School of Natural and Applied Sciences of

Ankara Yıldırım Beyazıt University

**In Partial Fulfilment of the Requirements for the Degree of Doctor of
Philosophy in Computer Engineering, Department of Computer Engineering**

by

Khaled Mohamed AB ALASHIK

July, 2021

ANKARA

Ph.D. THESIS EXAMINATION RESULT FORM

We have read the thesis entitled “**HUMAN IDENTIFICATION VERIFICATION FROM BIOMETRIC DORSAL HAND VEIN IMAGES BASED ON DEEP LEARNING GENERATIVE ADVERSARIAL NETWORKS**” completed by **KHALED MOHAMED AB ALASHIK** under the supervision of **PROF. Dr. Remzi YILDIRIM** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Ph.D.

Prof. Dr. Remzi YILDIRIM

Supervisor

Asst. Prof. Dr. Uraz YAVANOGLU

Jury Member

Asst. Prof. Dr. Gokhan GULTEKIN

Jury Member

Assoc. Prof. Dr. Gazi Erkan BOSTANCI

Jury Member

Asst. Prof. Dr. Yilmaz AR

Jury Member

Prof. Dr. Ergün ERASLAN

Director

Graduate School of Natural and Applied Sciences

ETHICAL DECLARATION

I hereby declare that, in this thesis which has been prepared in accordance with the Thesis Writing Manual of Graduate School of Natural and Applied Sciences,

- All data, information, and documents are obtained in the framework of academic and ethical rules,
- All information, documents, and assessments are presented in accordance with scientific ethics and morals,
- All the materials that have been utilized are fully cited and referenced,
- No change has been made on the utilized materials,
- All the works presented are original,

and in any contrary case of the above statements, I accept to renounce all my legal rights.

Date: July, 2021 **Signature** :.....

ACKNOWLEDGMENT

First of all, I would like to thank ALLAH Almighty for giving me the strength, ability, and determination to complete this thesis.

I hereby extend my sincere thanks to my study director, **Prof. Dr. Remzi YILDIRIM**, of the Graduate School of Natural and Applied Sciences, of Ankara Yıldırım Beyazıt University, for his patience, kindness, and support, and for his serious efforts, valuable guidance, and continuous support throughout the duration of the study.

I also extend my thanks and gratitude to the **members of my thesis committee** for their guidance, invitation, and valuable suggestions for the completion of this thesis.

On this occasion, I do not forget to have mercy on the souls of my dear **parents**, may God forgive them.

Moreover, I would like to extend my sincere thanks and gratitude to my **brothers and sisters** for their continuous support and advice.

Most of all, I will be forever grateful to my **family**, especially my **wife** and my dear **children** for their support, strength, and motivation. Many thanks to all **colleagues and friends** for the help all the time.

July 2021

Khaled Mohamed AB ALASHIK

HUMAN IDENTIFICATION VERIFICATION FROM BIOMETRIC DORSAL HAND VEIN IMAGES BASED ON DEEP LEARNING GENERATIVE ADVERSARIAL NETWORKS

ABSTRACT

For the purpose of this research, biometric hand vein technique was used to recognize individual identity. In this study, authentication was done using the hand “dorsal venous network” vein. For this, a new method DL-GAN has been developed by combining deep learning and productive adversarial network (GAN). With the DL-GAN authentication method, the authentication rate has been increased.

Many biometric methods, including vein information on the hand, are used to identify people. Information on how the superficial subcutaneous shallow vein, which can be detected by various technological methods such as infrared camera, is placed in the skin tissue is a new method of identifying people. Vein locations on the hand are idiosyncratic and are a good new option for identifying individuals. The vascular system on the upper hand and near the wrist, the subcutaneous artery on the upper hand, the veins and the vascular network (metecarpal, venous network, basilic) are used to authenticate.

In order to recognize the biometric hand vein image, appropriate software coding has been made for the MATLAB 2020a programming language. The developed DL-GAN method has been tested on two separate databases, Jilin University – hand held database and 11K hand held database. The results of the experiments performed on the hand-held vein data set, on the other hand, show that the DL-GAN method has reached 98.36% accuracy and has an error rate of 2.47% and a standard accuracy of 0.19%. The experimental results in the second data set, on the other hand, have an accuracy of 96.43%, an equal error rate of 3.55% and a standard accuracy of 0.21%. The improved DL-GAN method obtained better results from biophysical methods such as LBP, LPQ, GABOR, FGM, BGM and SIFT compared to the same databases.

Keywords: Identification, biometrics, detection system, dorsal hand veins, deep learning, generative adversarial network.

DERİN ÖĞRENME ÜRETKEN KARŞILIK AĞLARINA DAYALI BİYOMETRİK DORSAL EL DAMAR GÖRÜNTÜLERİNDEN İNSAN TANIMLAMA DOĞRULAMASI

ÖZ

Bu araştırmanın amacı, bireysel kimlik tanımak için biyometrik el üstü damar tekniği kullanılmıştır. Bu çalışmada, “dorsal venous network” damarı kullanılarak kimlik doğrulama yapılmıştır. Bunun için de derin öğrenme (deep learning) ve üretken çekişmeli ağ (GAN) birleştirilerek yeni bir metot DL-GAN geliştirilmiştir. DL-GAN kimlik doğrulama yöntemi ile kimlik doğrulama oranı artırılmıştır.

İnsanların kimliklerini tanımlamak için el üstü damar bilgileri de dahil olmak üzere pek çok biyometrik yöntem kullanılır. Kızılötesi kamera gibi çeşitli teknolojik yöntemlerle tespit edilebilen el üstü deri altı sığ damarın, el deri dokusuna nasıl yerleştirildiğine dair bilgiler, kişilerin kimliklerini tanımlamanın yeni bir yöntemidir. El üstündeki damar yerleşimleri kişilere özgü olup ve bireyleri tanımlamak için kullanılan iyi ve yeni bir seçenektir. El üstü ve bileğe yakın olan damar sistemi, elin üstündeki deri altı atar damar, toplar damar ve damar ağ sistemi (metecarpal, venous network, basilic) kimlik doğrulamak için kullanılır.

Biyometrik el üstü damar görüntüsünü tanımak için MATLAB 2020a programlama dili için uygun yazılım kodlaması yapılmıştır. Geliştirilen DL-GAN yöntemi, Jilin University– el üstü veri tabanı ve 11K el üstü veri tabanı olmak üzere iki ayrı veri tabanı üzerinde test edilmiştir. El üstü damar veri seti üzerinde yapılan deneylerin sonuçları ise, DL-GAN yönteminin %98,36 doğruluğa ulaştığını ve %2,47 hata oranına ve %0,19 standart doğruluğa sahiptir. İkinci veri setindeki deney sonuçları ise, doğruluğu %96,43 eşit hata oranı %3,55 ve standart doğruluk %0,21'dir. Geliştirilmiş DL-GAN yöntemi, aynı veri tabanları ile karşılaştırıldığında LBP, LPQ, GABOR, FGM, BGM ve SIFT gibi biyofiziksel yöntemlerden daha iyi sonuçlar elde edilmiştir.

Anahtar Kelimeler: Kimlik doğrulama, biyometri, tespit sistemi, el üstü damarları, derin öğrenme, üretken çekişmeli ağ.

CONTENTS

Ph.D. THESIS EXAMINATION RESULT FORM.....	ii
ETHICAL DECLARATION	iii
ACKNOWLEDGMENT	iv
ABSTRACT	v
ÖZ.....	vi
NOMENCLATURE.....	ix
LIST OF TABLES	x
LIST OF FIGURES	xi
CHAPTER 1	1
1. INTRODUCTION.....	1
1.1 Hand Biometric.....	4
1.2 Research Problem Statement	8
1.3 The Research Objective	9
1.4 The Research Contributions	10
1.5 Thesis Structure	12
CHAPTER 2	13
LITERATURE REVIEW	13
2. Background of Biometrics.....	13
2.1 Physiological Characteristics	15
2.1.1 Fingerprints	16
2.1.2 Palm Print	18
2.1.3 Face Recognition System	19
2.1.4 Hand Geometry	21
2.1.5 Iris Recognition	22
2.1.6 Dorsal Hand Veins	25
2.2 Behavioural Biometric Characteristics	26
2.2.1 Keystroke Dynamics	27
2.2.2 Mouse Dynamics.....	28
2.2.3 Voice Recognition.....	28
2.2.4 Signature Recognition	30
2.2.5 The Gait.....	32
2.3 Vein Pattern Recognition Applications	32
2.4 Deep Learning	38
2.4.1 Classification of DL Approaches	40
2.4.1.1 Deep Supervised Learning	41

2.4.1.2 Deep Semi-Supervised Learning.....	41
2.4.1.3 Deep Unsupervised Learning.....	42
2.4.1.4 Deep Reinforcement Learning.....	42
2.5 Generative Adversarial Network (GAN).....	44
2.5.1 GAN Network Structure.....	46
2.5.2 Conditional Generative Adversarial Network (CGAN).....	48
2.6 Related Works	50
2.6 DATASET	56
CHAPTER 3	58
METHODOLOGY	58
3.1 Image Acquisition.....	58
3.2 Preprocessing Images	60
3.3 Threshold Method.....	64
3.4 Generative Adversarial Network (GAN).....	66
3.5 Discriminator	72
3.6 Generator Classifier	75
3.7 Evaluation Criteria.....	76
CHAPTER 4.....	78
Experimental Results and Discussion	78
4.1 Experiments Results	78
4.2 Experiment Analysis.....	79
4.2.1 Jilin University-Dorsal Hand Vein Database Results	79
4.2.2 The 11K Hands Dataset Results.....	83
4.3 Discussion.....	87
CHAPTER 5.....	89
Conclusion and Future Work	89
5.1 Conclusion	89
5.2. Future Work.....	90
REFERENCES.....	92
APPENDICES	108
Appendix A – Artificial Neural Networks with Math.	109
Appendix B –Deep Learning	116
Appendix C – Generative Adversarial Networks (GAN).....	126

NOMENCLATURE

Acronyms

ACC	Accuracy
ANN	Artificial Neural Network
ACO	Ant Colony Optimization
BGM	Biometric Graph Matching
CNN	Convolution Neural Network
cGAN	Conditional Generative Adversarial Network
DBN	Deep Belief Network
DCGAN	Deep Convolutional Generative Adversarial Network
DG-GAN	Discriminative and Generative-Generative Adversarial Network
DL	Deep Learning
ECG	Electrocardiogram
EEG	Electroencephalography
EER	Equal Error Rate
FGM	Flamelt-Generated Manifolds
GAN	Generative Adversarial Network
HMM	Hidden Markov Model
LBP	Local Binary Pattern
LLBP	Local Line Binary Pattern
LPQ	Local Phase Quantization
NBM	Naive Bayes Model
RBM	Restricted Boltzmann Machine
RCN	Recurrent Comparative Network
ReLU	Rectified Linear Unit
ResNet	Residual Convolutional Neural Network
SIFT	Scale-Invariant Feature Transform
STD	Standard Deviation of The Accuracy

LIST OF TABLES

Table 1. 2 Table compares some of the biometric systems used lately.	14
Table 4. 2 Jilin University-Dorsal Hand Vein Database description	79
Table 4.3. Comparison of the ACC, STD, and EER.	81
Table 4.4. Dataset 2 description.	83
Table 4.5. Comparison of the ACC, the STD, and the EER.	84
Table 4.6. Experimental results for the DLGAN based on the databases that were used.	88
Table B. 1 Differences between the ML and the DL.	117
Table C. 1 Discriminator and generator losses equations.	128

LIST OF FIGURES

Figure 1. 1 Hand veins image.....	4
Figure 1. 2 Anatomy of the dorsal hand veins.	6
Figure 1. 3 Anatomy of skin tissue	7
Figure 2. 1 Types of biometric systems.	15
Figure 2. 2 Sample Fingerprint.	17
Figure 2. 3 Image of a right hand, which was scanned with a desktop scanner, that shows the main wrinkles and lines.....	19
Figure 2. 4 Face recognition.....	21
Figure 2. 5 Iris recognition.....	24
Figure 2. 6 Hand Vein pattern.	25
Figure 2. 7 Keystroke dynamics recognition.	28
Figure 2. 8 Voice Biometrics	29
Figure 2. 9 Biometric Authentication System.....	44
Figure 2. 10 GAN network architecture.....	48
Figure 2. 11 Conditional GAN.....	49
Figure 2. 12 Images from the JU-DHVD.....	56
Figure 2. 13 The 11K dataset.	57
Figure 3. 2 The DLGAN network that is used for identification of an individual based on the dorsal veins in their hand.	59
Figure 3. 3 Images showing the DHVs on the back of the hand using an IR camera.....	60
Figure 3. 4 Pre-processing diagram.	61
Figure 3. 5 . Shown above are images of a Gray scaled hand image and the veins within it (far left), the edge of the hand and the area of interest showing a Sobel edge (middle left), the segmentation area of interest extraction or the ROC by filling in the cavities (middle right), and the vein extraction of the hand (far right).	63
Figure 3. 6 Dorsal image preprocessing.....	64
Figure 3. 7 Learning in the generative adversarial network.....	68
Figure 3. 8 The architecture of the GAN. A false image is first generated by the generator from a random noise vector. Then, the discriminator faces the task of discriminating between the false images and those of the real training data.	70
Figure 3. 9 The general architecture of an ACGAN.....	71
Figure 3. 10 GAN Discriminator Model.....	73
Figure 3. 11 Real and Fake DHV	74
Figure 3. 12 Discrimination layers.....	75
Figure 4. 2 Accuracy and Loss of the JU-DHVD.	80
Figure 4. 3 Accuracy comparison of the method that was proposed herein against the other learning methods. The change rates comprised 6.39 max% and 1.67 min%.....	81

Figure 4. 4 Comparison of the SD of the method that was proposed herein with the other learning methods that were tested. The change rates were 0.288 max% and 0.018 min%.....	82
Figure 4. 5 Comparison of the EER of the method that was proposed herein with the other learning methods that were tested. The change rates were 7.42 max% and 1.45 min%.....	83
Figure 4.6. 11 K database accuracy and loss.....	84
Figure 4.7. Comparison of the ACC of the method that was proposed herein with the other learning methods that were tested. The change rates were 3.37 max% and 1.20 min%.....	85
Figure 4. 8 Comparison of the SD of the method that was proposed herein with the other learning methods that were tested. The change rates were 0.23 max% and 0.002 min%.....	86
Figure 4. 9 Comparison of the EER of the method that was proposed herein with the other learning methods that were tested. The change rates were 5.68 max% and 0.33 min%.....	86
Figure A. 1 NN architecture.....	109
Figure A. 2 Activation function.....	110
Figure A. 3 The 2 layers of the NN on the left comprise 1 HL that has 4 neurons, as well as 1 output layer that has 2 neurons, and 3 outputs. The 3 layers of the NN on the right comprises 3 inputs, 2 HLS that have 4 neurons as well as 1 output layer.....	110
Figure A. 4 Forward propagation.....	111
Figure A. 5 Back propagation.....	114
Figure B. 1 DL family.....	116
Figure B. 2 Differences between the DL and the ML.....	117
Figure B. 3 Types of ML.....	118
Figure B. 4 Convolutional NN architecture.....	119
Figure B. 5 CNN models.....	122
Figure C. 1 GAN architecture.....	126
Figure C. 2 Some of variants of the GAN.....	136

CHAPTER 1

1. INTRODUCTION

In the world today, the process of individual authentication is utilized widely for the purpose of security. This technique of identity verification is seen as an innovation that has resulted due to advances in modern technology. This innovation makes use of a number of different methods and techniques that are used in the verification of identities to increase and improve security measures in many areas. Among some of these varying techniques, the technique of biometric authentication has become the most popular and commonly used around the globe. It consists of an authentication process that makes use of the personal biometric properties of an individual. The most used physical biometric measurements of a person that are used for identification comprise fingerprints [1], irises [2], palmprints [3], faces [4], speech [5], retinas [6], and palmprint-based systems [7]. Nonphysical and/or behavioral biometric properties can also be utilized, which include things like DNA, psychological conditions, handwriting, and gait.

Biometric methods can be based on a person's face, but the accuracy of these face recognition methods depends on the state, gesture, and brightness of the image. In biometric methods based on face image processing, many factors play a role. We can mention the feeling inside the image, the person's age, the quality of ambient light, and the imaging angle [8].

Of course, fingerprints are another important biometric method; however, there are some challenges with a method such as this. As an example, an elderly person's fingerprints or those of individuals such as laborers are quite hard to recognize. The physical activity and hard work that is required in some jobs result in those individuals losing their fingerprints; hence identifying these people using this method would prove challenging [9]. There are some situations in which an individual might temporarily destroy their fingerprints on purpose or destroy them permanently via burns or acid so that they cannot be identified in the context of criminal activities [10].

Hence, a good method that can be used to identify people is through the use of information in the iris of their eye. The iris authentication method is highly accurate; unfortunately, however, it also presents with some challenges. One of these is that iris biometric detection necessitates the use of very specific cameras and hardware, in addition to the substrates that are required for this method, which all make implementing its use very expensive [11]. Palm line information is among the most extensively used methods for the identification of individuals. The literature comprises studies showing that both the groove pattern on the hand and arteries of the hand are unique and can be used in the identification of an individual with high accuracy. In this method, some tissue features can be made use of for identification, including bumps, wrinkles, and the folds inside of the palm of the hand [12, 13]. Different palm printing features, including things like the lines, wrinkles, and geometric features, have very unique and different patterns in each individual. Therefore, such methods can be made use of in both biometric and in individual identification. Palmprints possess a number of different features with regard to the location of the veins in the hand and fingerprints, and each set of these features can be utilized in the identification of an individual. In the analysis of hand images, the vascular location can be utilized, and fingerprints or the grooves in the hand can be made use of during printing [14, 15].

Fingerprint recognition, which is also known as hand printing, has been incorporated into many of the applications used for security. When compared to the other physiological features of an individual, finger and palmprints are significant factors that can be used in the identification of an individual.

Most of the studies in the literature have placed their focus on methods that are texture-, line-, subspace learning-, correlation filter-, local descriptor-, and orientation coding-based [16]. Due to scratches in the palm, they might be unusable for identification in individuals such as laborers or miners. Information that is in the groove of the palm can be exposed to constant changes or have damage to the tissue; however, the artery pattern in the hand remains constant and unchanging. The veins that are seen in the hand are made up of two parts, those on the back of the hand and those that are within the palm. Using the artery patterns and veins on the

back of the hand has proven to be a method of identification that is dependable cannot be imitated [17]. The veins within the palm are much more difficult to use for identification than those that are on the back, because the tissues in the palm are denser than those on the back. An appropriate method that can be used in the identification of the veins in the hand is infrared (IR) cameras. With IR cameras, it is possible to just aim at the hand and then capture and read the pattern in the veins. The main advantage of this method of identification is that it is not necessary for the individual to place their hand on any specific surface, as is required with manual printing methods.

The use of biometrics to distinguish evidence remains the best method to use since it makes a match with the person, and not with data that starts in one location and then moves on to another [18]. The human hand is a multi-fingered and prehensile appendage that is located at the end of the forearm of individuals. The dorsal venous framework of the hand comprises a network of veins that are located in the shallow band, framed by the dorsal metacarpal veins, on the back part of the hand. This framework extends to major veins, such as the cephalic and basilic veins. Biometrics is a field that is still developing, in which the fundamental freedoms bunches communicated worry over protection and character issues [19]. The biometric laws and guidelines in procedure and biometric industry measures are categorized into two principles, which comprise finishing and bifurcation focuses, which are used to remove the highlights of the veins. The vein highlights, such as the length, shape, thickness, and circulation, have been explored in order to locate the best possible representation of their design [19, 20]. The physical condition of the vascular tree, which is located subcutaneously on the back of the hand is composed of data that is appropriate for use in confirming the personality of a person to a sensible outcome [21], which is the basis of this research work. The condition of the veins that are in the finger is an example and its utilization for distinguishing proof [22]. The hand vein biometrics rule entails a non-invasive and automated examination of the structures of the veins, at a subcutaneous level, in the back of a hand to check the personality of people for biometric applications.

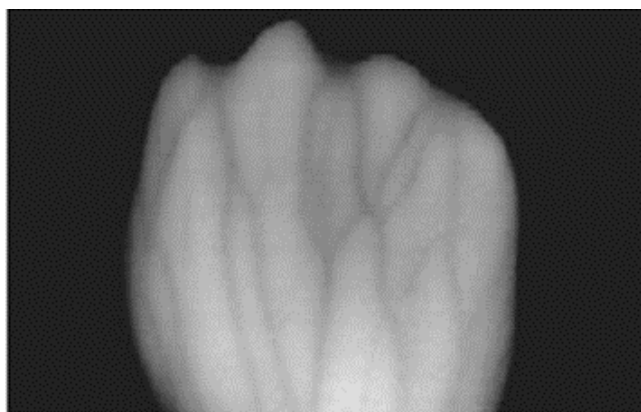


Figure1.1 Hand veins image

Biometrics is a computerized method that is used to identify a person by utilizing quantifiable organics, as indicated by anatomical and physiological and conduct attributes. The need the biometric framework since, presently we are living in a worldwide society of progressively frantic and unsafe individuals who can't be believed dependent on recognizable proof records, personality criminals take PIN, for example, date of birth to pull back cash from records [23]. Dorsal hand vein (DHV) acknowledgment framework is the distinguishing proof of the individual by perceiving the vein design displayed in the front and back of the hand [24]. An expanding request to receive the biometric geek to limit potential security hazards, a purchaser level biometric framework for programmed physical access control [25]. Through advancement of a close IR imaging gadget for biometric procurement of DHV examples and calculation effective picture handling strategies, they additionally announced that the proposed framework is appeared to have a scope of alluring activity attributes for certifiable use, which incorporate solid in recognizable proof, advantageous to utilize, high in security, quick accordingly, low in expense, and basic for establishment [25, 26].

1.1 Hand Biometric

Many features are used for security in hand biometric authentication. Its use has been widely preferred in recent years due to its effective authentication feature and its accuracy [27]. One of the biometric features of the human hand is the palm lines and

veins. It is used in authentication for security purposes and biometric identities because it is difficult to replicate personalized avatar traces and vessels [28, 29].

Apart from hand properties, there are other important biometric properties. One of them is DHVs [30, 31] and palm veins [32], and hand vein biometry and authentication are also used in many areas of daily life. The main areas of use are airports, banking systems, and everywhere where security is high. This new biometric technique is safer than other conventional biometric techniques. New biometric techniques brought by technology are also becoming widespread with new techniques daily. The main of these is hyperspectral technology [33, 34]. By combining the hyperspectral mode and technology with the biometric features of the other hand, higher security identification systems can be obtained [35].

The hyperspectral method or technique uses spectral information on a biometric hand image. This additional spectral information includes the dermal skin layers, epidermal layers, absorption coefficient, and hemoglobin. Thus, the reliability of identification is increased [35]. It is challenging to change and imitate the information obtained depending on the skin. Specific knowledge linked to the skin of an individual can be obtained using infrared and thermal cameras [36]. The spectral information obtained is used as a supplement to the hyperspectral information. The advantage of this is that it increases the accuracy rate in obtaining the biometric properties of an individual and correctly determining their identity.

As shown in Figure 1.2, the vein pattern is a vast network of blood vessels beneath the skin. Anatomically, aside from surgical intervention, vascular patterns in the body are distinct from each other, and are stable over long periods of time [36]. In addition, as blood vessels are hidden beneath the skin and are practically invisible to the eye, vein patterns are extremely difficult for intruders to copy when compared to other biometric features. The uniqueness, stability, and high immunity to vein pattern forgeries, makes these veins excellent biometric features, offering secure and reliable qualities for person identity verification [37].

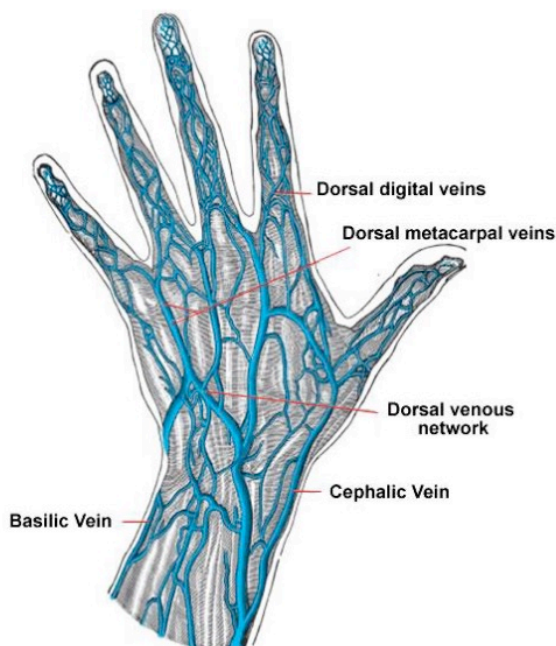


Figure 1.2 Anatomy of the dorsal hand veins

However, 3 key processes may alter vascular patterns in the hand: 1) natural changes in the vascular system throughout the life span, 2) natural changes in the vascular system, associated with disease, and 3) changes in the vascular system induced by other factors [38].

First, throughout life, the vein length extends while the body is growing, since the function of the vascular network is to provide oxygen to the entire body. Therefore, the vascular system must adapt to the size of the body. It will extend and shrink throughout life, with major changes before the age of 20 and minor changes during the aging process, from 20 years of age and onwards. Similarly, there is an inevitable decline in bone and muscle strength in the body, due to a time-dependent weakening of the blood supply to the muscles [39, 40].

Second, due to the dynamic character of the vascular system, it is sensitive to body conditions that deviate from normal or healthy forms. Diabetes, hypertension, atherosclerosis, metabolic diseases, or tumors can significantly remodel vascular systems. They induce effects on the mechanical properties of vessel walls, generating hemodynamic changes. The remodeling process results in thickening or thinning of

vessel walls in the lumen and in the external diameter. Another process influencing the vascular system during disease is angiogenesis, a hallmark of cancer and various ischaemic and inflammatory diseases [41].

Third, other factors that influence the vascular system are environmental temperatures, physical activities, and alcohol. This last factor should be considered not as the unhealthy use of alcohol causing permanent pathological changes in the vessels, but rather the temporary influence of alcohol while in the body. In this way, alcohol is a vessel dilator. Vessels also dilate if body temperatures are too high, as blood flow can increase by up to 150 times to expel excess heat. In cold weather, skin constricts the blood vessels and causes heat loss. During physical activity, blood vessels will also dilate to provide enough oxygen to the muscles [42].

As shown in Figure 1.3, the skin is made up of 3 layers: the epidermis, the dermis, and the hypodermis (also known as subcutaneous tissue). The top layer of skin is called the epidermis. It protects the underlying skin layers from the outside environment and contains cells that make keratin, a protein that waterproofs and strengthens the skin. The epidermis contains cells that express melanin, the dark pigment that gives skin its color. Other cells in the epidermis allow the sensation of touch and provide the body with immunity against foreign invaders like germs and bacteria. The bottom layer of the skin is the hypodermis [43]. It contains fat cells, or adipose tissue, which insulates the body to conserve heat. The layer between the epidermis and the hypodermis is the dermis, which contains cells that give the skin strength, support, and flexibility [44].

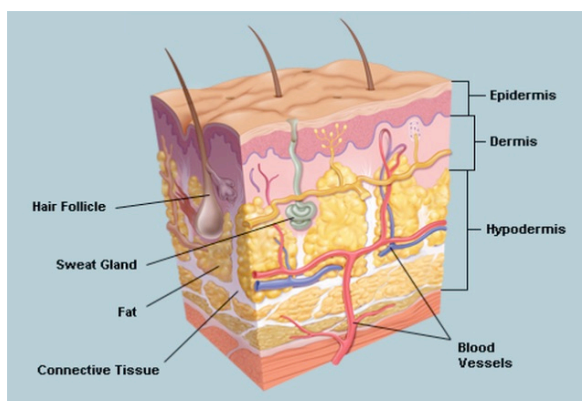


Figure 1.3 Anatomy of skin tissue

As a person ages, the cells in the dermis lose strength and flexibility, causing the skin to lose its youthful appearance. Aging changes the skin dynamics, and the skin becomes more transparent, which is caused by epidermal thinning, wherein the skin becomes slack due to the loss of elastic tissue and the loss of fat below the skin may result in a loosening ‘skeletal’ appearance.

1.2 Research Problem Statement

Organizations around the world today are witnessing more developments in internet technologies and the digital world. Companies and institutions seek a high level of security frameworks that are keen to start benefiting from the best innovations in the world of documented information protection and security using the best systems.

Effective organizations today are racing to use the latest innovative frameworks to have an effective security framework to protect their data and ensure an effective framework for dealing with users, for example, Personnel Recognition and Identification (DHVR).

The aim of using these approaches is to help organizations and organizations control a security framework effectively without any kind of misfortune. Therefore, any real-time security problem facing identification in the company or organization will avoid any kind of loss. Moreover, data security is a necessary issue also through the means by which the employee or users can be able to manage the data of the organization and the frameworks in which to talk about using the DHVR approach and identification to ensure privacy and security.

As a result of the previously performed analysis, it is possible to put forward the following problem statement:

- How can an individual be identified using the pattern structure in his vein?
- Starting from this stage, we can ask the following questions:
- How should the manual vein pattern recognition system be structured?

- What methods are available for use in identifying vein patterns, and how can we combine them?
- How to create an inexpensive setup for capturing vein patterns that will still provide consistent, high quality images?
- How can we test the performance of a system like this?

1.3 The Research Objective

Through an increasing emphasis on security, automated personal identification utilizing hand vein biometric features remains suitable and a hot topic in applications in both research and those that are practical in nature.

- The main goal of this research is to present a high-quality biometric authentication framework for high-security physical access control (PAC) based on the DHV pattern.
- To understand the steps of DHVR as well as identification for security behaviors of several users at any organization.
- To identify the degree of awareness of several users of their own role in the protection of the organization's information and that of their personal information, as well as how to behave to fulfill this role at any organization or firm (public/private).
- To investigate and enhance the size of loss that can be utilized in the case of information security weaknesses at any organization.
- To recognize the compliance of the users with the organization or the company information policies and controls.
- To identify the DHVR approach, which has the potential towards resolving any security problems that occur via external or internal users at the company or organization.

- To observe the advantages as well as the drawbacks of DHVR and identification for future research study.
- To gain a comprehensive understanding on the DHVR framework.
- To give an overview of the current state of the research based on the DHVR framework.
- To present the promising future research directions based on adaptive techniques on the DHVR framework.

1.4 The Research Contributions

This research study will start by segmenting the detection of the dorsal vein identification within the four required stages of the system processes, starting with acquiring an image a person as well as an image of a similar person. Four steps of the process will be performed: for example, checking the input image comparison and comparing the image resolution. In addition, an additional compression and matching process will be included to avoid the occurrence of re-inserted images. This study will also use a database to store images relevant to an organization, company employee, or user interested in using the system. Moreover, this research study will carry out dorsal vein identification via the use of deep learning (DL) with a multi-step approach that includes feature selection as well as image preprocessing, in addition to the use of the generative adversarial network (GAN) method to provide the most effective identification of the individuals by using the information that is taken from their dorsal vein images. Identity verification research that is based on biometrics was conducted, and the depth-preserving latent GAN (DLGAN) method was implemented and tested on the same databases, but in different ways. Validation of the obtained results using the DLGAN was performed using different methods, but under the same conditions, and the results were compared. While an approach that could create many educational and artificial examples is the application of DLGAN. One areas in which GAN learning has been implemented is in the field of image processing.

- The biggest challenge of biometric methods is that they are based on the information in the veins of the hand, and they examine most of the information on only one side of the hand.
- In most cases, the diagnostic error in this method can be significant.
- In recent years, DL methods with the GAN approach have become a convenient tool for pattern recognition and have high accuracy.
- A multi-step approach including feature selection and image preprocessing is used along with the GAN method to effectively identify individuals based on information from the veins on the backs of their hands.
- A variety of methods for biometric-based authentication has been introduced thus far.
- Although these methods are efficient, they also have challenges; for example, fingerprints present challenges such as fingerprint destruction due to activity and labor and are not always available.
- Using information from the veins of the hand, especially the veins on the back of the hand, which are more easily recognizable, is a new method in the field of biometrics that is used to identify individuals.
- Labeled data is the primary source of information for supervised classification, but unfortunately, the small volume of training samples or their excessive similarity to each other can lead to a poorly trained classification method.
- One approach that can create a large number of educational and artificial examples is to apply GAN deep learning.
- One field of GAN learning is the use of this method in image processing.
- The GAN is based on game theory and is an efficient tool for analyzing images, especially in the field of biometrics.
- This thesis uses a GANs to identify individuals based on the veins in their hands.
- The GAN network, like any deep-learning method, has a time overhead for generating instructional examples and training, so future research will attempt to use GPU architecture to expedite authentication operations.

1.5 Thesis Structure

This dissertation is structured into five chapters:

Chapter one: An overview of the main motivations, and the research problem statement, as well as the research objectives and contributions are summarized.

Chapter two: This chapter presents the literature review of DHVs, and the background of biometrics and datasets.

Chapter three: This chapter explains the proposed method in detail, describing the steps taken to implement the algorithm including the conducted experimental study. Pseudo code and other supported graphs and figures are also presented.

Chapter four: The experimental analysis and discussion are described in this chapter, which contains a complete description of the synthetic and real-world databases that have been used to perform the experiments. The evaluation of the obtained results and analysis of the experiments using several scenarios are also presented.

Chapter five: In the final chapter, the presented research is summarized and concluded with the findings and contributions of the dissertation. A review of all of the research objectives and results are reviewed. This chapter also emphasizes the contributions to science that this research work will provide. At the end of this chapter, some prospective aims for future work with regard to this research are presented.

CHAPTER 2

LITERATURE REVIEW

This chapter presents recent related research, including theoretical and methodological contributions, as well as substantive findings. Various related works are discussed and examined in this section. Indeed, the research methodology herein is at the crossroads of various research fields. To begin, the concept of clustering is introduced, as is referred to in this dissertation. Then, some terminologies in connection with the algorithm proposed are defined and various research studies relevant to the current discussion are examined. By analyzing various methods, the study explores the advantages and disadvantages of a particular method in order to give a more comprehensive overview of the relevant literature.

2. Background of Biometrics

The word biometric is made up of two Greek words: ‘bio’, which means life, and ‘metrics’, which means to measure. Biometrics comprises authentication that is based on something that an individual is, instead of something that they know. Passwords and smart cards are nowadays the default methods used for the authentication that is necessary to grant access to protected information [47]. Passwords are used commonly because they are a simple and inexpensive mechanism to implement and use. However, mostly due to weak character combinations and poor password practices, they are known for being a poor method of protection. They are also easy to steal and forget. Biometric recognition systems surpass these problems because they only depend on biological and behavioral characteristics, which are inherent to the human individuals.

For a number of decades now, researchers have been making improvements in biometrical methods and attempting to ascertain different body parameters or behavioral methods that are unique to an individual that can be implemented into a sample group and used to differentiate and identify a single individual, while obtaining the high precision and in a very short time.

Biometrics basically searched for data that remains constant over the course of the life of an individual or those that are not easy to simulate or intentionally change [48].

Any characteristic of the human body can be utilized in biometrics so long as it is able to fulfill the requirements below [49]:

- Universality - Every user should have the required trait.
- Uniqueness - The trait used should be sufficiently different for each user.
- Permanence - The trait used must be reasonably invariant over time.
- Measurability - The trait must have an easy acquisition and measurement and in addition, the obtained data must be easy to process.
- Performance - The system must be accurate, fast, and robust.

In order for a biometric system to be well-implemented, some factors that must be taken into account, such as the time response, precision, and cost. These will all be dependent on the software and hardware that is used in its implementation. Moreover, biometric identification has been becoming more and more commonly used as both software and hardware have seen a reduction in price [50].

Table 2.1 presents some of the spread methods that are used for biometric systems, as well as the uniqueness, stability, universality, acceptance, security, timeliness, ease of acquisition, accuracy, cost of implementation, devices that are required, as well as how well these methods have been accepted by society.

Table 1. 1 Table compares some of the biometric systems used lately.

Features	Fingerprint	Iris	Voice	Face	Signature	Hand Shape	Dorsal Hand Vein
Uniqueness	H	H	L	L	L	M	H
Stability	H	H	L	M	L	M	H
Universality	M	H	M	H	H	M	H
Acceptance	M	L	H	H	H	H	H
Security	H	H	L	L	L	M	H
Timeliness	H	H	L	L	L	M	H
Ease of Acquisition	M	M	M	H	H	H	M
Accuracy	H	H	M	M	L	M	H

A wide array of methods exist that can be used for biometric identification, each of which has advantages and disadvantages, which will be discussed throughout this chapter.

Figure 2.1 presents 2 types of biometric identification, physiological and behavioral. The differences in the principles for these 2 types will be explained, including the parameters that must be taken into account for each one [51]. Examples of both of these types of biometrics, which are used currently for the identification of individuals will be explained, beginning with a short overview in the next section.

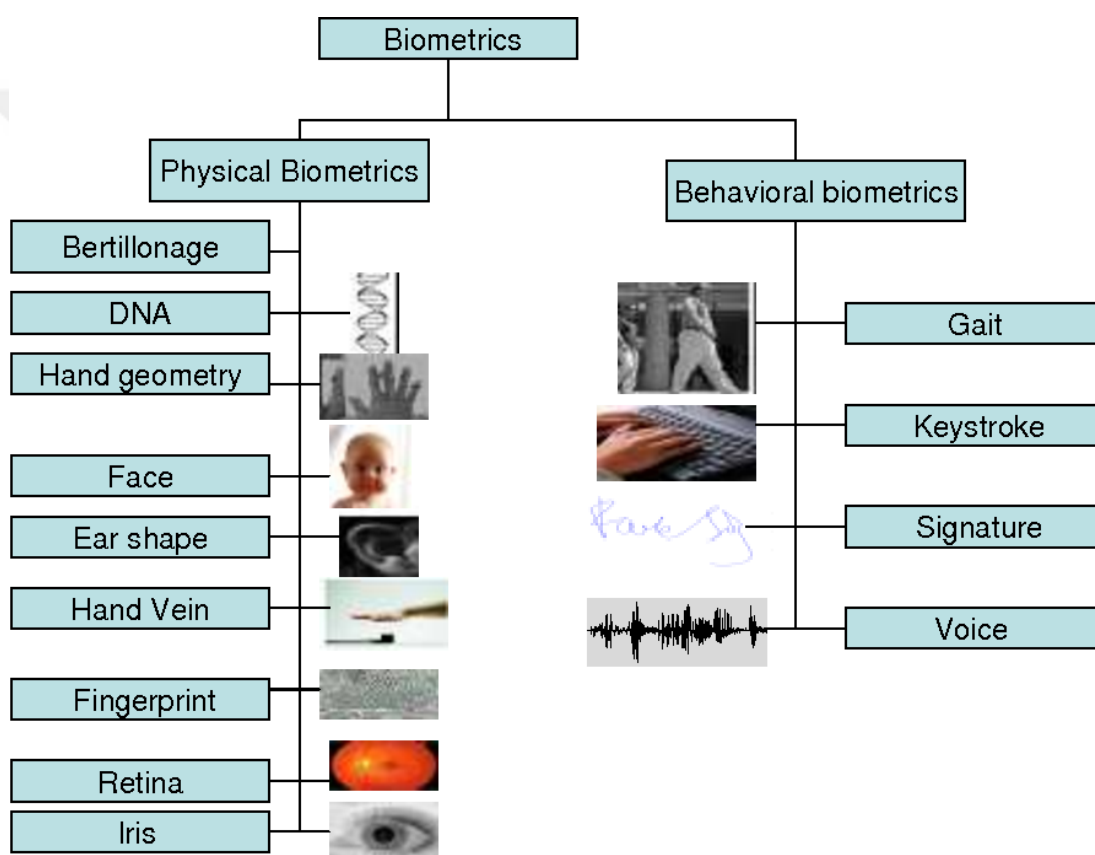


Figure 2.1 Types of biometric systems

2.1 Physiological Characteristics

As suggested by its name, the physiological type of biometrics is based on physiological features, such as the human face, fingerprints, and palmprints. The main advantage that is provided by method when compared to the behavioral type of biometrics is that the majority of these features do not change over time, which

means that, as time passes, they remain the same, and will not be affected by psychological state of an individual, such as their happiness or anger. However, these types of methods could be affected by things such as injuries or accidents, which can have a direct effect on the feature that is being utilized [52]. An example of this would be scratches on the hands or wounds to the fingers or palms, which can have an effect on the reading and recognition that is performed by the reading devices that are used for reading fingerprints and palmprints.

2.1.1 Fingerprints

One of the common methods used in physiological biometry is conducted using human fingerprints. Of these, fingerprint recognition is the most well-known method because of it being feasible, effective, reliable, fast, and very easy to apply in daily life [53].

A fingerprint biometric consists of a digital version of the method of ink and cardboard, which was used over a great deal of the 20th Century. Therefore, a fingerprint-based identifier is the oldest form of biometrics that is in use today. Due to this, it is the most extensively used modality of biometrics.

The scanning of fingerprints comprises the process of the acquiring the fingerprint characteristics of an individual and then quantifying the probability of being able to authenticate the identity of that the individual. Fingerprints comprise the physical dermal structures on the tips of the fingers that are defined prior to the birth of an individual. The furrow and ridge patterns on the surface of the skin of each finger, as seen in Figure 2.2, minutiae points, such as the local ridge characteristics that are located at the bifurcation or end of a ridge ending, and the image density are all unique to everyone. On the finger, the print pattern on the top joint, which is known as the distal interphalangeal joint, is the main location of focus. Historically, the process of fingerprinting has been used in the identification of criminals as well as in other forensic activities [54].



Figure 2.2 Sample Fingerprint

Due to the fact that fingerprints are the oldest form of biometric technology, there are many choices that can be made with regard to vendors and products. Fingerprint-based systems, which are also called finger imaging or finger scanning systems, involve the process of scanning one or more fingertips of an individual and then making a comparison of the scans of these fingertips with known images that are contained within a previously established fingerprint template that was created during enrollment into a program.

However, the process of scanning fingerprints has a variety of limitations that must also be considered. Authorities in the biometric industry have indicated that some 2% to 5% of the general population have physical limitations with regard to the use of fingerprint imaging technology. The ridges on the fingerprints of an individual can deteriorate both as they age and as the result of some types of occupational activities. It is a well-known fact that fingerprints possess a latency property. What that means is that individuals leave their fingerprints behind on almost every surface that they come into contact with, including things such as the drinking glass that they hold at cafes and restaurants, the doors that they open, and the phones and keyboards that they use at their jobs. This can have an impact on the privacy of the individual and encourage imposters to attempt to fool the system, which would decrease the security that would result from this biometric modality. Additionally, in a lot of countries, the process of fingerprinting has been culturally associated with the criminal element of society and forensics sciences; Hence, some people have a strong disinclination toward fingerprinting, meaning that it is not seen as being as socially acceptable as some of the other modalities, and this may significantly inhibit its global adoption.

Moreover, the fingerprint systems that are presently in use do not possess and form of liveness detection to ensure that the prints are in fact those of a real person; hence, a number of systems have been spoofed. If fingerprint-based systems included a liveness detection feature, this would enable them to have greater resistance to spoofing. However, the drawback of this would be that the volume of equipment required to construct them, and their price would tend to increase, which would neutralize their key advantages, that is, their size and price range. Finally, the negative aspect of the high volume of fingerprint-based products and vendors is that a number of fingerprint scanners have experienced a critical interoperability issues, despite the fact that there has been marked improvement as a result of the new standards that have been introduced.

2.1.2 Palm Print

The usefulness of palmprints has been gradually validated over the past 15 years. A handprint is defined as a small area on the surface of the palm that contains information that can be very useful an authentication system. There are 7 factors that influence biometrics recognition in each application, which comprise universality, exclusivity, permanence, achievement, performance, acceptability, and fraud. Handprint recognition was first introduced about a decade ago, and since then, it has gradually been attracting the attention of a wide range of researchers as a result of the richness of several of its several features. The palm of the hand comprises the inner surface that is between the fingers and the wrist. The area comprising the palm consists of many features that can be made use of as biometric features, which include things like the major lines, the geometry, wrinkles, the delta point, fine detailing, the datum point features, and the texture of the skin [55]. The basic lines are what is known as the fold creases, and their formation has to do with the movements made by the fingers, the structures of the tissue, and the skin's purpose. They are unique for each individual, where even in identical twins, the palmprints are not the same [56]. This is the result of the genetic code that is in DNA, which contains general instructions with regard to the way that the skin should be formed in a fetus as it is developing. However, with regard to the specific way in which it is formed, this results from the random events that occur to the location of the fetus in

the womb, which includes things such as a specific moment and/or the formation and the density of the amniotic fluid that surrounds the fetus. Palmprint verification involves the utilization of high- or low-resolution images. The majority of the research that has been conducted on handprint verification has made use of low-resolution images [57]. A palmprint verification system works based on 4 basic stages, which comprise the actual acquisition of an image of the palmprint, pre-processing, extraction of the features, and finally, matching, as can be seen in Figure 2.3. The handprint acquisition of an image of the palmprint is performed using a handprint scanner. Pre-processing consists of 2 parts, which are alignment of the image, and selection of the area. The image is aligned with the indication of key points. Determining the area of interest comprises cropping the image of the handprint from the image of the hand [58]. Extracting the features involves the selection of the distinctive features from previously processed palmprints. Matching compares the features of the captured image with the stored templates. The handprint image has a unique feature that will not change for a long period of time.

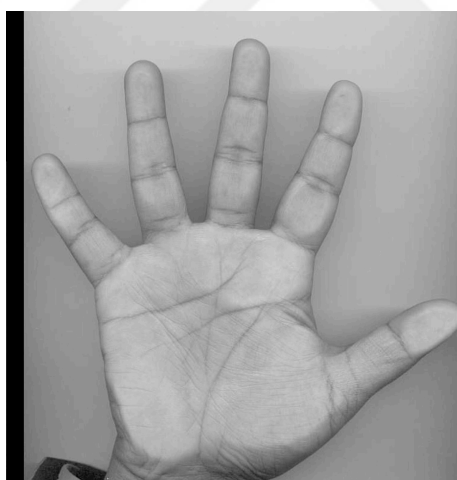


Figure 2.3 Image of a right hand, which was scanned with a desktop scanner, that shows the main wrinkles and lines

2.1.3 Face Recognition System

A face recognition system comprises a form of biometric technology that is used in the identification of verification of a user via a comparison of their unique facial features to those stored in a database. The use of face recognition has been rapidly increasing, as it is both an interesting, as well as challenging field of biometric

systems. Generally, it is used in areas of security that are able to be compared with other biometrics, such as fingerprints, irises, etc. Since user privacy and data privacy are known to be very important, and to be able to protect that, we must have security. Many types of security systems and techniques exist for securing data; however, each of these systems has some drawbacks and that can be overcome through face recognition. In a face recognition system, there are no passwords that must be remembered; thus, nobody can steal and use your password. In a face recognition system, your face is the password, which means that nobody can gain access to the data unless they have your permission. The very first face recognition system ever established was in 1960, and since then, the system has become greatly improved. Several technologies have emerged to optimize the system so as to attain better accuracy with regard to results, such as GABOR filters, Eigenface, neural networks (NNs), support vector machine (SVM), 3D recognition, and so on. As a result of these new technologies and algorithms, it has become more difficult to figure out which of the methods is the best one to use for security purposes. Face recognition has become one of the very best types of security systems [59].

In facial recognition, the features of the face are extracted, such as the shape of the nose and its size, the eyes, the lips, and so on. Each one of these features are different and unique to each individual. Once taken, these features are then stored in a database, and whenever it is requested to verify the user, the newly scanned features are then matched against those in the database. However, there are some disadvantages, as in the case of twins, where the good user cannot be distinguished. Second, if someone uses the user's photo, the system is not able to distinguish if what is being used is an actual person or an image.

In recent years, a lot of focus and attention has been directed toward all areas of biometrics, because this is the best and safest way to be able to get to know the user. This does not necessitate remembering a password and or the use of a plastic card; it is easy to use, and biologically, the appearance of humans never changes. Many different techniques exist in biometrics, including fingerprint recognition, iris recognition, palmistry, etc.; however, face recognition has been gaining more attention with regard to computer vision [60]. The first face recognition research was

conducted in the early 1950s and 1960s. However, face recognition remains difficult due to the complexity of the human face and the difficulty in extracting its features. In the last few decades, in the field of surveillance and security, face recognition has become the application that is most commonly used. Moreover, face recognition is a very challenging aspect of image analysis due to the fact that every individual has facial features that are different, as is seen in Figure 2.4. Face recognition is geared towards the visible spectrum due to several reasons[61].

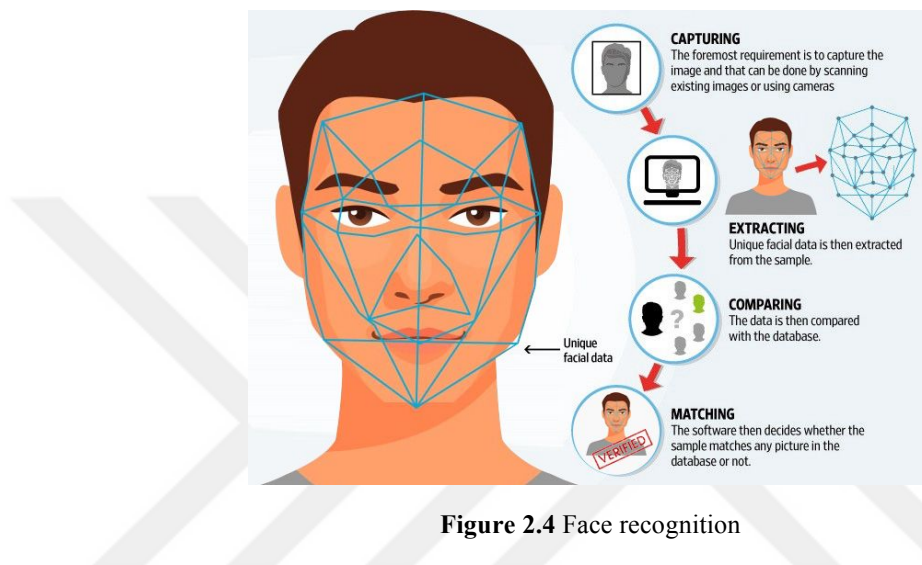


Figure 2.4 Face recognition

2.1.4 Hand Geometry

Biometrics systems that are based on hand geometry are among those that are the most used. The concept of hand geometry puts the focus on the physical structures of a hand in an outstretched position, such as things like the length, width, angle, and thickness of the palms and the fingers. Most often, the user will put their hand onto a reader that has studded guide pegs. There are advantages to manual engineering, which include the fact that they are naturally very easy to use, their templates normally use between 9 and 20 bytes of data, which is actually the smallest form in any of the biometric methods, have very low false-rejection and failure-to-register error rates, a lesser degree of privacy concerns when compared to fingerprint or face recognition, and these kinds of readers do very well even under harsh environmental conditions; for example, in commercial factories in which the users can be very dirty. This is a method that can easily be implemented at industrial sites, like warehouses or manufacturing facilities, such as for time, PAC, and attendance applications, where it is simple to use. Furthermore, one can easily combine hand geometry with

other biometrics, like fingerprint recognition, as well as vein pattern recognition in the fingers or the hand. However, unfortunately, the hand is not a unique feature, which means that engineering scanners have a false acceptance rate that has been reported to be about 0.1%. In addition, the features of the hand change more over time when compared to most other biometric modalities) [62].

The main disadvantage of using hand geometry lies in the fact that it is possible for the measurements to also match those of other individuals, which means that it is not a unique type of characteristic, and that is why it must work in conjunction with some other type of identification, like with a personal identification card or number. It is commonly used for task verification rather than identification. Additionally, hand engineering consists of some fundamental flaws, such as the fact that the readers are relatively large, the accuracy rates are moderate, and there are concerns with regards to sanitation because an individual must put their hands on the same surface as all of the other people who were scanned before them. Moreover, because such a small volume of information will be measured, hand geometry is unsuitable for most of the applications that are used for identification, especially those that will be used on a large population [54].

2.1.5 Iris Recognition

Eye recognition systems are a type of biotechnology that is one of the most accurate available today, and the retina information of almost anyone who is in good health can be recorded. There are no symmetrical iris or retina patterns, neither between the right and left eye, nor between twins that are identical. These 2 types of biometric methods are perceived to be similar by the public, even though they are widely different with regard to both their operation and the required level of user cooperation.

Retinal biometrics, which are also known as retinography, work by verifying the identity of an individual via a review of the unique blood vessel pattern in the retina. Retina scans pick up patterns of blood vessels in the optic disc, which is the center of the retina. The blood vessels in this area tend to remain the same throughout the life of an individual. A retina scanner works by shining a low-intensity IR light beam

into the eye. If the user wears glasses, they must be removed prior to scanning, without moving his/her head, with his/her eyes approximately about 30 cm from the scanner, and he/she has to focus their vision at a predetermined target with the IR beam directed at the pupil of his/her eye, in a similar manner to the process during an eye exam. The scanner then takes a measurement of the reflected light and is able to capture about 100 points of data.

Not only does it have a high resolution, but retina scanning results in an extremely compact template. Hence, modeling this method only requires a small amount of the memory storage that most of the other methods require. However, the retina can be prone to various diseases, which include retinopathy, glaucoma, age-related macular degeneration (AMD), and so on, which is more frequently observed in the elderly. Despite the fact that this type of biometrics does not result in any physical discomfort, a high number of users feel uncomfortable with a retina scan, as they are seen as intrusive and, to a certain extent, cumbersome. Primarily, retina scanning is used in applications where advanced safety is required, as no mechanism currently exists for retinal duplication and the retina of a body that is deceased is not able to be used for scanning.

The iris of an individual is distinct. The iris of the eye remains the same over time and is not likely to experience erosion or injury than the other measurable area of the body as shows in figure 2.5. The iris presents a structure that is rich in data and comprises fibrous as well as vascular tissues, which includes things such as freckles, pits, grooves, loops, fissures, and frills. However, the color of the iris color cannot be used. By measuring these properties as well as the inter-spatial relationships between them it is possible to obtain data that is useful for verifying or identifying an individual [54].

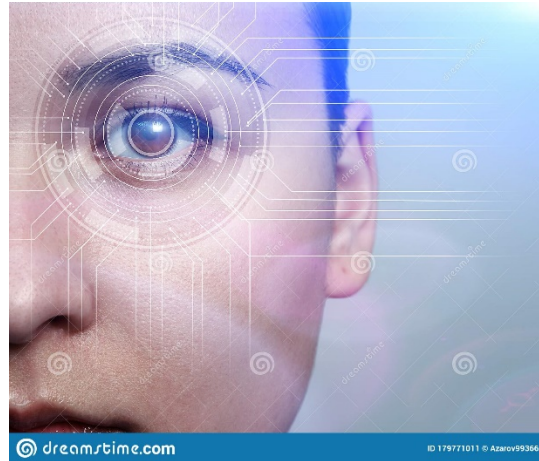


Figure 2.5 Iris recognition

Despite the fact that the majority of biometrics make use of between 13 and 60 distinct properties, the iris, due to its complexity, has been reported to consist of 266 degrees of freedom with regard to compositional contrast, which makes it the biometric that has most accuracy when compared to the other methods that are in use today. As a matter of fact, the iris of an individual contains almost 6 times more unique and measurable characteristics as what can be found in a fingerprint. However, to date iris-based biometric systems have not even begun to come close to attaining such a level of accuracy. At best, iris scanners are able to measure the properties within the iris from a distance of a meter or more away [63].

The process of collecting an image of an iris involves a greater degree of training and attention than is necessary in most of the other biometric methods. Even though this can be a high-resolution technique, it is not uncommon for there to be images that have low quality, which can result from eyelids that are droopy, pupils that are dilated, irises that are not centered, blurriness, and poor resolution. In some systems, difficulties such as these can result in failure rates that are higher than normal. Most of the iris recognition systems that are in use today are a little bit restricted as a result of the reading distance that is required, the fact that special lighting is needed, as well as the camera angle that is required, that is, the position of the eye in relation to the location of the camera. Moreover, iris scanning systems have the tendency to be some of the most expensive types of biometrics systems. However, they are able to work well in both identification and verification modes. Furthermore, iris scanning systems are not as intrusive when compared to retinal scanning. Another significant

advantage lies in the fact that it can be performed passively, which means that there is no active effort that must be put forth by the individual who is being scanned. The systems currently in existence are even able to be used if the individual has been fitted with glasses or contact lenses. That said, it is still possible for dirty glasses or lenses that have been extensively scratched lenses to interfere in the recognition process.

The iris scanner was originally designed to be used in PAC applications; However, it is currently being used in a wide array of settings, such as in facilities that require high security, commercial banking, and healthcare. Aa a matter of fact, the use of iris recognition has been increasing in many accessibility applications, both in the government and in commercial arenas.

2.1.6 Dorsal Hand Veins

This thesis focuses on the study of dorsal hand vein biometrics (DHV). The dorsal hand vein pattern is the vein pattern that has been preferred to represent a personal identity. To examine the images of the dorsal vein. Hand vein biometrics rely on using the pattern, structure, and location of veins in people's hands to determine who they are. The hand vein pattern has been found to be unique to people [21]. The network of blood vessels, including the veins, is responsible for transporting oxygen throughout the body through the blood. Some veins can be seen with the naked eye, but with near infrared light, the veins become more visible, near infrared light and optical window. An example picture of the dorsal side of the hand can be seen in figure 2.6. In this picture, some veins are visible to the naked eye, while others are hidden.



Figure 2.6 Hand Vein pattern

The fact that some of the veins lie under the skin, and thus hidden, makes impersonation more difficult when it is compared to other biometrics, such as face or fingerprints, which are easy to obtain from a specific person. The human vascular network develops during childhood up until the age of 20. After the age of 20, the vascular network does not usually change rapidly [38].

However, some diseases can cause angiogenesis, which is the formation of new blood vessels [64]. Consequently, this can lead to changes in the vein pattern of an individual, and thus require re-registration for it to be possible to recognize this person. For healthy people, slight changes can also occur due to temperature, alcohol intake, etc.; however, these are usually only minor changes, such as the vein size, and they do not affect the structure or pattern of the veins [38].

When compared to other physical properties, phlebitis pattern recognition is one of the most recent biometric techniques that has been researched and developed to solve many of the problems facing traditional biometric systems. This is because:

1. Palm vein image acquisition does not require direct contact with the vein pattern extraction sensor, and because non-contact forms are cleaner than all other forms of contact biometrics, user comfort is improved using venography technology [65].
2. The vein pattern does not change over time [66], and can represent the vitality of an individual, so cognitive performance can be improved using venography technique and a stable operation is expected [65].
3. Vein recognition technology is significantly less expensive than many other biometric techniques (such as the iris scanning technique) [65].

2.2 Behavioural Biometric Characteristics

Behavioral biometrics comprises the utilization of external patterns or the psychological reactions of an individual. This field consists of things such as signatures written by hand, voice patterns, and mouse and keystroke dynamics [67]. Features such as these, differently than physiological biometrics, may be affected

directly by the mood or emotions of an individual. As an example, if the individual is in a bad mood, they will likely have higher error rate of when performing actions such as typing, or they may type more slowly in general. For this reason, this method is general less accurate than that of physiological biometrics, because the mood of an individual can suddenly change as the result of a variety of external factors.

2.2.1 Keystroke Dynamics

Figure 2.7. illustrates, the method used in keystroke dynamics recognition, which first began to be used during the World War II. At that time, military intelligence made use of it for distinguishing the rhythms of morse code messages, to determine whether the code was that of an ally or enemy. Hence, in the beginning keystroke dynamics was considered to be a biometric solution that could be easily implemented with regards to hardware.

Keystroke dynamics is an inexpensive biometric technology that is based on the concept that each individual has their own timing pattern when they are typing, and this can be used to be able to identify which individual is typing. Once these dates have been recorded, they are then processed using some different techniques, which can include NNs or statistical classification. This form of technology has been becoming increasingly popular as a result of the fact that that there are now many more individuals who are using computers, cellular telephones, and other types of personal devices. This method does have disadvantages, as the data can vary based on the mood of the individual, such as being angry or sad. Additionally, behavioral biometrics are not as accurate as physiological biometrics.

For these above-mentioned reasons, behavioral biometrics are not as popular as physiological biometrics. Moreover, behavioral biometrics is seen as a method to be used for verification instead of actual identification. This type of biometrics can be normally used with the passwords, identification cards, or personal identification numbers (PINs) of a company for the purpose of identification.



Figure 2.7 Keystroke dynamics recognition

2.2.2 Mouse Dynamics

Mouse dynamics likewise keystroke dynamics describes an individual's behavior; and unlike keystroke, mouse dynamics has been begun to be studied more extensively the past three decades. Mouse dynamics started to gain more interest since the importance of user identification and verification in today's Internet-centered world is being increasing, protection of our accounts, bank movements, information from stranger hands is a higher priority. The application of this method is not only limited by, how the name suggests, computer mouse but it is also applicable to touch pads, on the other hand compared to keystroke dynamics, the information considered is less and a disadvantage as well as keystroke dynamics is that is difficult to mimic what you have done previously [68].

2.2.3 Voice Recognition

Today, biometrics for voice recognition are the most important area of research. Acoustic biometrics have been defined as the biometrics of the recognition of the individual who is speaking, which are shown in Figure 2.8. They are used for phone-based applications. Almost everyone has features of the human voice, and in twins, the voice can be replicated perfectly. Each individual has unique vocal patterns that are produced through a combination of factors that are both behavioral and physical. The physical characteristics consist of things such as the lips, vocal system, the nasal cavity, and both the shape and the size of the mouth, while the behavioral characteristics consist of things such as articulation, the focus and speed of the speech, as well as the dialects used.

Voice recognition is based on the way in which an individual speaks, while the focus is aimed at the speech that is produced by the vocal features, not on the speech itself or the voice. There is no need for any additional, special, and more expensive hardware. The phonotype features of speech are used by voice recognition to distinguish individuals. These patterns consist of behaviors (speaking style and tone of voice) and physical (the shape and size of the throat and mouth) features [69, 70]. The acoustic device is not affected, even by the cold, thus there is no negative impact on the accuracy.



Figure 2.8 Voice Biometrics

While the voice is being recorded, the individual is instructed to repeat a series of short phrases that will be used to prevent unauthorized access. For this, audio devices, like phones, normal microphones, and computer microphones are able to be used to capture the required sound. When using an analog to digital converter, the generated electrical signal, as a result of using a microphone, is converted to a digital one and then it is recorded as a digital version the voice of the individual. An appropriate matching algorithm will then compare any input audio with the digital audio that was stored previously and make an identification.

There are two types of voice recognition that depend on the independence of the speaker and the amplifier. A loudspeaker-based system relies on knowing the specific acoustic features of an individual. This requires that the system is learned and trained with regards to these traits, through a vocal training, specifically with regard to the accent and pitch of the speech. The feature known as independent voice recognition can recognize certain speech, such as the words and/or phrases recorded

by different users that have limitations with regard to the speech context. No system training is needed. Text-dependent, text-prompted, and text-independent are the 3 spoken input styles used by the audio system.

The speaker-dependent voice recognition design is quite challenging. Various factors, such as ambient noise, headphone diversity, same speaker tone, voice input system sensitivity, distance, and uniformity differences are made use of in the process of speech recognition system performance analyses. Voice recognition systems are commonly utilized in healthcare services, government offices, banks and banking services, applications with regard to entertainment, the smart cards used for personal identification numbers, access control, the authentication of customers, and various other security purposes [71].

Due to the fact that the voice of an individual will change with age, the system needs to process this change in the voice. There is a big difference between headphone recognition and speech recognition. Voice or stethoscope recognition is the used to identify an individual through tone, pitch, and tone of voice. Speech recognition is the understanding of what is spoken and is used for menu or map navigation and hands-free computing.

The disadvantage of this method is that it offers low resolution and a person who is sick with something like the common cold can make recognition of the task nearly impossible because their voice will be affected by the disease.

2.2.4 Signature Recognition

The signature made by an individual is a form of human biometrics that can possibly change as the result of various factors, such as their mood or the environment they are in. This would result in the dissimilarity of 2 signatures from the same person. A signature verification system would provide a solution in a situation such as this. This system can be broken down into 3 basic stages, which consist of 1) the acquisition of the data combined with pre-processing, 2) extraction of the features, and 3) verification.

For the identification of an individual, signature is among the most inexpensive types of biometrics, followed by the use of DNA, fingerprints and palmprints, the face, vein patterns, and the retina and the iris. Physiological traits such as these remain almost exactly the same throughout the life of an individual, unlike the human signature, which may vary depending on the with mood or age of the individual, or the environment that they are in. If an individual does not sign their name consistently each time, then there will be difficulty with regard to the identification and verification of their signature. Moreover, the database must be updated or, in some cases, changed at specified times to ensure that the system is working properly. Additionally, in order for a database to be considered as good, it must consist of a series of signatures that are from an individual that are almost, but not exactly, similar to attain better verification. To obtain this series of signatures, the same individual must sign their name several times, and this signature will then be used as reference. Within the series, the majority of the characteristics must stay the same in order for the confidence level to be determined, and the accuracy to be measured [72]. On the other hand, this method does have a disadvantage, because the required hardware may be quite expensive, and some individuals have insufficient motor coordination to result in a consistent result while they are writing.

Signature is one of human biometrics that may change due to some factors, for example age, mood and environment, which means two signatures from a person cannot perfectly matching each other. A Signature Verification System (SVS) is a solution for such situation. The system can be decomposed into three stages: data acquisition and preprocessing, feature extraction and verification.

In the field of human identification, signature is one of the cheapest biometric besides DNA, fingerprint, palm print, face, vein pattern, retina, and iris. These physiological traits are almost unchanged throughout a person's life, unlike signature that may change with mood, environment and age. A person who does not sign in a consistent manner may have difficulty in identifying and verifying his/her signature. The database should be changed or updated in a few specified periods to make sure the verification system is working properly. In addition, a good database must have a series of signatures from a person that are almost similar between each other for

better verification. A series of signatures means a same person will sign several times and the signature will be kept as reference. In the series, many characteristics must remain constant to determine the confidence level, measured in accuracy [72].

On the other hand, a disadvantage of this method is that the hardware needed could be highly cost and that some people do not have enough motor coordination to have a constant result in the writing.

2.2.5 The Gait

Biometric gait recognition is a process that consists of recognizing an individual based on the way in which they walk, which is a type of behavioral biometrics. Moreover, it is a form of remote biometrics, as it allows an individual of interest to be filtered in public locations, which can include shops, stations, and airports [73]. However, this approach is affected by several confounding factors, including variations in the footwear that an individual is wearing, the terrain at that location, if they are fatigued or injured, and the passage of time, thus all making gait biometrics very complicated. Despite these mentioned factors, the exploration and investigation of gait biometrics recognition has been increasingly within the scope of gender classification and age estimation combined with medical diagnoses, respectively [74].

2.3 Vein Pattern Recognition Applications

Biometrics technology is widely used in various applications. The majority of people make an association between biometrics and security applications, like the physical and logical access to a location. It is indeed true that physical access to secure buildings, data centers, laboratories, and many other types of facilities is a very commonly used and visible classical application of biometrics. In addition, logical access to businesses and company networks, personal computers, financial accounts (such as ATMs), and other types of virtual systems is an application that is both reliable and growing quickly. Obviously, DHV can easily solve physical and logical access control application issues [75]. However, there are many other applications, such as handprint recognition, which are a major problem for many companies, in which biometric technology can be solved efficiently and conveniently. And

members of a particular organization must verify their identities to receive the benefits of membership. Metrics can be used to determine whether a person is already part of the database, such as B. People seeking benefits, driving licenses, or ID cards. The passenger plane, ship, bus, or train; maintain a chain of evidence; or subscribe to confidential documents. Despite the fact that this wide variety of categories not only suitable uses for biometrics, but they are also the most important today [76].

The earliest success stories in biometrics are related to visual revenue-generating applications, including things like PAC, resetting passwords (such as logical access), and time/attendance. Moreover, it must also be understood that, as with other life-saving applications, the vein pattern recognition system is used more effectively, not as a safety accessory or stand-alone system, but as a key component of an end-to-end safety solutions. Therefore, well-trained system integrators play a very important role in vein pattern recognition solutions. For users who do not use it or find a way to avoid it, using it may be unacceptable. There are 5 main types of applications that use vein pattern recognition-based biometric applications, which comprise physical access, logical access, personnel, membership, and accountability [77]. Vein detection biometric technology is an integral part of a comprehensive solution that is used to combat identity theft, fraud, and unauthorized access. This form of technology is highly reliable, and the degradation of its performance is only seen in harsh environments, such as at construction sites or at locations where there is heavy traffic, like at schools or military facilities [78].

1. Physical access control is in reference to the policy, the process, as well as the technical control that are involved in controlling access to ATMs, rooms, buildings, or other physically defined structures. IN this sense, protection means actual physical security, despite the fact that in many situations, the phrase ‘physical access’ brings to mind the images of ID cards. The ID card and the kinship card are used separately and are obsolete. In the next 10 years, the human body will replace the physical key. In more and more cases,

biometric data is being used in PAC applications at the entrances and exits of buildings, floors, and offices to restrict access to authorized personnel. Companies are increasingly choosing to incorporate biometric technology into the layering process as part of a comprehensive security plan. Layering refers to the simultaneous use of biometric and non-biometric mechanisms of security, which include physical codes, such as radio-frequency identification or smart cards, and a PIN or a password. The layer provides higher security and flexibility when using authentication technology. The most commercially-available biometric application is physical access, which has the largest market share of all of the horizontal applications. The most used physical access solutions are manual design and fingerprint scanning. However, this is a market that can shift rapidly to other biometrics, like iris recognition. Scanning and vein pattern recognition are performed based on accuracy, ease of use, and user acceptance requirements. These standards are the main considerations for people [79]. Before an organization can require its employees to submit to biometric identifier collection, so as to be granted access to the resources of the organization, it may be useful to consider what the employees think about it, and other related aspects of it. Organizations can reduce the anxiety that is experienced by the users, that is, the employees, via the implementation of biometric solutions. The solution should consider the concerns of the individuals, including their privacy, their need for solutions with regard to their health, and any other cultural concerns that may arise. Currently, financial institutions, governments, and highly secure locations, such as data centers and power plants, constitute the most active

markets that are making use of biometric PAC systems. However, PAC applications are also utilized in other areas as well, such as luxury apartments, hotels, and villas. It is possible to integrate vein pattern recognition technology into the PAC system, such as in electronic door locking systems. Most vein pattern recognition access controls include vein image sensors, small screens, additional keyboards, and controllers that perform authentication processing and direct unlocking commands. The vein pattern recognition system can be directly connected to an electronic lock itself or to the security system that actually controls the lock.

2. **Ports** Airports, seaports, and land border points can be distinguished from each other in a very customer-friendly way and provide more convenience and security through biometric technology. Airports as well as seaports already use this kind of technology in the fields of customs, immigration, and quarantine. The vein pattern recognition scanner is obsolete. Border crossings usually use multi-biological methods in large-scale applications. An example of vein pattern recognition technology is the monitoring of passengers or personnel, such as at the Incheon International Airport in South Korea, while another is the Port of Halifax, in Nova Scotia, Canada [80].
3. **Financial Institutions** Most economic establishments care about safety and are actively searching for and putting into effect multi-aspect authentication solutions. American and Swiss banks have added get entry to manage to safes and different key regions with the again of Identical handheld scanners to update keys and input composite codes. Banco Bradesco, which is the biggest personal financial institution in Brazil, has incorporated Fujitsus

PalmSecure (palm vein scanners) into its ATMs, as it has a large quantity of different banks across the world. Financial security, inclusive of ATMs, price terminals, cashless systems, computerized test cashing, etc. Fourth, they're the packages being evaluated for equipping biometric systems. Banking is increasing assist for domestic solutions, that allows you to growth the want for precise biometric identifiers for logical get entry to person banking information. Like many establishments, economic establishments are actively searching for possibilities to offer new or advanced possibilities to their clients. Biometrics assist them in reaching this goal. SunFirst, which is located in St. George, Utah, USA, is the fastest developing network of financial institutions [81]. It has terrific popularity as a technologically superior economic institution. It uses the identical guide vein machine to defend access to the banks (IT) facts middle. The use of this vein pattern recognition machine affords an extra layer of safety for the financial institution and its clients, while at the same time preserving the facts middle control continually aware about who enters and exits for the duration of any given duration of time. SunFirst selected its vein pattern recognition answer as it become accurate, smooth to use, and had a quick and easy registration procedure. Therefore, as the need for more secure banking security has arisen with regard to ATMs, both the depositors and the banks have been rapidly following biometric solutions. As a matter of fact, the procedure started even earlier than the brand-new regulation become passed. Due to a cultural bias towards fingerprinting and the choice of fitness solutions, international banks are hastily adopting vein pattern recognition.

4. Buildings In Singapore, some famous homes, which include IBM Singapore, Caltex Tower, and Mizuho Bank, have followed Hitachi's venous structures as principal additives in their bodily get right of entry to protection structures. The University of Tokyo Hospital has introduced Fujitsu palm vein scanners to stable get right of entry to its room, rather than its previously and primarily used fingerprint-based system. A palm vein scanner has been set-up in the doorway of every room to prevent unauthorized entry. Most apartments, hotels, and motels have stopped using from normal keys and have begun to use digital keys, which are commonly primarily based completely on magnetic stripe cards. Building gets right of entry to is now migrated to biometric-enabled proximity playing cards and clever playing cards, or to biometrics alone [82]. Highly stable locations might also additionally require an extra guide for blanketed assets and accordingly, might also additionally screen individuals who constantly enter through more than one check point. For example, vein pattern recognition can help legal court docket employees to effortlessly gain access to the rooms, facilities, and areas of a secured building. Forgotten to get right of entry to codes or disclosure of PINS (accidental or intentional) grow to be non-elaborate whilst vein pattern recognition calls for verification of a people presence. This safety can expand to courtroom docket records structures [83].
5. Safety deposit boxes A safety deposit box is a private and individually secured container that can be used to store essential papers and a variety of valuables, such as jewelry. Safety deposit packing containers are approximately privateness and luxury at the same time as shielding one's

valuables in a secure and loss-resistant environment [84]. In a primary departure from the conventional card token and signature get admission to, biometric era is converting the manner wherein stable deposit clients can get admission to their funds. In banks, disposing of guide steps in signature card retrieval to offer clients with get admission to their belongings saves precious time, and clients price the velocity of get admission to. Not most effective can biometrics authenticate a visitor; however, they also can keep a digital file of everybody who enters or exits a protection deposit field. Many monetary establishments and lots of impartial protection deposit field operators at the moment are the usage of biometric locker access structures to permit without difficulty and licensed get admission to to secure deposit packing containers. Clients also are nicely privy to the dangers of identification theft. Simply the usage of biometrics gives some other layer of protection to shield valuables and purchaser identities. Using biometric identifiers enhances the safety of a safety deposit box, in addition to increasing comfort and privacy. Vein sample reputation era is the most powerful choice to guide this application.

2.4 Deep Learning

The field of DL includes NNs and similar types of machine learning (ML) algorithms that are comprised of 1, or more than 1, hidden layer (HL). That is to say, it comprises using a minimum of 1 artificial NN (ANN), and the computer then obtains some new data from the already available data through the use of many algorithms. DL that is supervised, semi-supervised, or unsupervised should be performed. Deep artificial NNs (DANNs) also have proved successful through a reinforcement learning approach [85]. DL involves the use of computational models that consists of a number of processing layers so as to learn to display the data with more than one level of abstraction. Technology such as this has improved dramatically the most

advanced technology that is used in a number of other areas, including speech recognition, visual object and object detection, and genomic and drug detection. DL that is used in combination with the reverse diffusion algorithm (RDA) is able to examine complex structures within a big dataset and make a determination about how a certain machine should restructure its internal parameters that are used in the calculation of the display that is in each layer of the display that was in the layer before it. Deep convolutional networks (DCNs) have revolutionized speech, image, audio, and video processing, whereas recurring networks deal with sequential data, like text and speech [85–87].

Despite the fact that DL methods are usually much better than traditional methods of biometric detection, some technical problems exist, as well as obvious issues, which include the availability, or lack thereof, of high-quality labeled training samples, its high computational cost, large storage that is required, hardware requirements, low degree of portability, and the complexity of the design model. Hence, making an attempt to solve these above-mentioned challenges comprises the future process of DL. There are a number of new types of problems with ML, which include semi-supervised, self-supervised, and uncontrolled learning, that have been researched in an attempt to reduce dependence that currently exists on labeled learning items. Compression technologies, including those like quantization, pruning, and data distillation, have been used to try to reduce both the computational and the storage costs. Moreover, light-weight DL models, including ShuffleNet and MobileNet were designed to be used in mobile environments and/or to increase their portability. In addition, much research is currently being aimed at making an attempt to improve the DL model interpretability. Therefore, due to their rapid development, DL technologies will definitely be an integral part of biometric cognition [88].

The DL method comprises just one method that has been successfully developed that is on the basis of computer systems in recent history [89]. This method allows there to be abstraction of multiple data points. It is also made up of a number of computational models and layers. This technique has been made use of in a number of different areas, such as translation, and the field of visual, voice, and speech object recognition and perception. Moreover, biometric identification is utilized to a great

extent in the field. Additionally, it has also been used in fields that include medicine and genes in industry [90]. Via utilization of the advantages of DL techniques, this method is able to extract and complete the features that are sought after in studies that have been conducted in line with that aim. As a result of this feature, it is able to be used in a wide variety of biometric applications as well as in industry. Additionally, it is able to provide a significant advantage to its users as a result of the fact that it comprises features such as natural language recognition and processing [87]. DL is used a great deal in biometrics, and with its use comes some great advantages. Even though the DL method performs quite a bit better than the other methods, there are still some disadvantages, which mainly include the need for human resources who have higher education levels or more qualified people, expenses such as calculation and storage costs, the hardware requirements of the computers that are being used, in addition to its portability and its complexity. These constitute the main problems that DL should aim to solve in future [85]. DL models are currently being used quite extensively in the electronic systems that are used for security purposes over recent years. These main systems comprise telephones and computers, as well as authentication and security. [90]. Due to the fact that the size of NNs is a critical factor, DL requires the use of high-performance software and hardware infrastructures. The areas of use for DL, in addition to positive and negative aspects of using it have been reported in the literature [91, 92].

2.4.1 Classification of DL Approaches

DL techniques are classified into 3 major categories, which are unsupervised, partially supervised (semi-supervised), and supervised. Furthermore, deep reinforcement learning (DRL), also known as RL, is another type of learning technique, which is mostly considered to fall into the category of partially supervised (and occasionally unsupervised) learning techniques.

2.4.1.1 Deep Supervised Learning

This technique deals with labeled data. When considering such a technique, the environs have a collection of inputs and resultant outputs $(x_i, y_i) : \rho$. For instance, the smart agent guesses if the input is x_i and will obtain it as a loss value. Next, the network parameters are repeatedly updated by the agent to obtain an improved estimate for the preferred outputs. Following a positive training outcome, the agent acquires the ability to obtain the right solutions to the queries from the environs. For DL, there are several supervised learning techniques, such as recurrent neural networks (RNNs), convolutional neural networks (CNNs), and deep neural networks (DNNs).

In addition, the RNN category includes gated recurrent units (GRUs) and long short-term memory (LSTM) approaches. The main advantage of this technique is the ability to collect data or generate a data output from the prior knowledge. However, the disadvantage of this technique is that decision boundary might be overstrained when training set does not own samples that should be in a class. Overall, this technique is simpler than other techniques by way of learning with high performance.

2.4.1.2 Deep Semi-Supervised Learning

In this technique, the learning process is based on semi-labeled datasets. Occasionally, GANs and DRL are employed in the same way as this technique. In addition, RNNs, which include GRUs and LSTMs, are also employed for partially supervised learning. One of the advantages of this technique is to minimize the amount of labeled data needed. On other the hand, one of the disadvantages of this technique is irrelevant input feature present training data could furnish incorrect decisions. Text document classifier is one of the most popular examples of an application of semi-supervised learning. Due to difficulty of obtaining a large amount of labeled text documents, semi-supervised learning is ideal for text document classification task.

2.4.1.3 Deep Unsupervised Learning

This technique makes it possible to implement the learning process in the absence of available labeled data (i.e., no labels are required). Here, the agent learns the significant features or interior representation required to discover the unidentified structure or relationships in the input data. Techniques of generative networks, dimensionality reduction and clustering are frequently counted within the category of unsupervised learning. Several members of the DL family have performed well on non-linear dimensionality reduction and clustering tasks, which include restricted Boltzmann machines (RBMs), auto-encoders, and GANs as the most recently developed techniques. Moreover, RNNs, including the GRUs and LSTM approaches, have also been employed for unsupervised learning in a wide range of applications. The main disadvantages of unsupervised learning are unable to provide accurate information concerning data sorting and computationally complex. One of the most popular unsupervised learning approaches is clustering [93].

2.4.1.4 Deep Reinforcement Learning

Reinforcement learning operates on interactions with the environment, while supervised learning operates on provided sample data. This technique was developed in 2013 with Google Deep Mind [94]. Subsequently, many enhanced techniques dependent on reinforcement learning were constructed. For example, if the input environment samples: $x_t \sim p$, agent predict, and the received cost of the agent is P , here is the unknown probability distribution, then the environment asks a question to the agent. The answer it gives is a noisy score. This method is sometimes referred to as semi-supervised learning. Based on this concept, several supervised and unsupervised techniques were developed. In comparison with traditional supervised techniques, performing this type of learning is much more difficult, as no straightforward loss function is available in the reinforcement learning technique. In addition, there are two essential differences between supervised learning and reinforcement learning: first, there is no complete access to the function, which requires optimization, meaning that it should be queried via interaction; and second,

the state being interacted with is founded on an environment, where the input x_t is based on the preceding actions [95, 96].

For solving a task, the selection of the type of reinforcement learning that needs to be performed is based on the space or the scope of the problem. For example, DRL is the best way for problems involving many parameters to be optimized. By contrast, derivative-free reinforcement learning is a technique that performs well for problems with limited parameters. Some of the applications of reinforcement learning are business strategy planning and robotics for industrial automation. The main drawback of reinforcement learning is that parameters may influence the speed of learning. Below are the main motivations for utilizing reinforcement learning:

- It assists you in identifying which action produces the highest reward over a longer period.
- It assists you in discovering which situation requires action.
- It also allows you to figure out the best approach for reaching large rewards.
- It also gives the learning agent a reward function.
- It cannot be utilized in all situations, such as:
 - In cases in which there is sufficient data to resolve the issue with supervised learning techniques.
 - It is computing-heavy and time-consuming, especially when the workspace is large.

The biometric authentication method is more reliable and easier than other methods [90]. It offers an end-to-end learning paradigm for combining DL, feature extraction and preprocessing and recognition based solely on biometric data [89]. Among the main advantages of DL is a strong learning ability, broad scope, adaptability, data-driven nature, good transferability, etc.

An automated biometric system aims to either correctly predict the identity of the instance of a modality or verify whether the given sample is the same as the existing sample stored in the database. Figure 2.9 presents a traditional pipeline of a biometric authentication system. Input data corresponds to the raw data obtained directly from the sensor(s). Segmentation, or detection, refers to the process of extracting the region of interest from the given input. Once the required region of interest has been extracted, it is preprocessed to remove the noise, enhance the image, and normalize the data for the subsequent regions. After segmentation and preprocessing, the next step in the pipeline is feature extraction. Feature extraction refers to the process of extracting unique and discriminatory information from the given data. These features are then used for performing the classification. Classification refers to the process of creating a model, which given a seen/unseen input feature vector can provide its correct label [97].

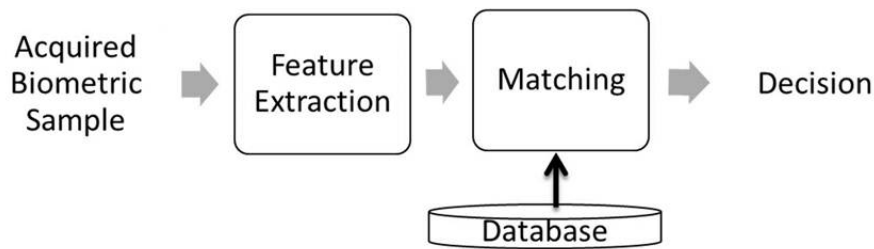


Figure 2.9 Biometric Authentication System

In this thesis, DL is used as a means to verify the advantages of modern technology in biometrics. These features are features that fit the auto pattern and provide an estimate of the density function. DL is integrated with the GAN to identify individuals through the dorsal vein. GAN is the latest model in the DL family.

2.5 Generative Adversarial Network (GAN)

GANs, which are generative models, were first presented by Goodfellow et al. [98]. Since their introduction, a number of applications of interest have appeared in the field of image creation as well as other areas, and GANs have been used to perform a range of computer vision tasks.

The main aspect with regard to GANs is that it has a zero-sum game and a maximum of 2 players. In the game, 1 of the players is granted benefits that are equal to if they

had lost the game to the 2nd player. Here, players are compatible with different GANs, which are called the discriminator and generator. The main objective for the discriminator is determining if the sample belongs to a distribution that is imaginary or one that is real. On the other hand, the generator has the aim of fooling the discriminator via the generation of a ‘dummy’ sample distribution. The discriminator produces the likelihood or chances that any given sample is a real sample. If there is a high probability value, then this indicates that the sample is quite likely to be a real one. Any value that is close to zero is an indication that the sample is imaginary. Any probability value that is close to 0.5 is an indication that the optimal solution has been created, which means that the discriminator was not able to distinguish between an imaginary sample and a real one.

Generally, obstetric models can be classified into 2 basic classes, the first of which comprises a traditional ML-based generation algorithm, which is made up of the RBM [99], naïve Bayes model [100], and hidden Markov model [101]. Some other examples also include DL models that comprise an automatic encoder [59] and GANs, as well as their derivative models. The GAN is a generative model which is able to generate target data that have latent variables. More specifically, the game training is conducted between the discriminator and generator in this type of model, and the generation of target variables is performed, using the real data distribution, by random variables that most usually follow a Gauss distribution. When a comparison is made with the traditional ML algorithm, this model is simpler and more effective, and it contains a greater number of application scenarios. Moreover, it can provide much better performance on large data sets, e.g., ImageNet [103] and CIFAR-100, when it is compared with the traditional algorithms.

It should also be noted that, through further research into GANs, it was determined that not only can they rapidly develop in image and video processing but can also be made use of in the fields of speech, text, and signal processing, including things like text-synthesized images, speech super-resolution, as well as EEG and ECG signal recognition. The recently published research studies have reported results that show how GANs are able to be combined in the signal communication field and have resulted in combinatorial optimization design [102]. Hence, this means that GANs

and their derivative models possess very strong vitality, and research regarding their applications have only begun [104], and have broad prospects for application in future development.

2.5.1 GAN Network Structure

GANs are popular due to the uniqueness of the architecture of their network model. GANs consist of 2 main parts, which comprise the generator and the discriminator. Both the generator and the discriminator are normally implemented by using NNs [105]. The inspiration behind the GAN model was derived from what is known as the mini-max 2-player game. The generators are responsible for generating samples that have almost real data distribution using random data. The discriminators must then discriminate between the real samples and the fake samples. The game training is utilized in the optimization of the weight parameters of the model between both of the networks so as to improve the ability that the model has to perform generalization. And then finally, the data distribution for the fake samples, which were generated by the generator, will be better aligned with the data distribution for the real samples, whereas the discriminator can use half of the probability when distinguishing the real samples and can use the other half when distinguishing the fake samples. For this model, an ideal state would be when the discriminator is not able to distinguish the fake samples from the real ones, which would result in it achieving an indivisible equilibrium state [106]. Figure 2.10 depicts the GAN network architecture. The implicit variable, Z , generates the fake samples, $G(z)$, using the generator. Then, the discriminator must determine the degree of authenticity of 2 types of input data, which of course, comprise the real and fake samples. Then, the discriminator must output the discrimination probability of the real and fake samples via the use of the objective function, and, as a last step, it must optimize both of the network structures. The GAN loss function was designed on the basis of the mini-max game of 2 different gamers, and it includes 2 NNs that must compete within the framework of a 0-sum game [107]. The discriminator must then to distinguish the data used for the input from the real data and then optimize the network model weights via the use of the backpropagation algorithm. The input

parameters that are used by the discriminator consist of X and θ^D , while the loss function used by the discriminator is given below:

$$V(D, \theta^D) = -E_{x: P_r(X)}[\log(1 - D(G(Z)))] \quad (2.1)$$

Here, P_r is the data distribution of the real samples. P_g is the data distribution of the fake samples that were generated via the generator. The input parameters that are used by the generator consist of Z and θ^G , while used by the generator is given below:

$$V(G, \theta^G) = -E_{z: P_g}[-\log(D(G(Z)))] \quad (2.2)$$

Both the generator and the discriminator optimize the weights, θ^G and θ^D , respectively, of each of the models by using the loss function. Neither the generator nor the discriminator changes each other's parameters for the model while they are being trained. The GAN will stop the training until both of the network structures are able to reach Nash equilibrium [108]. In the model for the GAN, there are 2 types of network structures that are trained by using the adversarial model. The GAN model's final objective function is as is shown in the equation below:

$$\min_G \max_D V(D, G) = -E_{x: P_{data}(x)}[\log D(x)] + E_{z: P_z(z)}[\log(1 - D(G(Z)))] \quad (2.3)$$

In the training of the GAN, both the generator and the discriminator are trained in an alternating manner. The discriminator, D , is trained first, and then, after that, the generator, G , is trained. At the point where one of these network structures is being trained, then the other structure is being fixed; thus, this alternates the training of the 2 networks, one-by-one. Now, theoretically, to be able to obtain an optimal solution for $V(D, G)$, first, the discriminator ought to be trained K times, and following that, then the generator ought to be trained just 1 time. However, in actual practice, when K equals 1, it has been observed to be better suited.

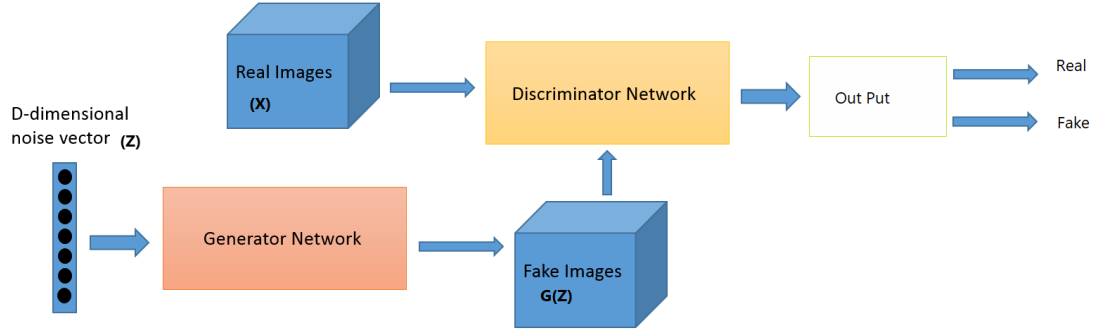


Figure 2.10 GAN network architecture

The cross-entropy loss problem $V(G, D)$ is commonly seen in classifications of a binary nature. Using this formula, it can be seen that this binary classification problem is for the discriminator, D . To be able to ensure that $V(G, D)$ is able to achieve a maximum value, it is possible to obtain $D^*(x)$ in the equation below following the derivation of $V(G, D)$:

$$D^*(x) = \frac{P_r(x)}{P_r(x) + P_g(x)} \quad (2.4)$$

It is possible to obtain the KL divergence of $P_r(x)$ and $P_g(x)$ via the introduction of the formula above into the objective function, which can better explain the training of the model.

2.5.2 Conditional Generative Adversarial Network (CGAN)

The basic architecture of the GAN [109] has no control over the modes of the data that are being generated. Van den Oord et al. [110], in their study, argued that the generation of class-conditioned images is able to significantly improve the quality of the images that are generated. A number of conditional-based GANs (cGANs) have been put forth that have learned to take samples from a conditional distribution, $p(x/y)$, rather than a marginal one, $p(x)$.

Variants of cGANs, as are shown in Figure. 2.11, are classified into 2 main types, which comprise supervised and unsupervised. The variants of the supervised type of cGANs must be given a pair of images as well as prior information like a class label

that corresponds to them. The prior information can include things such as textual descriptions, class labels, or data that has been taken from other modalities.

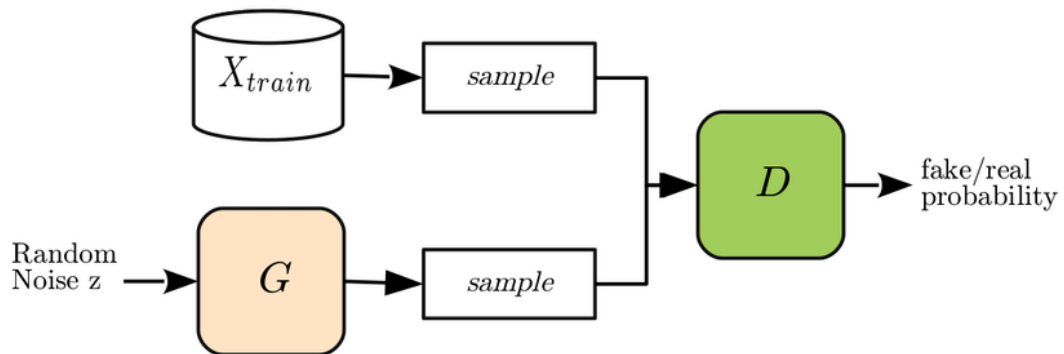


Figure 2.11 Conditional GAN

The authors in [111] proposed the cGAN, to be able to have some a control over the type of data that was being generated, via conditioning of the model on the prior information, y . In cGAN, the discriminator and the generator are both conditioned via feeding y into it as some additional input. By making use of this prior information, the cGAN will then generate some output images that have the desired properties.

The auxiliary classifier GAN (ACGAN) [112] comprises an architecture extension of the classical cGAN. In the ACGAN, the discriminator is given just the image, which is unlike the cGAN, which is given the image as well as the class label as the input. It is also modified to be able to distinguish the fake and real data, and to also reconstruct the class labels. Hence, in addition to the real/fake discrimination task, the discriminator is also able to predict the class label of the image via the use of an auxiliary decoder network.

CGAN is widely used in text, image, video, and prediction creation. Here, it will be explained how the proposed CGAN can help to generate the distribution of DHV images and then match them with an original image so as to be able to identify the image that will produce a lower number of matching errors. Using the method proposed herein, biocapture and DL training were both performed by the GAN, because GANs are very successful in DL and image processing. Moreover, these

networks are highly accurate as a result of increasing the size of the training data. As far as is known, this was the first time that a CGAN was implemented in this context.

2.6 Related Works

The literature contains studies on methods to be used in dorsal vein detection that generally follow procedures that are similar, such as feature extraction, pre-processing, and the measurement of similarity. These methods that have been described can roughly be separated into 2 main groups based on their characteristics in showing the vascular patterns, which comprise shape and texture. In this context, shape focuses on using the shape features that derive from the hand vein network and its structural composition to identify an individual. In some of the studies that were found in the literature, as well as early searches, it was found that the angles and positions of the short and straight vectors [113], as well as the small points [59, 114, 115], were used in the description of the distinction that was between the vessel shapes, which also included the intersections and endpoints. The authors in [116] put forth the proposal of a graphical representation of the shape of the vessels to be used for recognition, which was able to encode the resemblance of the information as well as the line segments, simultaneously, with just one model. Extraction of the region of interest (ROI) text was performed using the image of the vein for the latter, which was done as a function that was used in discerning various individuals. It is possible to use textures in an initial analysis of both of these main components, i.e. principal component analysis (PCA) [117], and in the linear discriminant analysis (LDA), which is performed using the entire image [118], as well as in local binary models (LBP) [119] and scale-invariant feature transform (SIFT) in some specific areas [120, 121]. The above-mentioned methods are also categorized on both generally and locally with regard to how the functions should be extracted. In order to assess the similarity, systematic methods make use of the details in the tissue or the shape of the dorsal artery on the whole, as well as the LDA- [116], PCA- [117], and graph [122, 123], based methods that belong to this pattern. Despite all of this, they appear to have some susceptibility to changes with regard to the light, distortion, as well as the obstacles, which can induce degradation of the output. The local elements, which remove the texture or the form features, as well as the LBP and SIFT, in a specific

area possess more resistance to these factors; hence, they dominate a problem such as this.

Coding methods are also used to attempt to encrypt and encode the information that is in images, an example of which is the use of GABOR and wavelet filters. Hybrid approaches make use of 2 or additional methods. Combined methods make use of the numerous benefits of authentication methods. The authors in [123] reported a fractal dimension box-counting technique that can be used for identification, which is based on venous vessel patterns. In biometric-based systems, the extraction of features from images is considered to be the most significant step. Their experimental research, which was based on data that was taken from the Bosphorous dorsal venous database, was indicative of the fact that, with regard to other methods that are well-known, the approach that they proposed showed positive and encouraging results. This method has the advantage of using a back structure to conduct the authentication, and this results in it having increased accuracy. The method has a main challenge, which comprises the fact that it is highly complex. The authors in [124] presented palm vein recognition that was based on the use of a competitive coding scheme, which utilized a multi-scale local binary pattern as well as ant colony optimization. Moreover, the ant colony algorithm that they used in their method allowed for the possibility of blocking points existing that were related to the quality of the image and also allowed for contrast problems, which could be observed in the images from the near-IR (NIR) spectrum, to be ignored. Next, the images were pre-processed followed by sorting via the use of the MLBP method which was performed using a competitive coding program. Some experimental results that have been obtained using the MS-PolyU database have also shown that the method proposed was highly accurate in its ability to identify an individual. The research put forth by [125] proposed a new method that used fuzzy curvature in the detection of hand veins via the use of thermal imaging. They first used an image of an arm that was taken using a thermal camera, and then they divided the image into 2 parts. They were then filtered via the use of a Gaussian high pass filter so as to smooth out and improve the contrast. The main topic of IR thermal imaging comprises the fusion method via the use of the fuzzy inference system, which is used as an expert system. Despite the fact that this method is highly accurate, the mechanism that they

proposed necessitates the use of special hardware. Moreover, it is very expensive to implement. In the study presented by [126], they presented a recognition system that can be used for partially-occluded DHVs, via the use of improved biometric graph matching. A noteworthy advantage of using this method lies in the fact that some parts of the dorsal areas could be lost as the result of damage, changes in the pigmentation, or tattoos, and this would then affect its ability to identify an individual. Hence, they are reconstructed in this method. The experimental results that they reported showed a desirable error rate as well as accuracy. Moreover, if an individual does have tattoos, using this method, it is possible to identify the individual anyway. In the research study of [127], hand vein recognition of the wrist, via the use of a local-line binary pattern (LLBP), was conducted to be able to identify a person based on the information taken from their hand veins. These researchers aimed at using the LLBP method while performing the extraction of the features. In their research study, they made use of the data from 50 individual datasets, which were taken directly via the use of IR cameras. Their experimental study reported an accuracy rate of approximately 96.50% using this method. In [128], the authors proposed the use of an iterative deep NN (DNN) in the verification of hand veins. In their thesis [129], they proposed the use of an iterative deep belief network (DBN) in the derivation of the vessel characteristics that were based on primary tag data. Some samples were then automatically generated through the use of very minimal prior knowledge and then were modified repeatedly via the use of the DBN. The venous features that resulted from this process were then used in the reconstruction of the training dataset, which was then used for retraining of the network. The experimental results that were in 2 general databases were able to show that the method that they proposed was effective; however, the degree of its accuracy was heavily dependent on the prototypes that were used for the training. The authors in [130] presented the use of vein biometric recognition that was used on a smartphone. The conducted research of [131], introduced on-the-fly finger vein-based biometric recognition that used DNNs. They proposed a DL architecture that was capable of capturing the finger vein structure via the use of a set of inexpensive cameras. In their study, they proposed the use of a diagnostic framework that used convolutional NNs. The experimental results showed that an accuracy rate for this method that was is high;

however, a significant challenge was that additional cameras had to be used in the implementation of the design.

For the extraction of the features, 2 categories exist, which comprise texture- and shape-based features [123, 132]. Texture-based utilizes changes in the tissue, as well as statistical characteristics of the tissue in the identification of an individual. Li and Kang, as one example, introduced a number of improved LBP models that were used in defining the characteristics in ship models [133, 134]. Premalatha and Kumar, in their study, examined local phase quantization (LPQ), as well as its alternatives, to determine an identifier in the identification of an individual [135]. In another study, Wei made use of hyperspectral images that had been taken of the DHV in the extraction of discriminative local features [51]. Wang used an improved model of the SIFT method in their study of cross-device hand vein recognition [136]. Meng and Wang used a GABOR filter, as well as its variants, in the minimization of the rotation and the effects of both noise and shifting problems, and also the maximization of the discriminatory power of the feature extraction [88]. Generally, the algorithms that are related to texture-based extraction of the features place their focus on changes in the texture. However, the basic geometric shape of the ship model is not considered, which therefore indicates an area that can be improved in the detection accuracy [121]. The extraction of features based on the use of tissue does not comprise a fundamental distinction that can be found between individuals and it is not a strong feature that can be used in the event of changes that can occur in lighting conditions [85, 132].

The field of DL includes NNs and similar types of machine learning (ML) algorithms that are comprised of 1, or more than 1, hidden layer (HL). That is to say, it comprises using a minimum of 1 artificial NN (ANN), and the computer then obtains some new data from the already available data through the use of many algorithms. DL that is supervised, semi-supervised, or unsupervised should be performed. Deep artificial NNs (DANNs) also have proved successful through a reinforcement learning approach [85]. DL involves the use of computational models that consists of a number of processing layers so as to learn to display the data with more than one level of abstraction. Technology such as this has improved dramatically the most

advanced technology that is used in several other areas, including speech recognition, visual object and object detection, and genomic and drug detection. DL that is used in combination with the reverse diffusion algorithm (RDA) is able to examine complex structures within a big dataset and make a determination about how a certain machine should restructure its internal parameters that are used in the calculation of the display that is in each layer of the display that was in the layer before it. Deep convolutional networks (DCNs) have revolutionized speech, image, audio, and video processing, whereas recurring networks deal with sequential data, like text and speech [85, 105].

Despite the fact that DL methods are usually much better than traditional methods of biometric detection, some technical problems exist, as well as obvious issues, which include the availability, or lack thereof, of high-quality labeled training samples, its high computational cost, large storage that is required, hardware requirements, low degree of portability, and the complexity of the design model [137]. Hence, making an attempt to solve these above-mentioned challenges comprises the future process of DL. There are several new types of problems with ML, which include semi-supervised, self-supervised, and uncontrolled learning, that have been researched in an attempt to reduce dependence that currently exists on labeled learning items. Compression technologies, including those like quantization, pruning, and data distillation, have been used to try to reduce both the computational and the storage costs. Moreover, light-weight DL models, including ShuffleNet and MobileNet were designed to be used in mobile environments and/or to increase their portability. In addition, much research is currently being aimed at making an attempt to improve the DL model interpretability. Therefore, due to their rapid development, DL technologies will definitely be an integral part of biometric cognition [85].

The DL method comprises just one method that has been successfully developed that is on the basis of computer systems in recent history [89]. This method allows there to be abstraction of multiple data points. It is also made up of several computational models and layers. This technique has been made use of in a number of different areas, such as translation, and the field of visual, voice, and speech object recognition and perception. Moreover, biometric identification is utilized to a great extent in the

field. Additionally, it has also been used in fields that include medicine and genes in industry [90]. Via utilization of the advantages of DL techniques, this method is able to extract and complete the features that are sought after in studies that have been conducted in line with that aim. As a result of this feature, it is able to be used in a wide variety of biometric applications as well as in industry. Additionally, it is able to provide a significant advantage to its users as a result of the fact that it comprises features such as natural language recognition and processing [87].

DL is used a great deal in biometrics, and with its use comes some great advantages. Even though the DL method performs quite a bit better than the other methods, there are still some disadvantages, which mainly include the need for human resources who have higher education levels or more qualified people, expenses such as calculation and storage costs, the hardware requirements of the computers that are being used, in addition to its portability and its complexity. These constitute the main problems that DL should aim to solve in future [85].

DL models are currently being used quite extensively in the electronic systems that are used for security purposes over recent years. These main systems comprise telephones and computers, as well as authentication and security. [90]. Due to the fact that the size of NNs is a critical factor, DL requires the use of high-performance software and hardware infrastructures. The areas of use for DL, in addition to positive and negative aspects of using it have been reported in the literature [91, 92].

The biometric authentication method is more reliable and easier than other methods [90]. It offers an end-to-end learning paradigm for combining DL deep learning, feature extraction and preprocessing and recognition based solely on biometric data [93]. Among the main advantages of DL deep learning is a strong learning ability, broad scope, adaptability, data-driven nature, good transferability, etc.

In this study, the technological advantages of the DL deep learning method were used for authentication. These advantages are automatic pattern fit features and provide an estimate of the density function. GAN is the newer model of the DL deep learning family. In the application, using this DL deep learning feature will be compared with the applications in two databases. Detailed specifications of Jill

University and the 11K database were written. The experiments are illustrated on the Jilin University – DHV dorsal hand vein database (Table-4.2) and the 11K hands dataset (Table-4.4), the most comprehensive databases, and comparative results are achieved, which shows the advantage of deep generative adversarial network features over light features in DHVR dorsal hand vein recognition.

2.6 DATASET

Herein, 2 different databases were made use of in the evaluation of the methods proposed, which comprised the Jilin University-DHV database (JU-DHVD) [126] as well as the 11K Hands dataset (11K-HD) [138]. The JU-DHVD comprises a number of real images as well as a great number of artificially made images that can be used in analyses and evaluations of biometric methods. This dataset comprised approximately 150 standard sample training images as well as 100 test images, which can be made use of to teach the method that is proposed, as can be seen in Figure 2.12.

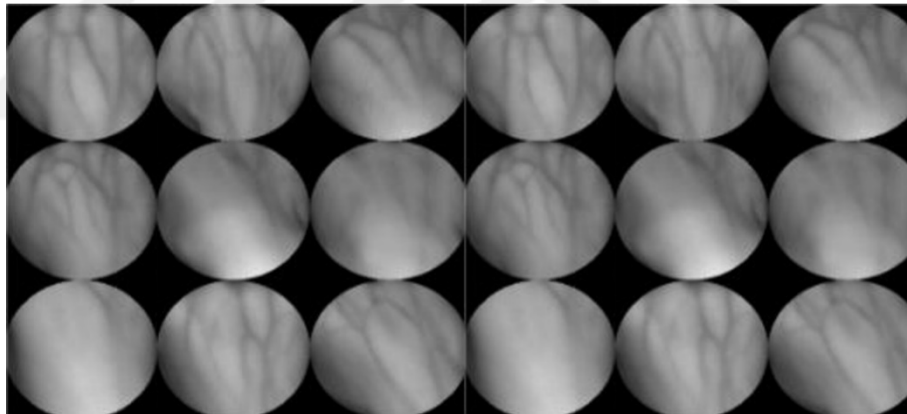


Figure 2.12 Images from the JU-DHVD

The 11K-HD comprises a set that consists of 11,076 manual images, each of which is 1600×1200 pixels, which were taken of 190 individuals who were between 18 and 75 years of age. Each of the individuals was instructed to open and then close their fingers on their right and left hands. The 11K-HD consists of 5680 images that were taken of the dorsal side of the hand; 2892 images that were taken of the right side of the hand, and 2788 images that were taken of the left hand; and 5396 images that were taken of the palm of the hand, including 2813 images that were taken of the palm on the right hand, and 2583 images that were taken of the palm of the left hand.

Each hand, both left and right, was imaged dorsally as well as palmarly, while using a plain white background with the hand placement for each of the hands at the same distance away from the camera. The dataset also consists of a number of manual images that have more detailed metadata, which can be seen in Figure 2.13 below. This dataset is available free of charge for academic use.



Figure 2.13 The 11K dataset

CHAPTER 3

METHODOLOGY

This chapter presents a biometric method that is based on venous processing for the identification of individuals. The method proposed herein makes use of GAN-based deep-forging technology for the identification of individuals, and a productive model is proposed to be used to identify individuals based on being able to identify their veins, which will be explained in the sections below.

As can be seen in Figure 3.1, this method consists of 7 steps, which comprise 1) reading of the image, 2) pre-processing, 3) the application of the threshold method, 4) use of the GAN, 5) use of the discriminators, 6) use of the generator classifiers, and 7) the performance evaluation. Each of these above-mentioned steps will be explained in detail below.

3.1 Image Acquisition

Image acquisition is understood as the data acquisition of the DHVs that define the process that begins from camera insertion to the final extracted information required by the system. These data should be representative of the dorsal veins of the individuals, and they should be able to present similar results between individuals to be able to select an appropriate identification system. The procedure for capturing information, converting it, extracting information about the user, and finally, comparing it with the database samples is the process of identifying the vein of the hand veins that are located on the back of the hand.

There are 2 ways to obtain images of the dorsal veins, and one of these methods is to use IVYRISE high-resolution cameras. These cameras can obtain from Amazon at an affordable price. Figure 3.1 presents a view from an IR camera to show how the veins of the hand of an individual appear (on the left). The same figure also shows an example of the dorsal veins, and the veins within them, of individuals after their initial treatment can be seen on the right [139].

The second method is to use specialized DHV databases for training and testing. In this thesis, 2 DHV databases were used, which were the JU-DHVD and the 11K-HD, which are available on the Internet for free.

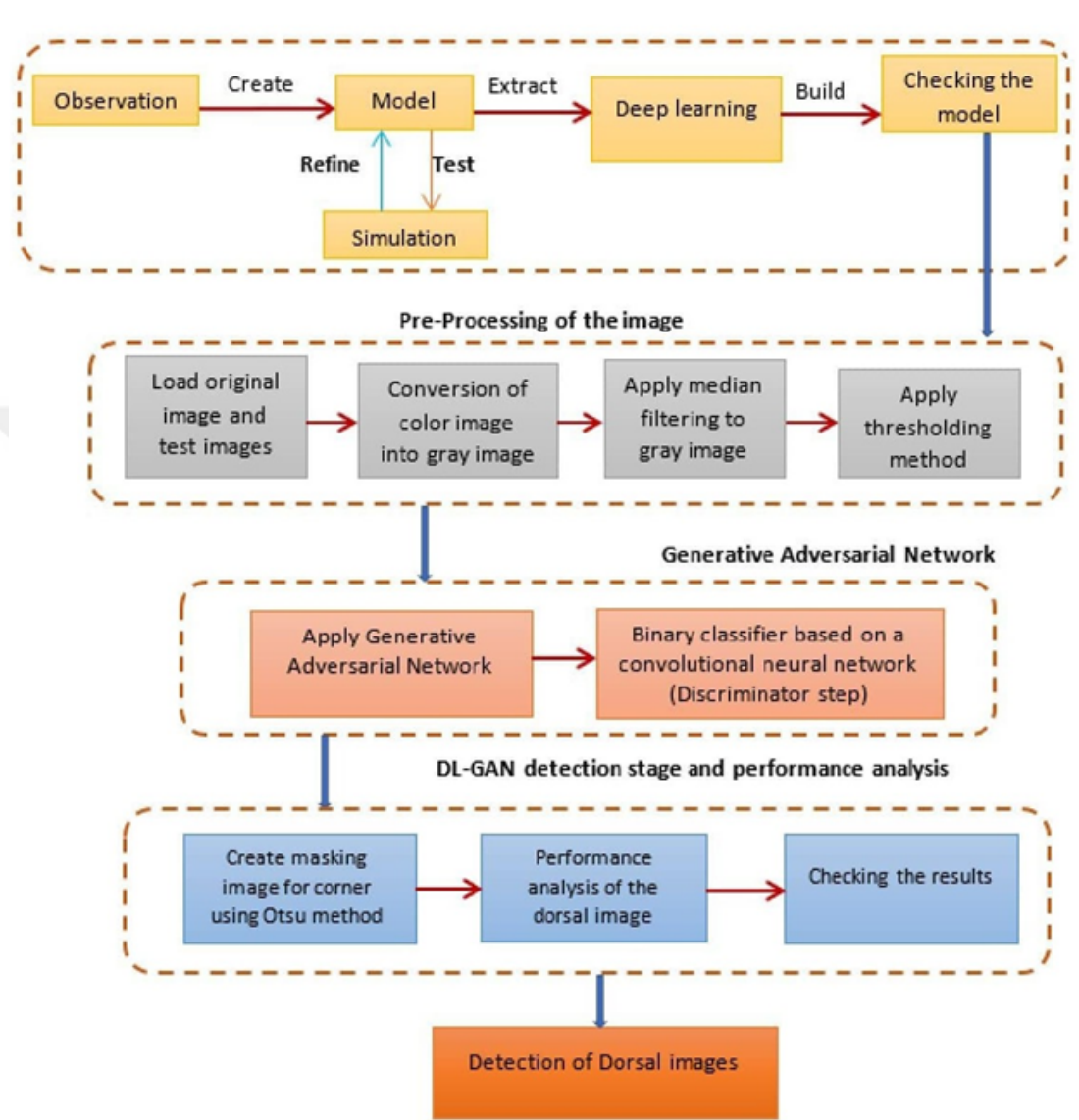


Figure 3.1 The DLGAN network that is used for identification of an individual based on the dorsal veins in their hand



Figure 3.2 Images showing the DHVs on the back of the hand using an IR camera

The position of the dorsal veins is under the skin. The patterns are usually extracted from the DHV. From the security side, the vein looks like an iris, but from the hand side, the DHV looks like a fingerprint. One of the exciting aspects of identifying a vein is the fact that the information is invisible, as it is hidden inside the body. A NIR sensitive optical sensor is used to capture the vein pattern image.

3.2 Preprocessing Images

Generally, most DHV images that have been collected have contained vein structures in addition to other information, including the fingers and knuckles, which were not made use of in this research.

Being able to measure the quality of the images is crucial in most of the image processing applications. Generally, a meter that is used to determine the image quality consists of 3 kinds of applications; hence it can be 1) used for monitoring of the quality of the image of a quality control system; 2) used for measuring image processing algorithms and systems; and 3) integrated into a system used for image processing to improve the algorithms and the parameter settings.

Once the vein images have been obtained, they cannot be used directly for the identification of an individual, as some pre-treatment processes must be performed beforehand due to the following reasons:

1. There may be some geometric differences, which can include rotation, shifting, and scale that is produced via the individual forming the different hand positions.
2. A ROI must be extracted if the image of the back of the hand contains any type of a background.
3. Irregular and unsteady lighting results in different image quality.
4. Noise inherent in photoreceptors. A schematic of the different stages of pre-treatment that must be performed is depicted in Figure 3.4. Each of these stages will also be described, in detail, in the subsections below.

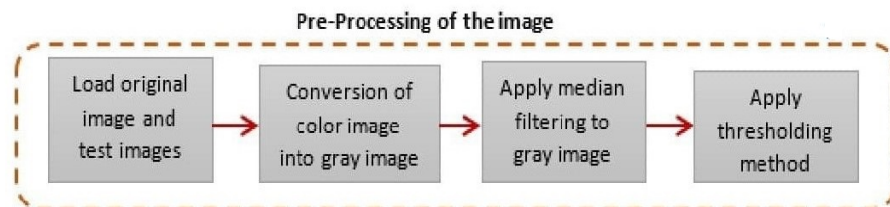


Figure 3.3 Pre-processing diagram

After the completion of this process, in this method, the output images should be good quality. However, during the process of obtaining these images, things such as the hand of the individual shaking or ambient noise will have an effect on the image quality. Therefore, these pre-processing steps must be performed on the images prior to the DL. This is a non-linear method that can be made use of during image processing to reduce the amount of salt and pepper noise. In this process, the intermediate filter moves through the pixels, pixel by pixel, and replaces each of the values with the average value that belongs to the pixels that is next to it. This neighbor style is what is known as a 'window'. Calculation of the median is performed by sorting all of the pixel values from the 'window' first, in the order in which they are numbered, and then by replacing that pixel with a medium pixel value. In this method, first, the images are converted to grayscale using a light intensity that is between 0 and 255. After this step, a median noise filter is used in the calculation of the midpoints of the neighborhood of the light intensity so as to eliminate the noise. In Eq. (3.1), as was suggested by [74], noise f :

$$f = \underset{(i,j) \in K}{\text{median}}\{H(i, j)\} \quad (3.1)$$

Here, K is the amount of adjacent and neighbor pixels of the central pixel; H is together with the noise of an image of a hand; f is the same as the image after the median filter. To be able to automatically authenticate the vein means that preprocessing of the images must occur first. If the quality of the image is enhanced prior to the authentication being performed, this will ensure the accuracy of the authentication. One of the methods that is used to increase image quality is the use of a histogram chart balancing method. Representation of the conversion function is given by $s = T(r)$, and then for balancing of the image, Eq. (3.2) can be used, which comprises an integral:

$$s = T(r) = (L-1) \int_0^r P_r(w) dw \quad (3.2)$$

Here, L is the level of light intensity that is highest within the image. This same value is 256 in the grayscale images. This relation can also be taken into consideration in the discrete state, which is represented as is shown in Eq (3.3):

$$S_k = T(r_k) = (L-1) \sum_{j=0}^k p_r(r_j) = \frac{(L-1)}{MN} \sum_{j=0}^k n_j, \quad k = 0, 1, 2, \dots, L-1 \quad (3.3)$$

Here, MN is the total number of pixels that are in the image; n_j is the number of pixels that have light intensity r_k . L is the number of brightness levels that are used in the image. Next, after performing the initial pre-processing and gray scaling of the images, the input image must be considered, in which the size will be normalized to 256×256 . In the 2nd stage, as shown in Figure 3.4, using the methods that can be utilized for hand edge detection, which include the Prewitt and Sobel method, extraction of the edge of the hand and surface area of the hand of the individual must be performed, which can then be subjected to further processing.



Figure 3.4 Shown above are images of a Gray scaled hand image and the veins within it (far left), the edge of the hand and the area of interest showing a Sobel edge (middle left), the segmentation area of interest extraction or the ROI by filling in the cavities (middle right), and the vein extraction of the hand (far right)

The steps listed below can be used in the extraction of the areas of interest and in identifying the vein pattern:

- ☆ **Frame capturing:** First a hand is inserted for a period of 2 or 3 seconds. After this, the cameras capture 34–36 frames of still images.
- ☆ **ROI extraction:** The image and the section of interest are extracted by hand. A circular section is used for extraction in some methods; however, such methods lack effectiveness because there are some very important vessels that can be used for recognition in the finger. The method proposed herein makes use of edge detection that is gradient-based, like Sobel, to obtain the edge of the image areas. The method proposed herein, it is possible to remove small areas from the image via thresholding of the size of the area of the hand and then filling in of the holes in the image.
- ☆ **Hand vein extraction:** Methods for image thresholding are used to perform hand vein extraction, which will be used as input into DL techniques. This thresholding of the image allows the hand veins to be detected as binary, and then they can be utilized as a model for learning in deep forging networks or GANs.

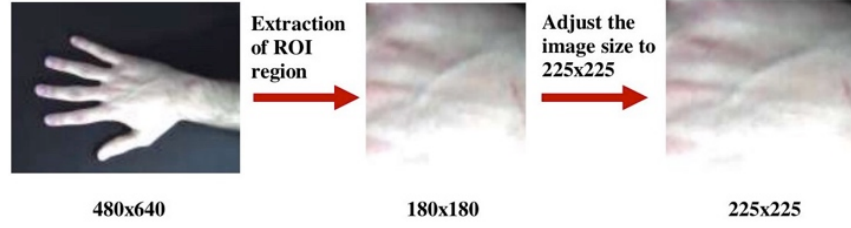


Figure 3.5 Dorsal image preprocessing

3.3 Threshold Method

One of the thresholding methods, known as Otsu thresholding, involves a procedure that is performed on the basis of statistical processes to be able to find optimal thresholds. For this method, the difference that exists between the mean of the light intensity and the variance in the light intensity on both sides of the thresholds are utilized for finding their optimal values. This is a definite method that can be used in the determination of thresholds. Using this algorithm, it is attempted to maximize the degree of variance in the light intensity on either side of the thresholds. When modeling the method that was proposed herein, an assumption was made that p_i was equal to the frequency of the number of pixels which had a light intensity of I , and could be calculated as shown in Eq. (3.4) below [140]:

$$p_i = \frac{n_i}{N}, p_i \geq 0, \sum_{i=0}^L p_i = 1 \quad (3.4)$$

Here, n_i represents the number of pixels that are within the entire image with a light intensity that is equal to I , N represents the total number of pixels within the image, and L represents the light intensity, which herein, was determined to be 255. Using the simplest example, it should be assumed that just one threshold exists, like t within the image, and this one threshold within the histogram is able to set the light intensity, from 0 to $t-1$, on the left. The light intensity, t to L , is also on the right, and in the gray scaled images. For Otsu thresholding, calculation of the total frequency or weight of the 2 classes on the left and the right of threshold t can be performed as in Eqs. (3.5) and (3.6), respectively, below. The sum of the 2 weights on the right and on the left of the classes is equal to 1, as can be seen in Eq. (3.7) below [140]:

$$w_0 = \sum_{i=0}^{t-1} p_i \quad (3.5)$$

$$w_1 = \sum_{i=t}^{L-1} p_i \quad (3.6)$$

$$w_0 + w_1 = 1 \quad (3.7)$$

Here, w_0 and w_1 represent the weighted classes for the left side and the right side, respectively. As the next step, calculation of the average weighted light intensity on the left side and the right side of the threshold is performed using to Eqs. (3.8) and (3.9), respectively. This can be used in the calculation of the average of these 2 classes of the left threshold as well as the right threshold, as can be seen in Eq. (3.10). By using the weights of these 2 classes as well as the calculated average, as shown in Eq. (3.11), an appropriate objective function can be defined with the goal of maximizing an optimal threshold [56], as shown below:

$$\mu_0 = \sum_{i=0}^{t-1} \frac{i \cdot p_i}{w_0} \quad (3.8)$$

$$\mu_1 = \sum_{i=t}^{L-1} \frac{i \cdot p_i}{w_1} \quad (3.9)$$

$$\mu_T = w_0 \mu_0 + w_1 \mu_1 \quad (3.10)$$

$$f = w_0 \cdot (\mu_0 - \mu_T)^2 + w_1 \cdot (\mu_1 - \mu_T)^2 \quad (3.11)$$

where μ_0 and μ_1 represent the average weighted light intensities within the pixels of the 2 classes, respectively, on the left side and the right side of the threshold; μ_T represents the average weighted light intensity of the 2 sides of the threshold; and f represents the objective function. If more than one threshold, like the m threshold, is used, then it is possible to present the objective function as seen in Eq. (3.12) below [124]:

$$f = w_0 \cdot (\mu_0 - \mu_T)^2 + w_1 \cdot (\mu_1 - \mu_T)^2 + \dots + w_m \cdot (\mu_m - \mu_T)^2 \quad (3.12)$$

where the aim is finding the optimal threshold that can be used to maximize the value of the objective function. These thresholds will thus be useable for performing the segmentation.

3.4 Generative Adversarial Network (GAN)

GANs consist of 2 networks, which comprise a generator and a splitter. GANs are the ultimate in advanced learning models [98]. Some algorithms to be used for artificial intelligence were developed by Goodfellow et al. in [91], including GANs. GANs comprise a group of algorithms that can be used for unsupervised learning. Their advantage is that they are able to produce images that are almost real by utilizing a scenario that comprises game theory and a 0-output model. Despite having these achieved such significant results, there are still significant challenges with regard to 3 main aspects when using GANs in real-world problems, which comprise 1) the creation of images of high quality, 2) being able to provide diversity, and 3) being able to obtain training that is sustainable, all of which will be focused on herein. Despite this, GANs still possess some problems. The most important of which is that they face difficulty when training then as well as when evaluating them. With these difficulties in their training, as well as maintaining balance between the product and the discriminator while training is not inconsequential. Generally, it is normal that the producer will learn well, i.e. to distribute the dataset fully [92].

Over recent years, DL methods that have utilized the GAN approach have presented a suitable tool that can be used for pattern recognition with high accuracy. Herein, the contribution of the current research is the use of a multi-step approach that includes both selection of the features and pre-processing of the image, in addition to the use of the GAN method, so as to identify individuals effectively using the information from DHV images. Research regarding biometric identity verification was conducted, and a method known as DLGAN was both developed and then tested using the same databases, but in different ways. Verification of the obtained results from the DLGAN was performed using different methods, but under the same conditions, and the results were compared. Although the implementation of the

DLGAN is an approach that can be used for the creation of many artificial as well as educational examples, one such areas for the use of GAN learning using this method of processing images.

The experimental results herein showed that there was an advantage of using the deep GAN features when compared to the light GAN features in the discrimination of the DHV networks. They also confirmed the necessity of performing the of the step of fine tuning.

GANs are able to provide improvements in the sequence generation with regard to representing the input features that are found in an entire body of work [141]. When the aim is to generate photo-quality images using real-world objects, LSTM RNNs are able to recall patterns as well as track the position in any generated sequence; however, they do have the potential of overcompensating, as it appears that there is a limit in their ability to be able to independently construct a complete work. The solution that was proposed solution in the research by Goodfellow added an extra level of careful inquiry to the sequence generation of the LSTM RNNs; hence, an RNN-based model was thus trained in order to critique the predicted output [141].

This was able to provide the adversarial nature that has been alluded to with regard to the naming convention of the GAN, which consists of 2 NNs being trained using the same data. The GAN that is presented in Figure 3.6 is made up of 2 different NNs. Of these 2, one has been abbreviated as G in the figure, which refers to the generator, while the other has been abbreviated as D in the figure, which refers to the discriminator. First, the generator produces some random images. Then, the discriminator conducts an examination of the images and instructs the generator as to the similarity of the images when compared to the actual image, as can be seen below [53].

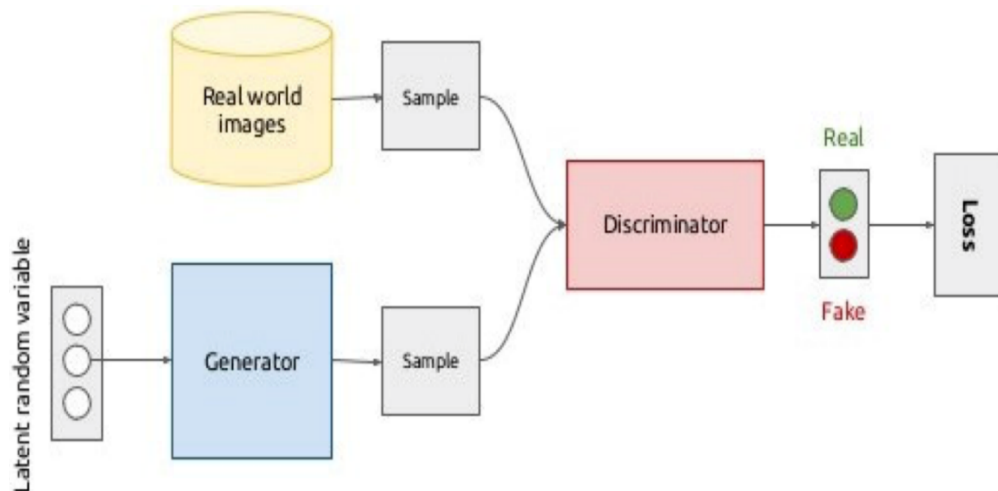


Figure 3.6 Learning in the generative adversarial network

In the first stage, the generating model uses the random noise signals as the input and then it will produce a distorted image as the output. Step by step, aided by the discriminator, it will create some images of the class that is desired, which appear to be real. Next, a distinctive network, which will act as a kind of competitor to the generator, uses these generated images of the generator, in addition to a certain class of real images, as the input and it will then evaluate to what degree the generated images will be evaluated with the real image. Then, after completing several stages of the network operation, a point is reached in which the discriminator is not able to recognize if the images that have been produced are in fact real or not. At this point, the generator produces some images of a certain class, which is actually the class that the discriminator was trained with, which did not previously exist. Hence, DL is similar to giving providing a thief with a picture of a hand. In the 2nd step, this image is checked against the image of the hand, and if similarity cannot be confirmed, then the image is given back to the forger, which will apply some noise and also slightly change it, so as to provide an increased level of adaptation with that image. This process, which comprises moving the hand image goes on to the point that the fake image will be a slight bit different than the real image. In the method that is proposed, the training is conducted at 2 different levels. The DHVs contain noise and they are matched against the original image in order to select an image that will produce a lower number of matching errors.

Moreover, A GAN is used to perform the DL for the biometric training and adaptation. GANs have very successful been used for DL and the processing of images. Furthermore, it is possible to achieve high accuracy in these networks via increasing of the size of the training data that are used. A limitation of many of the DL methods lies in the fact that they do require a considerable amount training data to be able to achieve high accuracy. Moreover, in most cases, insufficient training samples are available. As an example, there may be limited data that pertains to the effect of manual veins, which will therefore result in a decreased chance that the biometric accuracy will increase with a decrease in the number of training samples. On the other hand, an advantage of learning for a GAN lies in the fact that is able to generate a lot of random instances using the game theory. Classifiers like as CNNs usually necessitate the use of a large dataset for the training. GANs have the advantage of being able to create many of artificial samples for use in the training, resulting in increased accuracy, which will thus increase ability of this method with regard to image classification. Herein, a method is presented that is able to use these hostile networks. An estimation will be performed to determine if these artificial examples that have been added to the main set of training database images are able to improve the test database accuracy.

A major issue that can be observed in the training of DL algorithms for performing classification tasks on biometric-based images is class imbalance. A commonly used technique to address class imbalance when training the data is the geometric augmentation of existing images. Such augmentation consists of the use of geometric as well as other transformations in the creation of new images from the existing images, which has been proven as a technique used to improve the generalizability of DL models. That said, the optimal methods to be used for augmentation of the training images are usually unknown prior to training. Therefore, a great deal of tuning is required. Moreover, there are not many methods that can be used for augmenting a given image, and it is not possible to create an unlimited number of images by simply making use of data augmentation.

This all led to GANs being invented [140], which have recently gained a great deal of attention as a result of their ability to synthesize images that appear to be realistic from white noise vectors. The GAN architecture that was originally proposed is made up of 2 DCNNs, in which one competes against the other. Figure 3.7 presents the basic GAN architecture:

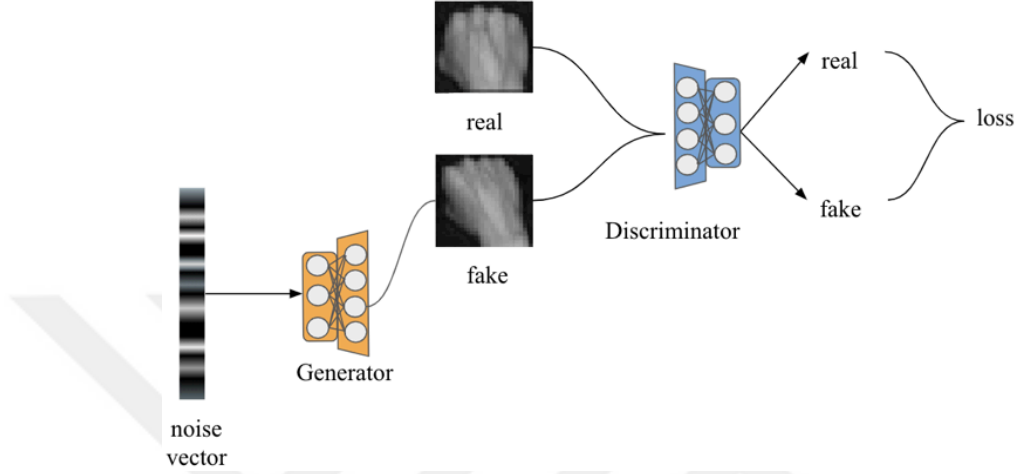


Figure 3.7 The architecture of the GAN. A false image is first generated by the generator from a random noise vector. Then, the discriminator faces the task of discriminating between the false images and those of the real training data

Figure 3.7 shows the 2 competing DCNNs, which comprise the generator, which generates the synthetic images from noise, and the discriminator, which discriminates the genuine samples from the false ones. Commonly, the learning objective of a GAN is for the generator to be able to fool the discriminator with the false images, which become increasingly realistic. From a mathematic perspective, this learning objective can be expressed as a minimax loss, as given below:

$$\min_G \max_D V(D, G) = E_{x: P_{data}(x)} [\log D(x)] + E_{z: P_z(z)} [\log(1 - D(G(z)))] \quad (3.13)$$

Here, x and z represent the inputs for the discriminator. In Eq. 13, the training objective for the discriminator is maximizing the $\log(D(x)) + \log(1 - D(G(z)))$, as well as the probability that correct labels will be assigned to training images as well as the fake ones that were generated by the generator. The generator is then trained to minimize the $\log(1 - D(G(z)))$, as well as the inverted log probability of the prediction of the fake images of the discriminator. Such minimization is difficult to implement; therefore, it is sought to instead maximize $D(G(z))$.

The original architecture of the GAN can perform synthetization of realistic images. But it is only able to synthesize them randomly and often experiences mode collapse (MC), which occurs in a situation in which the generator selects the easiest class within the dataset in order to fool the discriminator successfully. The images that result from this will lack diversity and they usually all are of the same class. In practice, MC occurs quite often as a result of class imbalance in the data used for the training. One method that can be used to address MC and correct the problems that appear in the unconditional image synthesis comprises the incorporation of side information. A cGAN is a very common kind of GAN in which a generator is used to conditionally generate images on the basis of their class labels [111]. The ACGAN is kind of cGAN that makes use of an extra auxiliary classifier for assigning the class labels correctly to the synthesized images [142]. Figure 3.8 depicts the general ACGAN architecture:

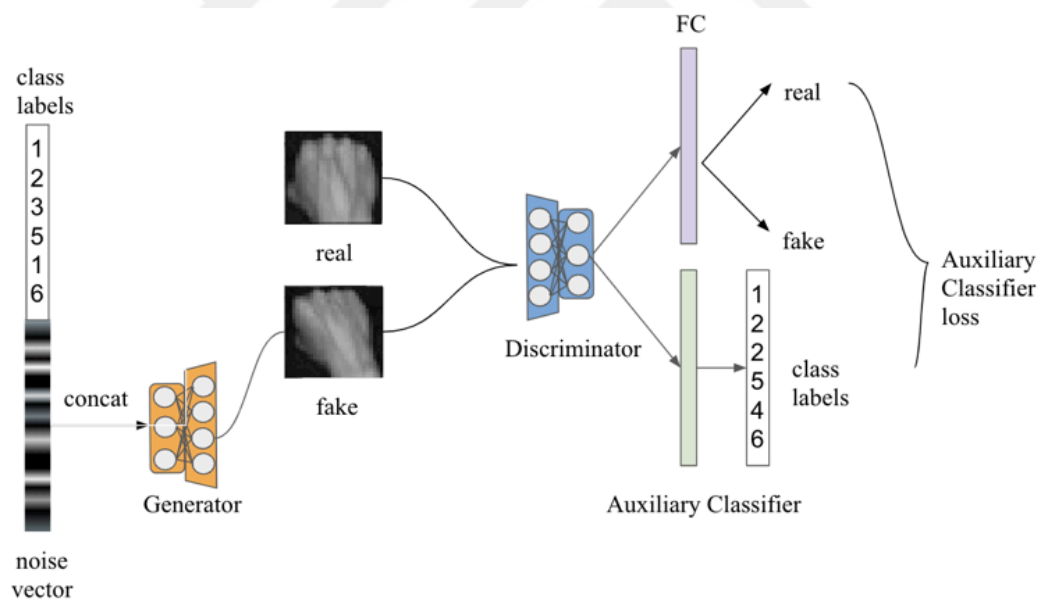


Figure 3.8 The general architecture of an ACGAN

ACGAN makes use of extra class labels while performing image synthesis. This is done conditionally so as to synthesize the images, as well as to bring about improved quality.

Because the ACGAN comprises an additional auxiliary classifier in its discriminator, its objection functions can be defined as given below:

$$L_s = E[\log P(S = \textit{real} \mid X_{\textit{real}})] + E[\log P(S = \textit{fake} \mid X_{\textit{fake}})] \quad (3.14)$$

$$L_c = E[\log P(C = c \mid X_{\textit{real}})] + E[\log P(C = c \mid X_{\textit{fake}})] \quad (3.15)$$

In addition to the production of probability distribution $P(S|X) = D(X)$ of the possible image sources, the discriminator of the ACGAN also consists of an auxiliary classifier that is capable of producing probability distribution $P(C|X) = D(X)$ of the class labels. The objective functions that were given in Eq. (3.14) and (3.15) can be defined as the log likelihoods of the correct source, L_s , and the correct class L_c , of the image.

Pix2pix is a commonly used variant of cGAN that is very popular. It is very often used in tasks with regard to image-to-image translation. Pix2pix was built on a commonly used segmentation network that is known as UNet and makes use of adversarial learning to be able to achieve a modality transfer. In Pix2pix, the generator is generally a UNet, or any other type of encoder/decoder network, and the discriminator is a convolutional-type Patch GAN classifier. Dissimilar to the other cGANs, pix2pix makes use of a dual objective function, which combines the adversarial loss with the L1 loss, as seen below:

$$G^* = \arg \min_G \max_D L_{cGAN}(G, D) + \lambda L1 \quad (3.16)$$

3.5 Discriminator

The discriminator uses an example that is taken from the domain as the input, which can be either real or generated, and then it predicts the binary class label for the real or fake/generated image as shown in Figure 3.9.

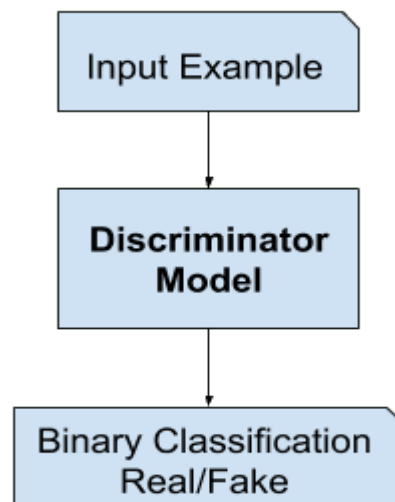


Figure 3.9 GAN Discriminator Model

The real examples are taken from the dataset used for the training. The examples that are generated are used as the output via the generator. The discriminator is a general classification-type model. Following the training process, the discriminator is not kept as only the generator is needed now. Occasionally, the generator can be reused, because it learned how to effectively extract the features from the examples within the problem domain. In fact, some, or sometimes, all the feature extraction layers are used to transfer the learning applications via the use of similar, or the same, input data [143].

The discriminator has also been conditioned, which means that it was given an input image, which was real or fake, as well as the additional input, which is shown in Figure 3.10 below. With a classification label-type conditional input, it would be expected by the discriminator that the input is from the same class, which would, in turn, teach the generator to begin to generate examples of that class so that it could fool the discriminator.

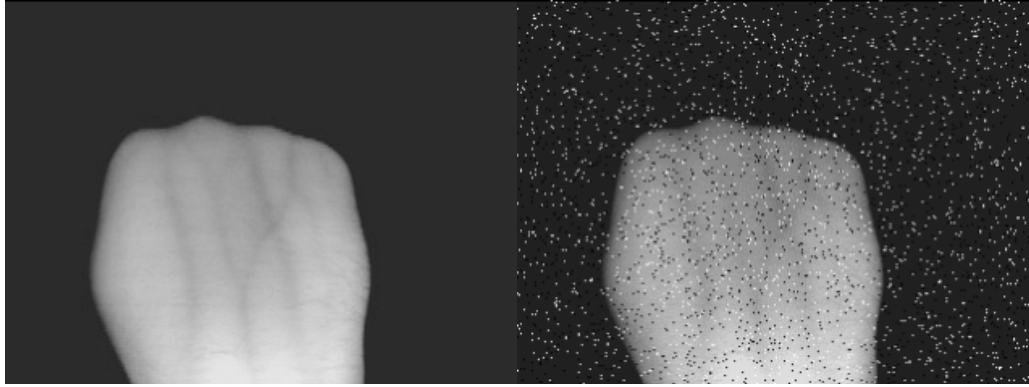


Figure 3.10 Real and Fake DHV

In deep forging, the discriminator is a binary classifier for which the inputs are classified as either real or fake. This distinguishing architecture was designed on the basis of a convolutional NN. Specifying the input of the section to be distinguished does not need to be performed first, and the inputs are considered as $N_x \times N_y \times N_c$. Here, $N_x \times N_y$ represents the image size. This was chosen as 256×256 , which comprises the light intensity channel of the used images. In the grayscale images, this is equal to 1. The outputs that are used to distinguish a scalar number is 0 and 1, thus it is binary and it determines if the image is the real one or not. In most instances, a sigmoid activity function, like that shown in Eq. (3.17) is utilized for the differentiating output. When using a sigmoid function, it is possible to normalize the output to between 0 and 1:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3.17)$$

The settings below can be utilized in the method proposed herein for the discriminator:

- ☆ The 0 layer, i.e. the input layer, has input images that are 256×256 pixels.
- ☆ HL-1: Utilizes masks that are 5×5 pixels as well as 32 feature maps.
- ☆ HL-2: Utilizes masks that are 5×5 pixels as well as 64 feature maps.

☆ HL-3: Makes use of the full connection that is between layer 2 and layer 3.

☆ HL-4: Makes use of the full connection that is between layer 3 and layer 4.

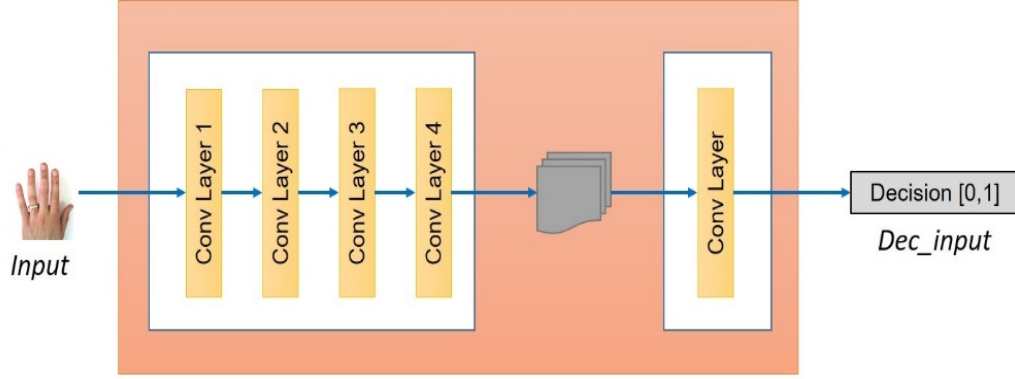


Figure 3.11 Discrimination layers

3.6 Generator Classifier

This uses a 3-layer deconvolutional NN. It has outputs which are samples that comprise $N_x \times N_y \times N_c$. The activity function $f(x) = \tanh(x)$ is also used by the output, which will change in the interval of $[-1, 1]$. In this part of the GAN network, three layers are active as follows:

- The 0 layer: Comprises the input layer that will receive the noise vector.
- HL-1: Connected to the 0 layer and learns using masks that are 5×5 pixels as well as a feature vector that has 32 modes.
- HL-2: Has masks that are 5×5 pixels as well as a feature vector that has 32 modes.
- The output layer has outputs that are equal to 256×256 pixels (the same as the original).

The DLGAN network is based on game theory. The generator and the discriminator perform a minimization/maximization game, respectively, as is given in Eq. (3.18) below [98]:

$$\min_G \max_D V(D, G) = E_{x: P_{data}(x)} [\log D(x)] + E_{z: P_z(z)} [\log(1 - D(G(Z)))] \quad (3.18)$$

Here, $p_{data}(x)$ represent the actual training examples, while $E_{z_{P_z}(z)}$ represent the fake ones that were created via the use of the noise process.

3.7 Evaluation Criteria

An analysis and evaluation of the authentication method that was proposed herein was performed using the indicators related to this field. This can be done using the criterion of an indicator for the average accuracy of identification, as is given in Eq. (3.19) [144]:

$$ACC_{Mean} = \frac{\sum_{i=1}^{Test} ACC_i}{Test} \quad (3.19)$$

Here, ACC_{Mean} , ACC_i , and Test, respectively, represent the average accuracy of authentication, the accuracy within a test, and the number of tests. In order to perform a calculation of the accuracy for an experiment that is formulated with ACC_i , the number of matching modes should be divided by the total number of test specimens, and the accuracy of an experiment should be calculated as shown below in Eq. (3.20) [144]:

$$ACC_i = \frac{\text{recognized images}}{\text{Total images}} \quad (3.20)$$

The standard deviation (STD) is also a significant indicator of the evaluation that has been conducted. A reduction in the STD is an indicator of increased stability in the authentication algorithm. The STD criterion is shown in Eq. (3.21) below:

$$ACC_{std} = \sqrt{\frac{\sum_{i=1}^{Test} (ACC_i - ACC_{mean})^2}{Test}} \quad (3.21)$$

The equal error rate (EER) is a biometric security system algorithm that is utilized to determine the threshold values beforehand for the false acceptance and rejection rates. If the rates are equal, then that common value is the ERR. This value is an

indication that the number of false acceptances is the same as that of the false rejections.

The EER represents the ERR of the improved biometric graph matching (IBGM). The lower the EER , the higher the accuracy of the system will be.

$$EER = \frac{\sum_{i=1}^N EER_i}{N} \quad (3.22)$$

Here, EER_i is the ERR in the i th experiment and N is the total number of samples.

CHAPTER 4

Experimental Results and Discussion

In this chapter, the test results implemented for the implementing system, which were mentioned in the previous chapter, are presented to determine the best recognition results that can be achieved for identifying individuals using feature vectors that were gained from the feature extraction stage.

MATLAB version 2020a was used in the tests and the training in the Windows 10 operating system environment, using the following laptop:

Device name: MSI. Processor Intel (R) Core (TM) i7-7700HQ CPU @ 2.80GHz 2.80 GHz.

Installed RAM 16.0 GB (15.9 GB usable). Product ID 00342-41367-87450-AAOEM. System type 64-bit operating system, x64-based processor.

4.1 Experiments Results

In the first stage, the generating model uses the random noise signals as the input and then it will produce a distorted image as the output. Step by step, aided by the discriminator, it will create some images of the class that is desired, which appear to be real. Next, a distinctive network, which will acts as a kind of competitor to the generator, uses these generated images of the generator, in addition to a certain class of real images, as the input and it will then evaluate to what degree the generated images will be evaluated with the real image. Then, after completing several stages of the network operation, a point is reached in which the discriminator is not able to recognize if the images that have been produced are in fact real or not. At this point, the generator produces some images of a certain class, which is actually the class that the discriminator was trained with, which did not previously exist. Hence, DL is similar to giving providing a thief with a picture of a hand. In the 2nd step, this image is checked against the image of the hand, and if similarity cannot be confirmed, then the image is given back to the forger, which will apply some noise and also slightly change it, so as to provide an increased level of adaptation with that

image. This process, which comprises moving the hand image goes on to the point that the fake image will be a slight bit different than the real image. In the method that is proposed, the training is conducted at 2 different levels. The DHVs contain noise and they are matched against the original image in order to select an image that will produce a lower number of matching errors. Moreover, A GAN is used to perform the DL for the biometric training and adaptation. GANs have very successful been used for DL and the processing of images. Furthermore, it is possible to achieve high accuracy in these networks via increasing of the size of the training data that are used. A limitation of many of the DL methods lies in the fact that they do require a considerable amount training data to be able to achieve high accuracy. Moreover, in most cases, insufficient training samples are available. As an example, there may be limited data that pertains to the effect of manual veins, which will therefore result in a decreased chance that the biometric accuracy will increase with a decrease in the number of training samples.

4.2 Experiment Analysis

To perform the evaluation and implementation the method that was proposed herein, it was implemented using 2 datasets. The results were first analyzed on the JU-DHVD [126], and the results on the the 11K-HD are given in the next section [138]. It was possible to diagnose it using similar methods, such as LBP [133], LPQ [135], GABOR [135], SIFT [136], flamelt-generated manifolds (FGM) [145] and BGM [145]. Three different thresholds exist in in the evaluation of Otsu thresholds. Pre-processed of the images is performed prior to entering the GAN, and then a comparison of the results of the output is performed against the desired methods.

4.2.1 Jilin University-Dorsal Hand Vein Database Results

This database contains 150 images for 50 different persons. Note that each of these subjects has 5 images of left and right dorsal veins. Table 4.1 shows the number of images used for training and testing the network.

Table 4. 1 Jilin University-Dorsal Hand Vein Database description

Dataset	Number of subjects	Training	Testing	Total training images	Total testing images
Jilin University-Dorsal	150	3 images of	2 images of	100	50
Hand Vein Database		each person	each person		

Once the image acquisition, features to be used, and matching process have been completed, it is time to define and characterize the performance of the proposed system and the matching methods. The results of each method will be studied from a statistical point of view.

Training was performed using 100 samples of DHVs from 50 different subjects (3 samples per individual). A total of 50 samples were used for testing and the others were reserved for mutual verification.

All of the samples were taken through an automated system that required the individuals to remove their hands from the device between each sample. With the consequent changes in the position, rotation, level of curvature of the back of the hand, and the consequent change in the dorsal vein pattern, common problems were found under normal use.

The ROC curve, as can be seen in Figure 4.1, presents the validation ACC (98.3%) via the use of the DLGAN, and in contrast, the loss error (1.67) was calculated with different threshold values for the dorsal as well as the palmar images for the 50 people. The individuals and images were randomly selected each time.

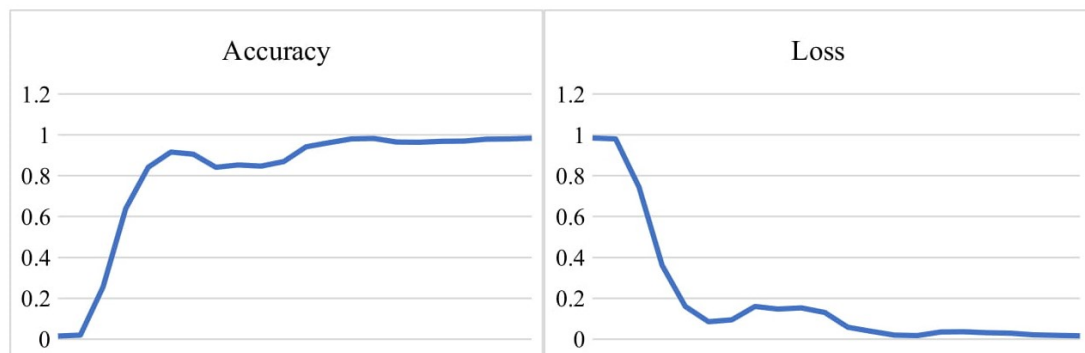


Figure 4.1 Accuracy and Loss of the JU-DHVD

In Table 4.2, the proposed method is compared with the LBP, LPQ, GABOR, SIFT, FGM, and BGM methods in three indicators of accuracy (ACC), the standard deviation of accuracy (STD), and equal error rate (EER) in identifying individuals:

Table 4.2. Comparison of the ACC, STD, and EER.

METHOD	ACC	STD	EER
LBP	91.97	0.48	8.28
LPQ	92.23	0.43	8.17
GABOR	93.38	0.39	7.18
SIFT	96.69	0.21	3.92
FGM	94.03	0.30	5.56
BGM	95.47	0.29	9.89
DL-GAN	98.36	0.192	2.47

The method that was proposed herein was observed to have greater efficiency when compared to the DLGAN, the LBP, the LPQ, the GABOR, the SIFT, the FGM, and the BGM for the 3 indicators of the ACC, the STD, and the mean error in the identifications of the individuals due to the fact that it exhibited greater accuracy. The method that was proposed herein had less errors well as STD in the experiments that were performed when a comparison was made with the other methods.

Figure 4.2. depicts the comparison that was performed of the accuracy index (AI) of the method that was proposed herein with other methods.

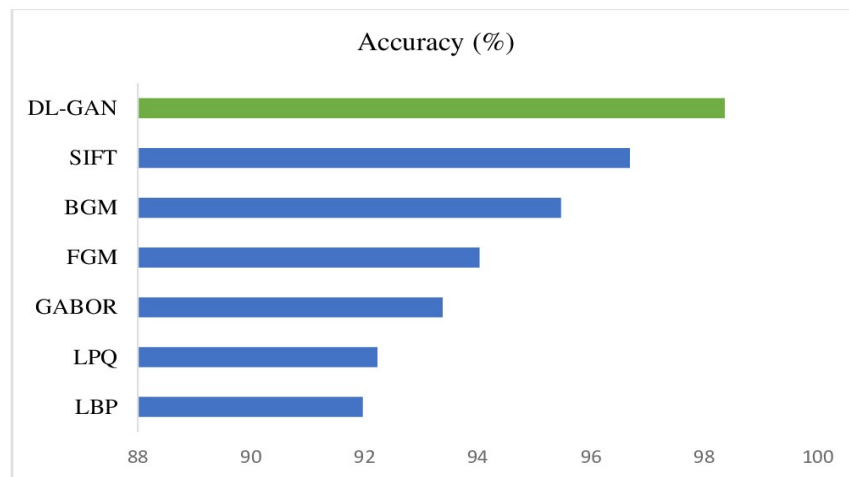


Figure 4.2 Accuracy comparison of the method that was proposed herein against the other learning methods. The change rates comprised 6.39 max% and 1.67 min%

The results proved that the method that was proposed herein had good accuracy, with 91.97% for the LBP, 92.33% for the LPQ, 93.38% for the GABOR, 96.69% for the SIFT, 94.03% for the FGM, 95.47% for the BGM, and 98.36% for the DLGAN. As can be seen, the method that was proposed herein ranked 1st for the AI, while the BGM method ranked 2nd. The method that was proposed herein achieved $\sim 1.67\%$ greater accuracy than the SIFT. The lowest performance was seen by the LBP method, which had accuracy of $\sim 91.97\%$.

Figure 4.3 shows a comparison of the method that was proposed herein with the other methods that are in the SD index of accuracy and error, respectively. The SD analysis results showed the SD for identification in the methods tested, which comprised 0.48 for the LBP, 0.43 for the LPQ, 0.39 for the GABOR, 0.21 for the SIFT, 0.3 for the FGM, 0.29 for the BGM, and 0.192 for the DLGAN. The method that was proposed herein exhibited the lowest SD of the methods tested, which showed that it had greater stability than the other methods.

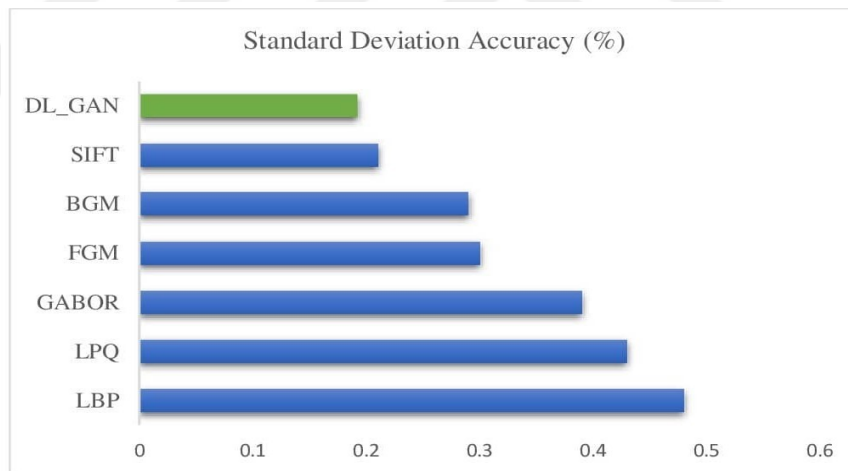


Figure 4.3 Comparison of the SD of the method that was proposed herein with the other learning methods that were tested. The change rates were 0.288 max% and 0.018 min%

Figure 4.4 presents the ERR of the methods tested, which comprised 8.28 for the LBP, 8.17 for the LPQ, 7.18 for the GABOR, 3.92 for the SIFT, 5.56 for the FGM, 9.89 for the BGM and 2.47 for the DLGAN. The best of the performances observed was determined for the method that was proposed herein, while the worst was found with the BGM. The experiments showed that the lowest possible error was observed

with the SIFT method when it was compared to the other methods that were tested, after the method that was proposed herein.

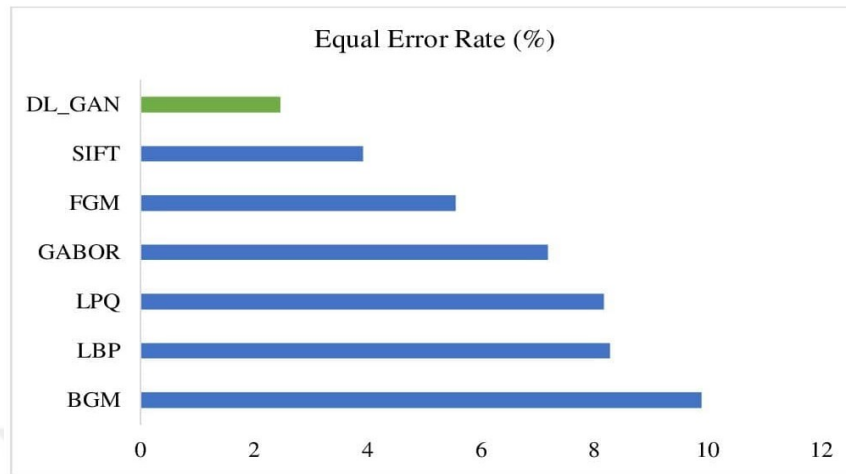


Figure 4.4 Comparison of the EER of the method that was proposed herein with the other learning methods that were tested. The change rates were 7.42 max% and 1.45 min%

4.2.2 The 11K Hands Dataset Results

This dataset contains 11,076 images of both sides of the hand, the left and right hands of 190 individuals of different ages and genders. Among them are 5396 images of the palm of the hand and the 5680 images of the back of the hand (dorsal), including 2892 images of the right side and 2788 of the left hand. The dorsal and palmate of each hand were photographed against a uniform white background and placed approximately the same distance from the camera. The dataset is free of charge for academic use. This work was used in both training and testing experiments on the dorsal veins side only, where 30 images were taken for everyone on average and were divided into 70% for training and 30% for tests, as shown in Table 4.3.

Table 4.3. Dataset 2 description.

Dataset	Number of subjects	Training	Testing	Total training images	Total testing images
11K data set	5680	21images of each person	9 images of each person	3976	1704

The experiments were performed using 5680 samples of DHVs from 190 different subjects (30 samples per individual). 3976 samples were used for training while another 1704 were reserved for mutual verification.

Figure 4.5 shows the accuracy obtained by DLGAN method (96.43%), and in contrast the loss error (3.57) was calculated with different threshold values for each of the dorsal and palmar images for 190 individuals. Individuals and images are randomly picked each time.

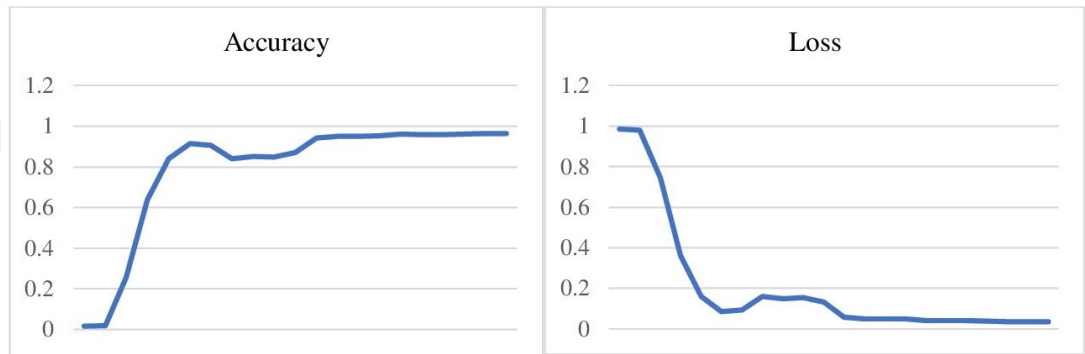


Figure 4.5 11 K database accuracy and loss

Table 4.5 shows a comparison of the method that was proposed herein with the LBP, the LPQ, the GABOR, the SIFT, the FGM, and the BGM for the 11K hand dataset with regard to the ACC, the STD, and the EER in the identification of the individuals.

Table 4.4. Comparison of the ACC, the STD, and the EER.

METHOD	ACC	STD	EER
LBP	92.86	0.41	8.42
LPQ	93.12	0.39	7.97
GABOR	94.21	0.37	7.23
SIFT	95.23	0.23	3.88
FGM	94.12	0.31	6.23
BGM	94.23	0.28	9.23
DL-GAN	96.43	0.21	3.55

The experiments in Table 4.4 show the accuracies of the methods that were tested, which comprised 92.86% for the LBP, 93.12% for the LPQ, 94.21% for the GABOR, 95.23% for the SIFT, 94.12% for the FGM, 94.23% for the BGM, and 96.43% for the DLGAN. The method that was proposed herein ranked 1st with regard to the AI, followed by the SIFT method, which ranked 2nd. The method that was proposed herein had $\sim 1.20\%$ greater accuracy than that of the BGM. The worst of the performances was observed for the LBP method, with an accuracy of $\sim 92.86\%$. A comparison was performed with regard to the AI between the method that was proposed herein and all other methods that were tested, which is given below in Figure 4.6.

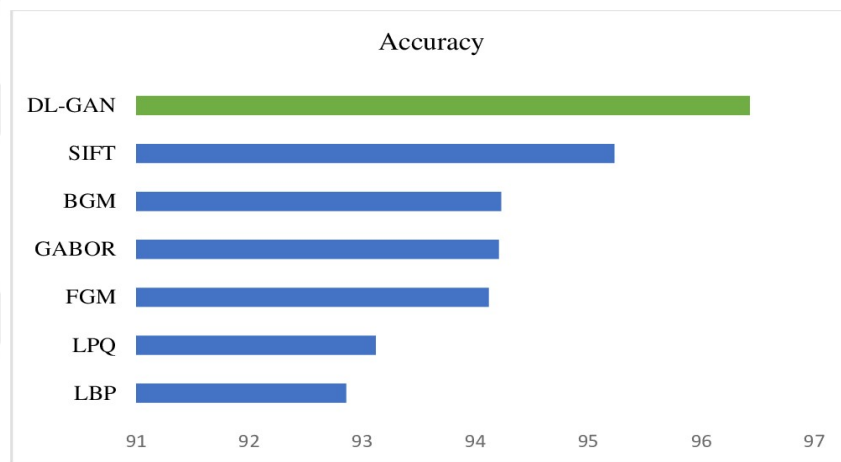


Figure 4.6 Comparison of the ACC of the method that was proposed herein with the other learning methods that were tested. The change rates were 3.37 max% and 1.20 min%

As can be seen in Figure 4.7 below, the method that was proposed herein was compared with all of the other methods with regard to the SD index of accuracy and error. The SD analysis showed the SD for identification for all of the methods that were tested, which comprised 0.41 for the LBP, 0.39 for the LPQ, 0.37 for the GABOR, 0.23 for the SIFT, 0.31 for the FGM, 0.28 for the BGM, and 0.21 for the DLGAN. The method that was proposed herein exhibited the lowest SD when compared to all of the other methods, showing that it had greater stability than the other methods that were tested.

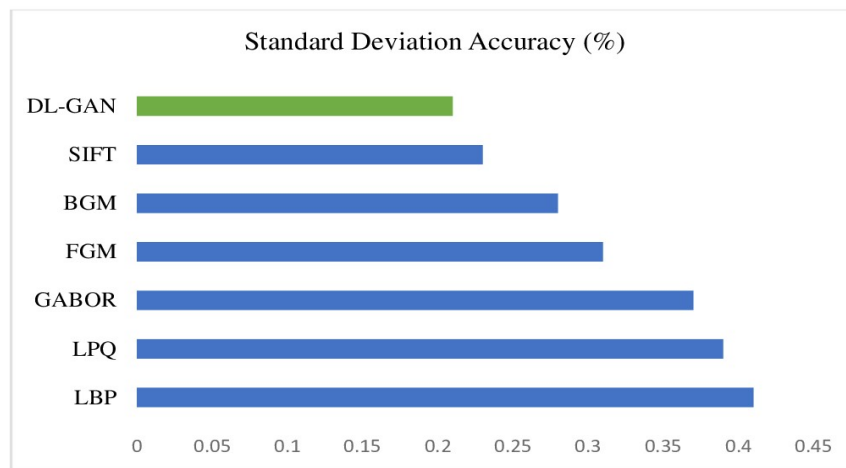


Figure 4.7 Comparison of the SD of the method that was proposed herein with the other learning methods that were tested. The change rates were 0.23 max% and 0.002 min%

The experiments in Table 4.8 show the EER of the methods that were tested, which comprised 8.42 for the LBP, 7.37 for the LPQ, 7.23 for the GABOR, 3.88 for the SIFT, 6.23 for the FGM, 9.23 for the BGM, and 3.55 for the DLGAN. The best of the performances was seen for the method that was proposed herein, while the worst was seen for the BGM method. The experiments showed that the SIFT method exhibited the lowest possible error when it was compared to all of the other methods, after the method that was proposed herein.

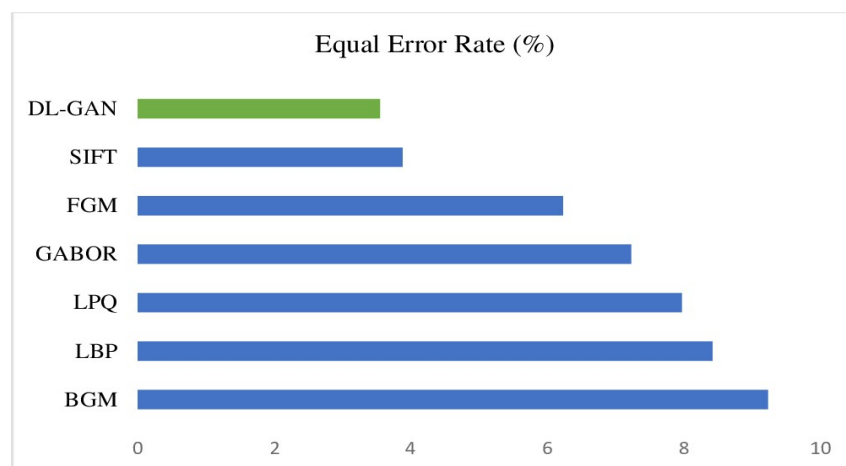


Figure 4.8 Comparison of the EER of the method that was proposed herein with the other learning methods that were tested. The change rates were 5.68 max% and 0.33 min%

4.3 Discussion

The visibility and the clarity blood vessels that are located under the skin varies with different users and is dependent on their physiological differences, the ambient temperature, the humidity, the position of the hand, the health of the individual, and the thickness of the subcutaneous layer of fat within the hand. Variations in the structure and the size of the parts of the veins, such as the collagen fibers, the lipid membrane, and nuclei, all have an affect on the scattering properties of the venous material and the alteration of the photon flux distribution. As well as the palm clarity, in the obtained images, the dorsal veins can also be affected by any hairs, moles, or scars that are on the skin surface, as well as the pigmentation of the skin. Hence, the approach that was investigated with regard to personal authentication was not only based on the venous network topology, it was also based on the pigmentation of the skin and the lipid deposits that can be found on the back of the hand.

In the obtained images, poor vision of the veins leads to extraction of the venous structure that will be inaccurate, which, in turn, can result in false points or some that are missing, which degrade system performance. The average number of fine details that could be extracted from images of low quality, that is to say, those that have poor vascular clarity, is limited.

It should be noted that the results of the experiments on the accuracy in identifying people using the DLGAN method obtained the highest rate when compared to some of the methods used in this field, such as the LBP, the LPQ, the GABOR, the SIFT, the FGM, and the BGM, and this indicated the success of testing the DLGAN method to generate pictures by the generator and the classification between the real and fake by discriminator.

By comparing the method proposed in this thesis with the other methods when testing the STD, the lowest percentage between them was obtained. Moreover, when evaluating the calculation of the ERR, the best possible result was compared among methods used, as shown in Table 4.5.

Table 4.5 Experimental results for the DLGAN based on the databases that were used.

Data Base	11 K Database			Jilin uni. Database		
Method	ACC	STD	EER	ACC	STD	EER
LBP	92.86	0.41	8.42	91.97	0.48	8.28
LPQ	93.12	0.39	7.97	92.33	0.43	8.17
GABOR	94.21	0.37	7.23	93.38	0.39	7.18
SIFT	95.23	0.23	3.88	96.69	0.21	3.92
FGM	94.12	0.31	6.23	94.04	0.30	5.56
BGM	94.23	0.28	9.23	95.47	0.29	9.89
DL-GAN	96.43	0.21	3.55	98.36	0.19	2.47

CHAPTER 5

Conclusion and Future Work

5.1 Conclusion

In the research that was presented herein, the investigation of a DHV system was performed to be able to verify biometric identity. DL and the GAN were used together toward this aim. The main advantage of using DL was that improvements could be made with regard to biometric identity verification of human hand veins. The DLGAN that was developed herein was tested on 2 databases. The obtained experimental results from the JU-DHVD experimental results showed the accuracy of the method that was proposed herein when compared to the other methods. The accuracies were 91.97% for the LBP, 92.33% for the LPQ, 93.38% for the GABOR, 94.03% for the SIFT, 96.69% for the FGM, 95.47% for the BGM, and 98.36% for the DLGAN. The rate of changes achieved were 6.39 max% and 1.67 min%. The experimental results using the 11K-HD showed the accuracies of the methods that were tested, which comprised for 92.86% the LBP, 93.12% for the LPQ, 94.21% for the GABOR, 95.23% for the SIFT, 94.12% for the FGM, 94.23% for the BGM, and 96.43% for the DLGAN. The rate of changes achieved were 3.37 max% and 1.20 min%. When the method that was proposed herein was compared with the LBP, the LPQ, the GABOR, the SIFT, the FGM, and the BGM for the same benchmark, but by using various other experimental scenarios, it was observed that the DLGAN was able to achieve recognition rates (RRs) with a discriminator size of 3, which gave a RR of 0.9000 and an epsilon of 1.0000e-08. As can be seen by these parameters, the method that was proposed herein outperformed most of the other methods that were tested. The main reason for this was that there was limited specific data that could be used for fine-tuning. Additionally, when compared with the traditional feature-based methods, DGAN feature-based methods do not require as much pre-processing and are more appropriate for applications in a real-world setting. The advantage that the DLGAN provides is that a backhand structure can be used for authentication, thus increasing its accuracy. The main challenge of the DLGAN is that the method is highly complex. However, an important advantage of the DLGAN lies in the fact that

some of the dorsal points could be lost due as the result of damage to the hand, changes in the pigmentation, or the presence of tattooing, which would all affect the ability to identify an individual. An advantage of this method is that it can also create a large volume of fake samples to be used in the training, which increases its accuracy, which, in turn, increases the power of the DLGAN with regard to image classification. A disadvantage of the DLGAN lies in the fact that it does necessitate the use of a large dataset, and the preparation of this dataset is somewhat difficult. In the case of a distorted hand, identification would be very difficult. If the person were to lose their hand at that time, recognizing the person would be impossible. Hence, the limitations of the DLGAN comprise the fact that it necessitates the use of a considerable amount of training data for it to be able to achieve the highest accuracy.

5.2. Future Work

Further work with regard to the research herein should consider the following aspects:

- 1) An improved capture device could be constructed, an improved handle box could be constructed to make its use more comfortable, and reduce the geometrical distortion within the vein images that are obtained. Factors that influence the illumination, including the intensity of the light, the distance, and the angle, could be subjected to further experiments in order to increase the contrast of the images, thus making the hand dorsal veins more visible.
- 2) Even though a database that comprised 2040 images of the hands of 102 participants was created, it must be further expanded, via the inclusion of more hand vein images, so as to facilitate a performance evaluation that will be more robust and accurate, and a variety of ethnic groups and varying age groups should be added to enable an analysis of the influence that race and age have on the recognition. Also included should be hand vein images that are acquired prior to and following the consumption of alcohol, so as to study the influence that alcohol has on the recognition.

- 3) In order to achieve higher RR for a larger database, feature extraction and pattern recognition methods that work more efficiently should be utilized. The potential of multifeature fusion features was observed herein. Some of the more complicated classifiers, including those like random tree and random forest may be more suitable. Moreover, hand dorsal vein patterns were shown to be an effective and efficient biometric feature to be used in personal identification herein. With just a bit more research, the model should be able to achieve an RR of 100%.



REFERENCES

- [1] F. Liu, G. Liu, Q. Zhao, and L. Shen, "Robust and high-security fingerprint recognition system using optical coherence tomography," *Neurocomputing*, vol. 402,14-28, 2020.
- [2] M. Trokielewicz, A. Czajka, and P. Maciejewicz, "Post-mortem iris recognition with deep-learning-based image segmentation," *Image and Vision Computing*, vol. 94, 103866, 2020.
- [3] L. Fei, B. Zhang, S. Teng, Z. Guo, S. Li, and W. Jia, "Joint Multiview Feature Learning for Hand-Print Recognition," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, 9743-9755, 2020.
- [4] M. He, J. Zhang, S. Shan, M. Kan, and X. Chen, "Deformable face net for pose invariant face recognition," *Pattern Recognition*, vol. 100,107113, 2020.
- [5] Q. Abbas and M. E. Ibrahim, "DenseHyper: an automatic recognition system for detection of hypertensive retinopathy using dense features transform and deep-residual learning," *Multimedia Tools and Applications*, vol. 79, 31595-31623, 2020.
- [6] S. Veluchamy and L. Karlmarx, "HE-Co-HOG and k-SVM classifier for finger knuckle and palm print-based multimodal biometric recognition," *Sensor Review*, 2020.
- [7] M. O. Oloyede, G. P. Hancke, and H. C. Myburgh, "A review on face recognition systems: recent approaches and challenges," *Multimedia Tools and Applications*, vol. 79, 27891-27922, 2020.
- [8] A. R. Patil, A. D. Rahulkar, and C. N. Modi, "Designing an Efficient Fingerprint Recognition System for Infants and Toddlers," in *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 2019,1-7.

- [9] W. Yang, S. Wang, J. Hu, G. Zheng, and C. Valli, "Security and accuracy of fingerprint-based biometrics: A review," *Symmetry*, vol. 11,141, 2019.
- [10] C. Rathgeb, J. Wagner, and C. Busch, "SIFT-based iris recognition revisited: prerequisites, advantages and improvements," *Pattern Analysis and Applications*, vol. 22, 889-906, 2019.
- [11] N. Alay and H. H. Al-Baity, "Deep Learning Approach for Multimodal Biometric Recognition System Based on Fusion of Iris, Face, and Finger Vein Traits," *Sensors*, vol. 20, 5523, 2020.
- [12] Y. Pititheeraphab, N. Thongpance, H. Aoyama, and C. Pintavirooj, "Vein pattern verification and identification based on local geometric invariants constructed from minutia points and augmented with barcoded local feature," *Applied Sciences*, vol. 10, 3192, 2020.
- [13] A. Herbadji, Z. Akhtar, K. Siddique, N. Guermat, L. Ziet, M. Cheniti, et al., "Combining Multiple Biometric Traits Using Asymmetric Aggregation Operators for Improved Person Recognition," *Symmetry*, vol. 12, 444, 2020.
- [14] B. Hartung, D. Rauschnig, H. Schwender, and S. Ritz-Timme, "A simple approach to use hand vein patterns as a tool for identification," *Forensic science international*, vol. 307, 110115, 2020.
- [15] W. Nie, B. Zhang, and S. Zhao, "A Novel Hyperspectral Based Dorsal Hand Recognition System," in *Proceedings of the 2019 11th International Conference on Machine Learning and Computing*, 2019, 362-367.
- [16] D. P. Wagh, H. Fadewar, and G. Shinde, "Biometric Finger Vein Recognition Methods for Authentication," in *Computing in Engineering and Technology*, ed: Springer, 2020, 45-53.
- [17] K. Přihodová and M. Hub, "Biometric Privacy through Hand Geometry-A Survey," in *2019 International Conference on Information and Digital Technologies (IDT)*, 2019, 395-401.

- [18] A. H. H. Alasadi and M. H. Dawood, "Dorsal hand-vein images recognition system based on grey level co-occurrence matrix and Tamura features," *International Journal of Applied Pattern Recognition*, vol. 4, 207-225, 2017.
- [19] B. Sontakke, V. Humbe, and P. Yannawar, "Dorsal hand vein authentication system: a review," *International Journal of Scientific Research Engineering & Technology*, vol. 6, 511-514, 2017.
- [20] N. Charaya and P. Singh, "Human Authentication Based On Dorsal Hand Veins: A," *International Journal of Pure and Applied Mathematics*, vol. 119, 2175-2185, 2018.
- [21] A. M. Badawi, "Hand Vein Biometric Verification Prototype: A Testing Performance and Patterns Similarity," *IPCV*, vol. 14, 3-9, 2006.
- [22] Y. Matsuda, N. Miura, A. Nagasaka, H. Kiyomizu, and T. Miyatake, "Finger-vein authentication based on deformation-tolerant feature-point matching," *Machine Vision and Applications*, vol. 27, 237-250, 2016.
- [23] S. Sharavanan, "Application of Image Enhancement and Segmentation Techniques on Vein Pattern for Better Identification," *Vinayaka Missions University*, 2016.
- [24] R. Sudhakar, S. Bharathi, and V. Gurunathan, "Analysis of Hand Vein Images Using Hybrid Techniques," *Hybrid Intelligence for Image Analysis and Understanding*, 315-338, 2017.
- [25] Y. Wang, W. Xie, X. Yu, and L.-K. Shark, "An automatic physical access control system based on hand vein biometric identification," *IEEE Transactions on Consumer Electronics*, vol. 61, 320-327, 2015.
- [26] S. H. Benziane and A. Benyettou, "Dorsal hand vein identification," *International Journal of Computer Science and Information Security*, vol. 14, 423, 2016.

- [27] B. L. Tait, "Aspects of Biometric Security in Internet of Things Devices," in *Digital Forensic Investigation of Internet of Things (IoT) Devices*, ed: Springer, 2021, 169-186.
- [28] M. R. A. Matos, "Identifikace uživatele na základě biometrických parametrů," 2016.
- [29] D. Bhattacharyya, R. Ranjan, F. Alisherov, and M. Choi, "Biometric authentication: A review," *International Journal of u-and e-Service, Science and Technology*, vol. 2, 13-28, 2009.
- [30] B. Ammour, L. Boubchir, T. Bouden, and M. Ramdani, "Face-iris multimodal biometric identification system," *Electronics*, vol. 9, 85, 2020.
- [31] J. Wayman, A. Jain, D. Maltoni, and D. Maio, "An introduction to biometric authentication systems," in *Biometric Systems*, ed: Springer, 2005, 1-20.
- [32] K. Sauerwein, T. B. Saul, D. W. Steadman, and C. B. Boehnen, "The effect of decomposition on the efficacy of biometrics for positive identification," *Journal of forensic sciences*, vol. 62, 1599-1602, 2017.
- [33] J. Priesnitz, C. Rathgeb, N. Buchmann, C. Busch, and M. Margraf, "An overview of touchless 2D fingerprint recognition," *EURASIP Journal on Image and Video Processing*, vol. 2021, 1-28, 2021.
- [34] C. Wilson, *Vein pattern recognition: a privacy-enhancing biometric*: CRC press, 2010.
- [35] D. D. Zhang, *Palmprint authentication*: Springer Science & Business Media, 2004.
- [36] A. W.-K. Kong, D. Zhang, and G. Lu, "A study of identical twins' palmprints for personal verification," *Pattern Recognition*, vol. 39, 2149-2156, 2006.
- [37] W. K. Kong and D. Zhang, "Palmprint texture analysis based on low-resolution images for personal authentication," in *Object recognition supported by user interaction for service robots*, 2002, 807-810.

- [38] X. Zhou, K. Zhou, and L. Shen, "Rotation and translation invariant palmprint recognition with biologically inspired transform," *IEEE Access*, vol. 8, 80097-80119, 2020.
- [39] C. L. Chowdhary, *Intelligent Systems: Advances in Biometric Systems, Soft Computing, Image Processing, and Data Analytics*: CRC Press, 2019.
- [40] I. A. Abdullah and J. J. Stephan, "A Survey of Face Recognition Systems," *Ibn AL-Haitham Journal For Pure and Applied Science*, vol. 34,144-160, 2021.
- [41] J. M. Chaves-González, M. A. Vega-Rodríguez, J. A. Gómez-Pulido, and J. M. Sánchez-Pérez, "Detecting skin in face recognition systems: A colour spaces study," *Digital Signal Processing*, vol. 20, 806-823, 2010.
- [42] S. A. Alazawi, R. A. Hasan, and A. A. Abdulhameed, "Hand Recognition System Based on Invariant Moments Features," in *Research in Intelligent and Computing in Engineering*, ed: Springer, 2021, 429-440.
- [43] P. Chaudhari, H. Agrawal, and K. Kotecha, "Data augmentation using MG-GAN for improved cancer classification on gene expression data," *Soft Computing*, 1-11, 2019.
- [44] A. Nadort, "The hand vein pattern used as a biometric feature," *Master Literature Thesis of Medical Natural Sciences at the Free University, Amsterdam*, 2007.
- [45] P. Carmeliet and R. K. Jain, "Angiogenesis in cancer and other diseases," *nature*, vol. 407, 249-257, 2000.
- [46] J.-C. Lee, "A novel biometric system based on palm vein image," *Pattern Recognition Letters*, vol. 33, 1520-1528, 2012.
- [47] J.-W. Wang and T.-H. Chen, "Building palm vein capturing system for extraction," in *2011 21st International Conference on Systems Engineering*, 2011, 311-314.

- [48] S. Barra, M. De Marsico, M. Nappi, F. Narducci, and D. Riccio, "A hand-based biometric system in visible light for mobile environments," *Information Sciences*, vol. 479, 472-485, 2019.
- [49] L. Fei, B. Zhang, Y. Xu, Z. Guo, J. Wen, and W. Jia, "Learning discriminant direction binary palmprint descriptor," *IEEE Transactions on Image Processing*, vol. 28, 3808-3820, 2019.
- [50] D. Zhong, X. Du, and K. Zhong, "Decade progress of palmprint recognition: A brief survey," *Neurocomputing*, vol. 328, 16-28, 2019.
- [51] W. Nie, B. Zhang, and S. Zhao, "Discriminative local feature for hyperspectral hand biometrics by adjusting image acutance," *Applied Sciences*, vol. 9, 4178, 2019.
- [52] D. Huang, R. Zhang, Y. Yin, Y. Wang, and Y. Wang, "Local feature approach to dorsal hand vein recognition by centroid-based circular key-point grid and fine-grained matching," *Image and Vision Computing*, vol. 58, 266-277, 2017.
- [53] W. Wu, S. J. Elliott, S. Lin, and W. Yuan, "Low-cost biometric recognition system based on NIR palm vein image," *IET Biometrics*, vol. 8, 206-214, 2019.
- [54] S. Ma, Z. Tao, X. Yang, Y. Yu, X. Zhou, and Z. Li, "Bathymetry retrieval from hyperspectral remote sensing data in optical-shallow water," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, 1205-1212, 2013.
- [55] M. I. Lakhal, H. Çevikalp, S. Escalera, and F. Ofli, "Recurrent neural networks for remote sensing image classification," *IET Computer Vision*, vol. 12, 1040-1045, 2018.
- [56] M. A. Ferrer, A. Morales, and A. Díaz, "An approach to SWIR hyperspectral hand biometrics," *Information Sciences*, vol. 268, 3-19, 2014.

- [57] L. Wang and G. Leedham, "Near-and far-infrared imaging for vein pattern biometrics," in 2006 IEEE International Conference on Video and Signal Based Surveillance, 2006, 52-52.
- [58] H.-s. Choi, "Apparatus and method for identifying individuals through their subcutaneous vein patterns and integrated system using said apparatus and method," ed: Google Patents, 2001.
- [59] J. Cross and C. Smith, "Thermographic imaging of the subcutaneous vascular network of the back of the hand for biometric identification," in Proceedings The Institute of Electrical and Electronics Engineers. 29th Annual 1995 International Carnahan Conference on Security Technology, 1995, 20-35.
- [60] H. Rubin, "Aging process: Part x-the skin, the skeleton and the brain," 2003.
- [61] O. A. Carretero, "Vascular remodeling and the kallikrein-kinin system," The Journal of clinical investigation, vol. 115,588-591, 2005.
- [62] A. Nadort, "The Hand Vein Pattern Used as a Biometric Feature: Literature thesis for Master of Science programmed Physics of Life," Nadort,— Amsterdam: Medical Natural Sciences at the Free University, 2007.
- [63] B. M. Carlson, Human Embryology and Developmental Biology E-Book: With STUDENT CONSULT Online Access: Elsevier Health Sciences, 2018.
- [64] H. G. Zheng, "Static and dynamic analysis of near infra-red dorsal hand vein images for biometric applications," University of Central Lancashire, 2017.
- [65] E. Papakonstantinou, M. Roth, and G. Karakiulakis, "Hyaluronic acid: A key molecule in skin aging," Dermato-endocrinology, vol. 4, 253-258, 2012.
- [66] J. M. Waller and H. I. Maibach, "Age and skin structure and function, a quantitative approach (I): blood flow, pH, thickness, and ultrasound echogenicity," Skin research and technology, vol. 11, 221-235, 2005.
- [67] R. V. Yampolskiy and V. Govindaraju, "Behavioural biometrics: a survey and classification," International Journal of Biometrics, vol. 1,81-113, 2008.

- [68] N. Zheng, A. Paloski, and H. Wang, "An efficient user verification system via mouse movements," in Proceedings of the 18th ACM conference on Computer and communications security, 2011,139-150.
- [69] A. Eriksson and P. Wretling, "How flexible is the human voice?-a case study of mimicry," in Fifth European Conference on Speech Communication and Technology, 1997.
- [70] S. Furui, "Recent advances in speaker recognition," Pattern recognition letters, vol. 18, 859-872, 1997.
- [71] J. Ker, L. Wang, J. Rao, and T. Lim, "Deep learning applications in medical image analysis," Ieee Access, vol. 6,9375-9389, 2017.
- [72] A. N. Azmi, D. Nasien, and F. S. Omar, "Biometric signature verification system based on freeman chain code and k-nearest neighbor," Multimedia Tools and Applications, vol. 76, 15341-15355, 2017.
- [73] D. Gafurov, "A survey of biometric gait recognition: Approaches, security and challenges," in Annual Norwegian computer science conference, 2007, 19-21.
- [74] S. Minaee, A. Abdolrashidi, H. Su, M. Bennamoun, and D. Zhang, "Biometric recognition using deep learning: A survey," arXiv preprint arXiv:1912.00271, 2019.
- [75] S. Bhowmick, S. Nagarajaiah, and A. Veeraraghavan, "Vision and Deep Learning-Based Algorithms to Detect and Quantify Cracks on Concrete Surfaces from UAV Videos," Sensors, vol. 20, 6299, 2020.
- [76] B. Bhanu and A. Kumar, Deep learning for biometrics: Springer, 2017.
- [77] L. Boubchir and B. Daachi, "Recent Advances in Biometrics and Its Applications," ed: Multidisciplinary Digital Publishing Institute, 2021.

- [78] Y. Li, T. Zhang, S. Sun, and X. Gao, "Accelerating flash calculation through deep learning methods," *Journal of Computational Physics*, vol. 394,153-165, 2019.
- [79] I. Serna, A. Peña, A. Morales, and J. Fierrez, "InsideBias: Measuring bias in deep networks and application to face gender biometrics," in *2020 25th International Conference on Pattern Recognition (ICPR)*, 2021, 3720-3727.
- [80] E. J. da Silva Luz, G. J. Moreira, L. S. Oliveira, W. R. Schwartz, and D. Menotti, "Learning deep off-the-person heart biometrics representations," *IEEE Transactions on Information Forensics and Security*, vol. 13,1258-1270, 2017.
- [81] E. Maiorana, "Deep learning for EEG-based biometric recognition," *Neurocomputing*, vol. 410, 374-386, 2020.
- [82] A. B. López, "Deep Learning in Biometrics: A Survey," *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal*, vol. 8,19-32, 2019.
- [83] R. Vargas, A. Mosavi, and R. Ruiz, "Deep learning: a review," *Advances in intelligent systems and computing*, 2017.
- [84] Y. Deng, "Deep learning on mobile devices: a review," in *Mobile Multimedia/Image Processing, Security, and Applications 2019*, 2019, 109930A.
- [85] A. B. J. Teoh and L. Leng, "Special Issue on Advanced Biometrics with Deep Learning," ed: Multidisciplinary Digital Publishing Institute, 2020.
- [86] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning. *nature* (2015)," May; 521 (7553): 436 10.1038/nature14539, 2015.
- [87] N. A. Al-johania and L. A. Elrefaei, "Dorsal hand vein recognition by convolutional neural networks: feature learning and transfer learning approaches," *Int. J. Intell. Eng. Syst*, vol. 12, 178-191, 2019.

- [88] Z. Meng and X. Gu, "Hand vein identification using local Gabor ordinal measure," *Journal of Electronic Imaging*, vol. 23, 053004, 2014.
- [89] N. C. Thompson, K. Greenewald, K. Lee, and G. F. Manso, "The computational limits of deep learning," *arXiv preprint arXiv:2007.05558*, 2020.
- [90] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, 436-444, 2015.
- [91] Y. Bengio, I. Goodfellow, and A. Courville, *Deep learning vol. 1: MIT press Massachusetts, USA*, 2017.
- [92] J. D. Kelleher, *Deep learning: MIT press*, 2019.
- [93] M. M. Saeed, Z. Al Aghbari, and M. Alsharidah, "Big data clustering techniques based on Spark: a literature review," *PeerJ Computer Science*, vol. 6, 321, 2020.
- [94] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, et al., "Human-level control through deep reinforcement learning," *nature*, vol. 518, 529-533, 2015.
- [95] M. Z. Alom, T. M. Taha, C. Yakopcic, S. Westberg, P. Sidike, M. S. Nasrin, et al., "A state-of-the-art survey on deep learning theory and architectures," *Electronics*, vol. 8, 292, 2019.
- [96] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "A brief survey of deep reinforcement learning," *arXiv preprint arXiv:1708.05866*, 2017.
- [97] K. Alashik, S. Hussin, R. YILDIRIM, and A. ALGUTTAR, "Dorsal Hand Vein Identification Based on Deep Convolutional Neural Networks and Visualizing Intermediate Layer Activations," *Avrupa Bilim ve Teknoloji Dergisi*, 512-521.

- [98] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, et al., "Generative adversarial networks," arXiv preprint arXiv:1406.2661, 2014.
- [99] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep learning for computer vision: A brief review," *Computational intelligence and neuroscience*, vol. 2018, 2018.
- [100] Z. Jiang, "Cai," A Novel Bayes Model: Hidden Naive Bayes," *Knowledge and Data Engineering*, " IEEE Transactions on, vol. 21, 1361-1371, 2009.
- [101] I. IN'TRODLJCHION, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," *Readings in speech recognition*, 267, 1990.
- [102] Y. Pu, Z. Gan, R. Henao, X. Yuan, C. Li, A. Stevens, et al., "Variational autoencoder for deep learning of images, labels and captions," arXiv preprint arXiv:1609.08976, 2016.
- [103] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*, 2009, 248-255.
- [104] M. Arjovsky and L. Bottou, "Towards principled methods for training generative adversarial networks," arXiv preprint arXiv:1701.04862, 2017.
- [105] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning. nature 521 (7553), 436-444," *Google Scholar Google Scholar Cross Ref Cross Ref*, 2015.
- [106] A. Antoniou, A. Storkey, and H. Edwards, "Data augmentation generative adversarial networks," arXiv preprint arXiv:1711.04340, 2017.
- [107] D. He, W. Chen, L. Wang, and T.-Y. Liu, "A game-theoretic machine learning approach for revenue maximization in sponsored search," arXiv preprint arXiv:1406.0728, 2014.

- [108] I. Goodfellow, "Nips 2016 tutorial: Generative adversarial networks," arXiv preprint arXiv:1701.00160, 2016.
- [109] D. Warde-Farley and B. Xu, "Generative adversarial nets," Advances in Neural Information Processing Systems (NIPS), 2014.
- [110] A. v. d. Oord, N. Kalchbrenner, O. Vinyals, L. Espeholt, A. Graves, and K. Kavukcuoglu, "Conditional image generation with pixelcnn decoders," arXiv preprint arXiv:1606.05328, 2016.
- [111] M. Mirza and S. Osindero, "Conditional generative adversarial nets," arXiv preprint arXiv:1411.1784, 2014.
- [112] A. Odena, C. Olah, and J. Shlens, "Conditional image synthesis with auxiliary classifier gans," in International conference on machine learning, 2017, 2642-2651.
- [113] S. Shao, D. Zhou, R. He, J. Li, S. Zou, K. Mallery, et al., "Risk assessment of airborne transmission of COVID-19 by asymptomatic individuals under different practical settings," Journal of aerosol science, vol. 151,105661, 2021.
- [114] K. Wang, Y. Zhang, Z. Yuan, and D. Zhuang, "Hand vein recognition based on multi supplemental features of multi-classifier fusion decision," in 2006 International Conference on Mechatronics and Automation, 2006, 1790-1795.
- [115] C.-L. Lin and K.-C. Fan, "Biometric verification using thermal images of palm-dorsa vein patterns," IEEE Transactions on Circuits and systems for Video Technology, vol. 14, 199-213, 2004.
- [116] X. Zhu and D. Huang, "Hand dorsal vein recognition based on hierarchically structured texture and geometry features," in Chinese Conference on Biometric Recognition, 2012, 157-164.
- [117] M. H.-M. Khan, R. K. Subramanian, and N. A. M. Khan, "Representation of hand dorsal vein features using a low dimensional representation integrating

cholesky decomposition," in 2009 2nd International Congress on Image and Signal Processing, 2009, 1-6.

- [118] M. H.-M. Khan and N. A. M. Khan, "Investigating linear discriminant analysis (lda) on dorsal hand vein images," in Third International Conference on Innovative Computing Technology (INTECH 2013), 2013, 54-59.
- [119] Y. Wang, K. Li, and J. Cui, "Hand-dorsa vein recognition based on partition local binary pattern," in IEEE 10th International conference on signal processing proceedings, 2010, 1671-1674.
- [120] D. Huang, Y. Tang, Y. Wang, L. Chen, and Y. Wang, "Hand vein recognition based on oriented gradient maps and local feature matching," in Asian conference on computer vision, 2012, 430-444.
- [121] D. Huang, Y. Tang, Y. Wang, L. Chen, and Y. Wang, "Hand-dorsa vein recognition by matching local features of multisource keypoints," IEEE transactions on cybernetics, vol. 45, 1823-1837, 2014.
- [122] A. Kumar and K. V. Prathyusha, "Personal authentication using hand vein triangulation and knuckle shape," IEEE Transactions on Image processing, vol. 18, 2127-2136, 2009.
- [123] D. Huang, X. Zhu, Y. Wang, and D. Zhang, "Dorsal hand vein recognition via hierarchical combination of texture and shape clues," Neurocomputing, vol. 214, 815-828, 2016.
- [124] Y. Aberni, L. Boubchir, and B. Daachi, "Palm vein recognition based on competitive coding scheme using multi-scale local binary pattern with ant colony optimization," Pattern Recognition Letters, vol. 136, 101-110, 2020.
- [125] O. Alpar, "A novel fuzzy curvature method for recognition of anterior forearm subcutaneous veins by thermal imaging," Expert Systems with Applications, vol. 120, 33-42, 2019.

- [126] F. Liu, S. Jiang, B. Kang, and T. Hou, "A recognition system for partially occluded dorsal hand vein using improved biometric graph matching," *IEEE Access*, vol. 8, 74525-74534, 2020.
- [127] A. Achban, H. A. Nugroho, and P. Nugroho, "Wrist Hand Vein Recognition Using Local Line Binary Pattern (LLBP)," in *2019 5th International Conference on Science and Technology (ICST)*, 2019, 1-6.
- [128] H. Qin, M. A. El Yacoubi, J. Lin, and B. Liu, "An iterative deep neural network for hand-vein verification," *IEEE Access*, vol. 7, 34823-34837, 2019.
- [129] B. Qin, Z. Chen, M. Sun, and Q. Sun, "Active Disturbance Rejection Control of Ship Course Based on Deep Belief Network," in *2019 IEEE 8th Data Driven Control and Learning Systems Conference (DDCLS)*, 2019, 29-36.
- [130] R. Garcia-Martin and R. Sanchez-Reillo, "Vein biometric recognition on a smartphone," *IEEE Access*, vol. 8, 104801-104813, 2020.
- [131] R. S. Kuzu, E. Piciucco, E. Maiorana, and P. Campisi, "On-the-fly finger-vein-based biometric recognition using deep neural networks," *IEEE Transactions on Information Forensics and Security*, vol. 15, 2641-2654, 2020.
- [132] X. Li, D. Huang, R. Zhang, Y. Wang, and X. Xie, "Hand dorsal vein recognition by matching width skeleton models," in *2016 IEEE International Conference on Image Processing (ICIP)*, 2016, 3146-3150.
- [133] K. Li, G. Zhang, Y. Wang, P. Wang, and C. Ni, "Hand-dorsa vein recognition based on improved partition local binary patterns," in *Chinese Conference on Biometric Recognition*, 2015, 312-320.
- [134] W. Kang and Q. Wu, "Contactless palm vein recognition using a mutual foreground-based local binary pattern," *IEEE transactions on Information Forensics and Security*, vol. 9, 1974-1985, 2014.

- [135] K. Premalatha and A. Natarajan, "Hand vein pattern recognition using natural image statistics," *Defence Science Journal*, vol. 65, 150, 2015.
- [136] Y. Wang and X. Zheng, "Cross-device hand vein recognition based on improved SIFT," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 16, 1840010, 2018.
- [137] A. Mousavi, A. B. Patel, and R. G. Baraniuk, "A deep learning approach to structured signal recovery," in *2015 53rd annual allerton conference on communication, control, and computing (Allerton)*, 2015, 1336-1343.
- [138] M. Afifi, "11K Hands: Gender recognition and biometric identification using a large dataset of hand images," *Multimedia Tools and Applications*, vol. 78, 20835-20854, 2019.
- [139] K. M. Alashik and R. Yildirim, "Human Identity Verification from Biometric Dorsal Hand Vein Images Using the DL-GAN Method," *IEEE Access*, 2021.
- [140] Z. Wang, Q. She, and T. E. Ward, "Generative adversarial networks in computer vision: A survey and taxonomy," *arXiv preprint arXiv:1906.01529*, 2019.
- [141] Z. Wang, Q. She, and T. E. Ward, "Generative adversarial networks in computer vision: A survey and taxonomy," *ACM Computing Surveys (CSUR)*, vol. 54, 1-38, 2021.
- [142] B. Küçükuğurlu and E. Gedikli, "Symbiotic Organisms Search Algorithm for multilevel thresholding of images," *Expert Systems with Applications*, vol. 147, 113210, 2020.
- [143] M. Frid-Adar, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "Synthetic data augmentation using GAN for improved liver lesion classification," in *2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)*, 2018, 289-293.

- [144] R. B. Trabelsi, A. D. Masmoudi, and D. S. Masmoudi, "Hand vein recognition system with circular difference and statistical directional patterns based on an artificial neural network," *Multimedia Tools and Applications*, vol. 75, 687-707, 2016.
- [145] S. M. Lajevardi, A. Arakala, S. Davis, and K. J. Horadam, "Hand vein authentication using biometric graph matching," *IET Biometrics*, vol. 3, 302-313, 2014.
- [146] H. Hewamalage, C. Bergmeir, and K. Bandara, "Recurrent neural networks for time series forecasting: Current status and future directions," *International Journal of Forecasting*, vol. 37, 388-427, 2021.

APPENDICES

Appendix A: Artificial neural networks with Math.

Appendix B: Deep learning.

Appendix C: Generative Adversarial Networks.



Appendix A – Artificial Neural Networks with Math.

In this appendix, we provide an overview of ANNs from a mathematical point of view.

A NN is comprised of neurons, which are also known as nodes, which are all connected in one way or another. Each of these neurons has a number, and each of these connections has a weight.

The neurons, or nodes, are divided between the input, the hidden layer, and the output layer. In practice, many layers exist and there is no number of layers that is the best.

The NNs that make up DL are part of ML.

Why DL?

1. DL is the preferred choice over shallow level learning when there is a huge amount of data that is either labeled or unlabeled.
2. DL can provide state-of-the-art level performance with tasks that involve images, text, or sound, in addition to a number of advances in computer vision, speech recognition, and NLP.
3. DL provides feature or abstract representation; thus less time needs to be spent on feature engineering.

Neuron, or node, is a computational unit that uses the input to perform some calculations and it then produces an output.

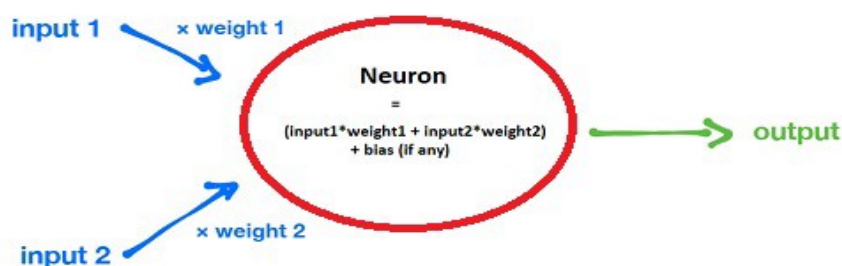


Figure A.1 NN architecture

Figure A.1 In NNs, there is an input, as well as some weights/parameters that are applied to the dot product of these 2 vectors, which produces the result, which would comprise a continuous value –infinity to infinity.

An activation function, which was used for restricting the output values, squashes them and a value in the range that is determined based on the type of the activation function that is used is thus produced.

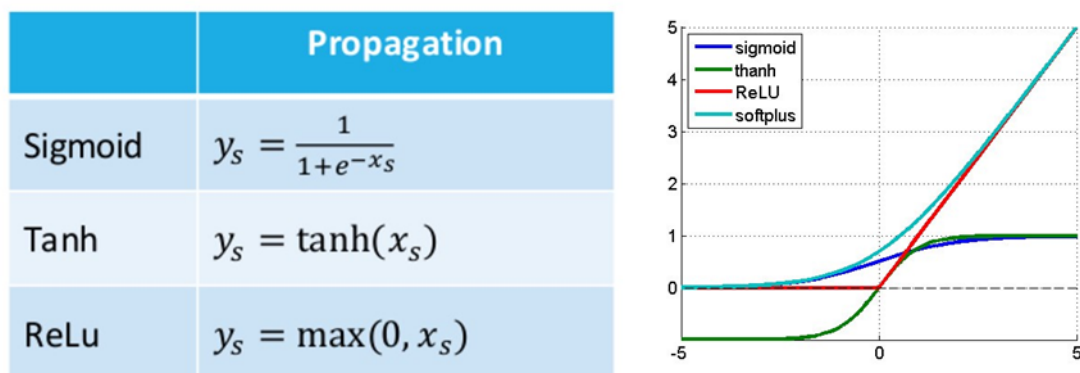


Figure A.2 Activation function

These 3, as shown in Figure A.2, are the activation functions that are used most often. They have a Sigmoid that ranges from 0 to 1, a Tanh that ranges from –1 to 1, and a Relu that ranges from 0 to +infinity.

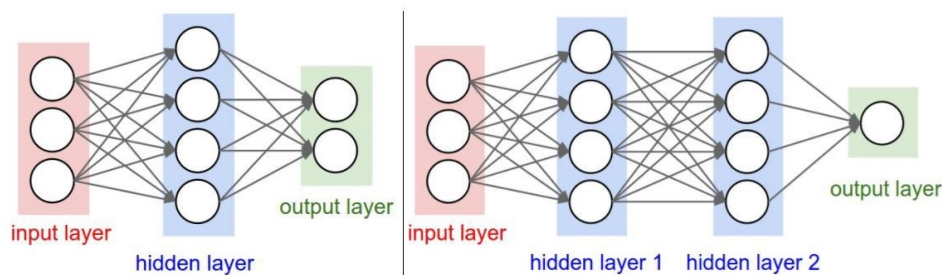


Figure A.3 The 2 layers of the NN on the left comprise 1 HL that has 4 neurons, as well as 1 output layer that has 2 neurons, and 3 outputs. The 3 layers of the NN on the right comprises 3 inputs, 2 HLS that have 4 neurons as well as 1 output layer

The output from one of the layers would then become the input for the next layer.

Here, there are 3 layers:

1. The input layer, which comprises set of input neurons, of which each of the neurons represents each of the features in the dataset. This layer takes in the input and passes it on to the next layer.
2. The HL comprises a set of (n) number of neurons, of which each of the neurons has a weight, that is, a parameter, that is assigned to it. This layer takes in the input from the layer before it and performs the dot product of the inputs and the weights, and it then applies the activation function, as was shown above, so as to produce the result and then pass those data on to the next layer.
3. The output layer, which is basically the same as the HL, except this will give the result, that is the outcome/class/value.

1. Forward propagation

Forward propagation is an easy process, in which the inputs are fed forward through each of the layers that are in the network, and then the outputs that are from the layer before it is used as the inputs for the next layer. This is all done, of course, after the data was fed in as the inputs.

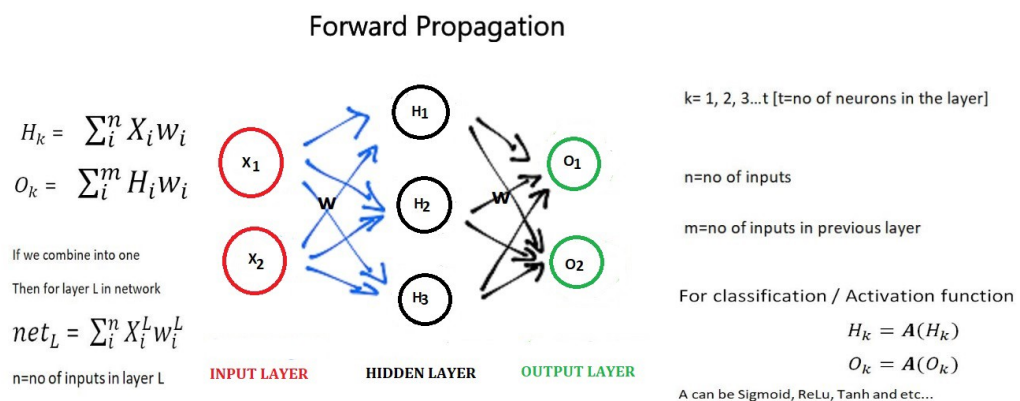


Figure A.4 Forward propagation

Moving forward through the network is what is known as a forward pass. To do this, a formula is iteratively used in the calculation of each of the neurons in the next layer. The neurons that are used for the activations are represented by a , while the weights are represented by w , and the biases are represented by b , which has been cumulated in the vectors.

$$a^{(l)} = \sigma(Wa^{l-1} + b) \quad (\text{A.1})$$

From here we move forward, until an output is obtained. How good this output, $\hat{y}$ is, can be measured using a cost function, C , as well as the desired result in the output layer, y , and this is done for each of the samples. This is most commonly known as the mean squared error (MSE):

$$C = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (\text{A.2})$$

When given the first result, it is necessary to go back and make an adjustment to the weights and biases, so the cost function, can be optimized, which is known as a backwards pass. Essentially, it is aimed to adjust the entire NN so the output value will be optimized. In actuality, this is how the algorithm is instructed as to whether performed well or not. Attempts keep being made to optimize the cost function by running new observations from the dataset.

To be able to update the network, the gradients are calculated, which provide small nudges/updates to the individual weights in that are in each of the layers.

$$\frac{\partial C}{\partial w^{(L)}} = \frac{\partial C}{\partial a^{(L)}} \frac{\partial a^{(L)}}{\partial z^{(L)}} \frac{\partial z^{(L)}}{\partial w^{(L)}} \quad (\text{A.3})$$

Put simply, each of the weights is gone through, such as in the output layer, and then the learning rate value is subtracted, and then we multiply that by the cost of a specific weight that is from the original value that the specific weight had.

$$w^{(L)} = w^{(L)} - \text{learningrate} \times \frac{\partial C}{\partial w^{(L)}} \quad (\text{A.4})$$

Something known as mini-batches are added, where the gradient of a certain number of defined observations for each mini-batch are averaged.

2. Back propagation

The main goal here is updating each weight in the network so they result in the predicted output being closer to the target output. This will then minimize the error for each of the output neurons as well as the network. Thus far, the total error that need to be minimized has been obtained.

Next, it is necessary to calculate the terms that are given below:

1. How much has the total error changed in terms of the result? or How much of a change was there in the results?
2. How much has the result changed in terms of the sum? or How much of a change was there in the sum?
3. How much has the sum changed in terms of the weights? or How much of a change was there in the weights?

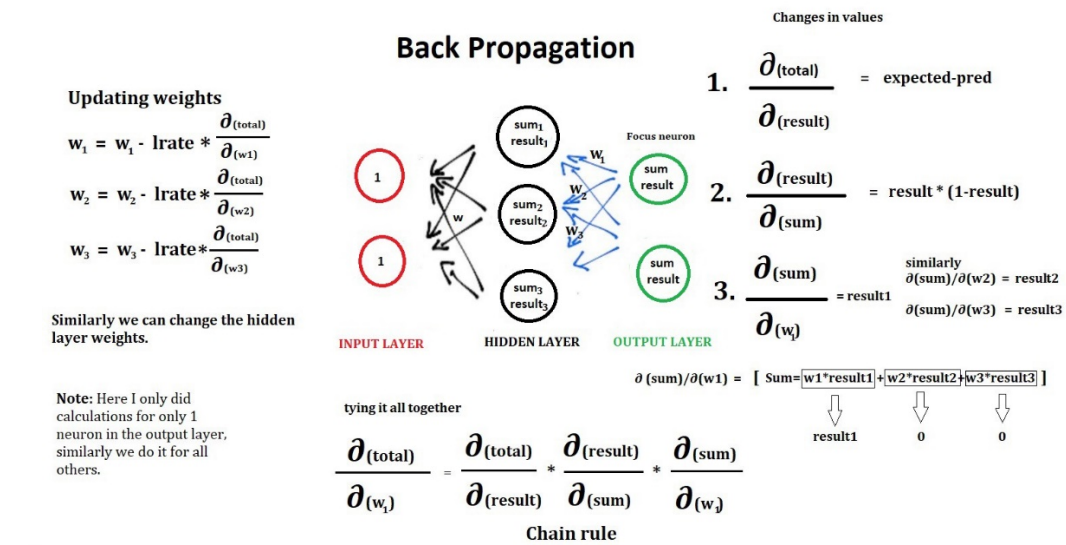


Figure A.5 Back propagation

First, the relevant equations must be defined. Note that the indexing that was explained previously is not used here, and each of the layers is abstracted to rather than each of the weights, biases, or activations:

$$z^{(L)} = w^{(L)} \times a + b \quad (\text{A.5})$$

$$a^{(L)} = \sigma(z^{(L)}) \quad (\text{A.6})$$

$$C = (a^{(L)} - y)^2 \quad (\text{A.7})$$

Here, L is the last layer, w is the weight, l is the layer, a is the activation, y is the output value, b is the bias, and c is the cost function.

Three equations for calculating the gradient.

It is necessary to move backwards within the network to be able to update the weights and biases. Let us discuss now how that is done with math. One equation is used for the for the weights, one equation is used for the for the for biases, and one equation is used for the for the for activations:

$$\frac{\partial C}{\partial w^{(L)}} = \frac{\partial C}{\partial a^{(L)}} \frac{\partial a^{(L)}}{\partial z^{(L)}} \frac{\partial z^{(L)}}{\partial w^{(L)}} \quad (\text{A.8})$$

$$\frac{\partial C}{\partial b^{(L)}} = \frac{\partial C}{\partial a^{(L)}} \frac{\partial a^{(L)}}{\partial z^{(L)}} \frac{\partial z^{(L)}}{\partial b^{(L)}} \quad (\text{A.9})$$

$$\frac{\partial C}{\partial a^{(L-1)}} = \frac{\partial C}{\partial a^{(L)}} \frac{\partial a^{(L)}}{\partial z^{(L)}} \frac{\partial z^{(L)}}{\partial a^{(L-1)}} \quad (\text{A.10})$$

These equations will only measure the ratio of how a specific weight will affect the cost function that is aimed to be optimized. Optimization is performed by moving toward the output of the equations. Each part of the derivative taken from the weights and from the biases will be saved in one of the gradient vectors, which have a number of dimensions that is equal to the number of weights and biases. The gradient is depicted below by the triangle (∇), and n represents the number of the weights and the biases:

$$-\nabla C(w_1, b_1, \dots, w_n, b_n) = \begin{bmatrix} \frac{\partial C}{\partial w_1} \\ \frac{\partial C}{\partial b_1} \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial C}{\partial w_n} \\ \frac{\partial C}{\partial b_n} \end{bmatrix} \quad (\text{A.11})$$

Appendix B –Deep Learning

DL is a branch of ML that is completely based on artificial neural networks, as shown in Figure B.1. As NNs will mimic the human brain, so DL is also a kind of mimic of the human brain. In DL, we do not need to explicitly program everything. The concept of DL is not new. It has been around for a couple of years now. There is hype about it nowadays because earlier there was not that much processing power or a lot of data. Over the last 20 years, the processing power has increased exponentially, and thus DL and ML have been the result.

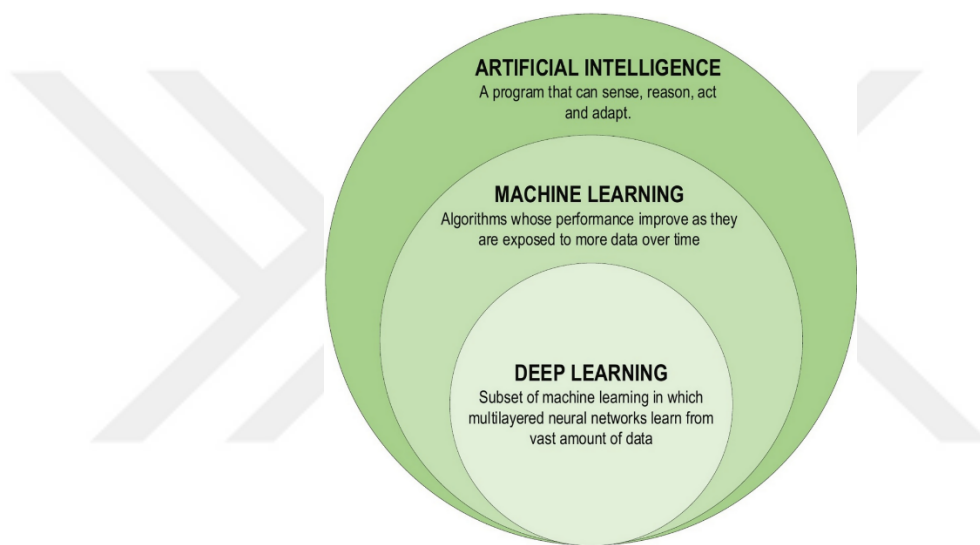


Figure B.1 DL family

DL is an artificial intelligence function that uses multiple layers to extract higher-level features from raw input stages. For instance, while inserting an image input into a deep neural network, the edges are identified in the first layers, and the most obvious features are extracted, then the higher layers identify deeper features. DL, which is currently one of the most important and practical ML techniques, has been very successful in many applications such as image analysis, speech recognition, and text recognition. This technique teaches how to detect and classify objects by extracting features from input images through two strategies: supervised and unsupervised with a unique architectural characteristic.

DL has shown high-level efficiency in various fields such as image classification, object detection, video processing, natural language processing, and speech recognition. Given that each application has its algorithm and processing method, various models and algorithms of DL have been introduced over the past few years.

B.1. Deep Learning vs. Machine Learning

DL and ML are the 2 most trending technologies in the world today. These technologies are often used interchangeably. While DL is a subset of ML, many people get confused between these 2 terminologies. Table B.1 and Figure B.1 show the differences between the DL and the ML.

Table B.1 Differences between the ML and the DL.

Machine Learning	Deep Learning
Works on small amount of Dataset for accuracy.	Works on Large amount of Dataset.
Dependent on Low-end Machine.	Heavily dependent on High-end Machine.
Divides the tasks into sub-tasks, solves them individually and finally combine the results.	Solves problem end to end.
Takes less time to train.	Takes longer time to train.
Testing time may increase.	Less time to test the data.

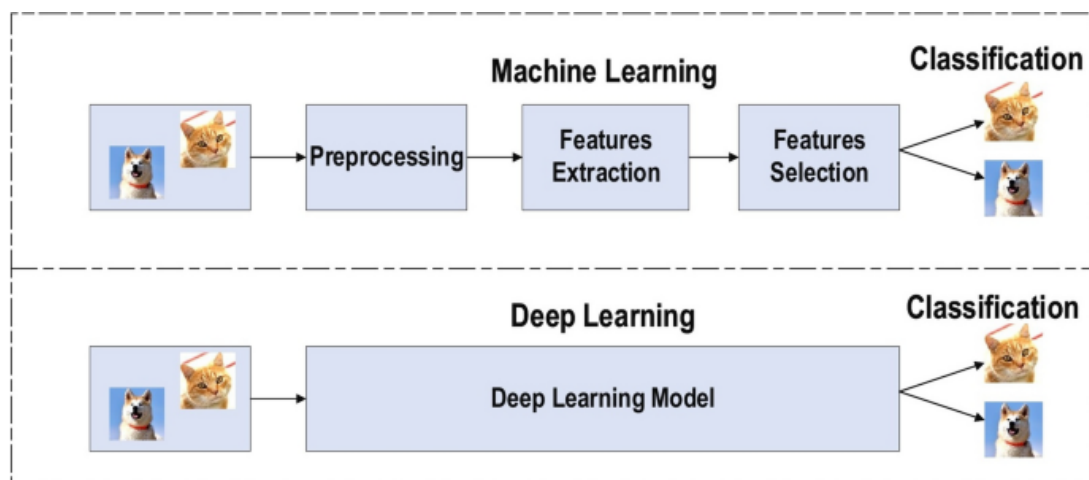


Figure B.2 Differences between the DL and the ML

ML is a subset of artificial intelligence associated with creating algorithms that can change themselves without human intervention to get the desired result, by feeding themselves through structured data.

DL is a subset of ML where algorithms are created and function similarly to ML, but there are many levels of these algorithms, each providing a different interpretation of the data it conveys. This network of algorithms is called artificial neural networks. In simple words, it resembles the neural connections that exist in the human brain.

B.2. Types of Deep Learning Approaches

DL approaches can be categorized as follows: supervised, semi-supervised or partially supervised, and unsupervised. Figure B.2 shows the pictorial diagram.

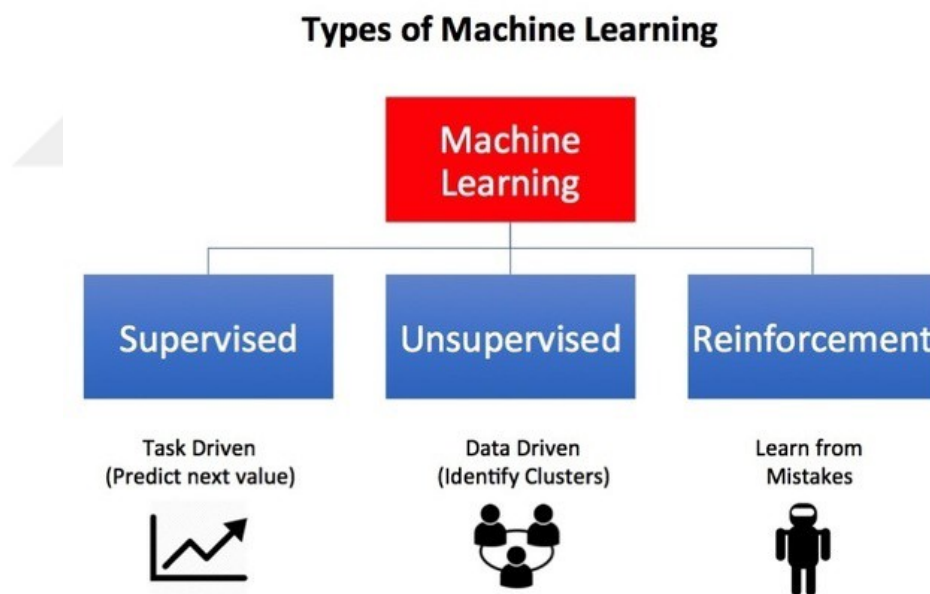


Figure B.3 Types of ML

B.3. Challenges of DL

There are several challenges for DL:

- Big data analytics using DL.
- Scalability of DL approaches

- The ability to generate data which is important where data is not available for learning the system (especially for computer vision task, such as inverse graphics).
- Energy efficient techniques for special purpose devices, including mobile intelligence, FPGAs, and so on.
- Multi-task and transfer learning or multi-module learning. This means learning from different domains or with different models together.
- Dealing with causality in learning.

B.4. Convolutional Neural Network

A typical CNN consists of many layers of hierarchy, with some layers for feature representations and others as a type of conventional neural networks for classification. There are two altering types of layers called convolutional and subsampling layers. The convolutional layers perform convolution operations with several filter maps of equal size, while subsampling layers reduce the sizes of proceeding layers by averaging pixels within a small neighborhood.

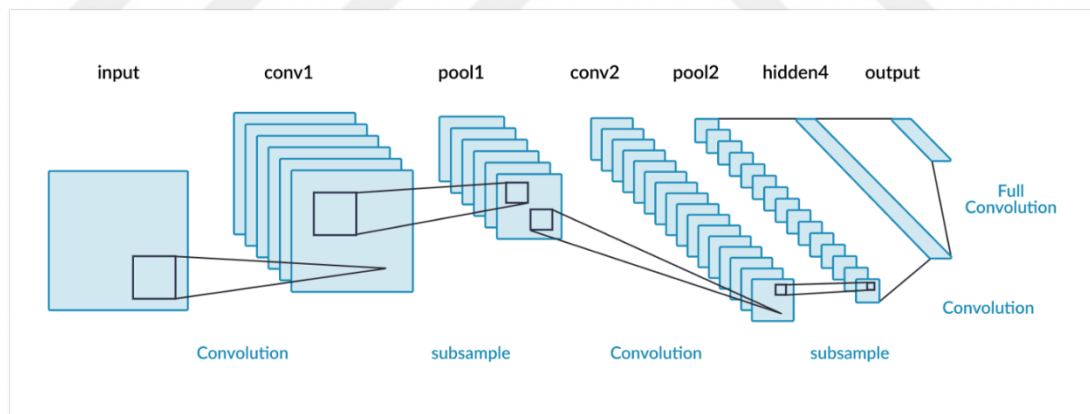


Figure B.4 Convolutional NN architecture

B.4.1. Convolutional Layer

The convolutional layer (Conv) is the main layer of this neural network architecture that determines the amount of output associated with the input to the receiver. Conv layer consists of filters, each composed of neurons that the neurons inside each filter are considered a kernel. Convolution kernels divide input images into smaller parts,

commonly known as receiver fields. The concept behind cutting and dividing the input into smaller sizes emphasizes key information making it easier to extract important features. The outputs are identified by the cores that process the length and width of the data.

B.4.2. Pooling Layer

The pooling layer usually comes between convolution layers. The main task of this layer is to reduce the dimensions of the feature map and the parameters. Using pooling operations helps to extract a combination of features. Reducing the size of the feature map to smaller subsets and excessive connections regulates the complexity of the network and increases performance. Pooling is the most widely discussed among the three layers.

B.4.3. Fully Connected Layer

In convolutional neural network architecture, one or more fully connected layer is defined after the last pooling layer. FC layer converts two-dimensional feature maps created from previous layers into a one-dimensional vector for displaying more features. Fully connected layers, composed of nearly 90% of a CNN model, act as a traditional neural network. It gives a predefined length vector of extracted features from its output, which can be used for classification purposes or considered a feature vector for reprocessing.

B.4.4. Activation Function

Activation functions are a set of mathematical equations that determine the output of each layer in the neural network. These functions are connected to each neuron in the network. This technique determines whether the functions should be activated or not due to the task of each neuron. Activating functions means deciding to stop processing and sending output to the next layer. Various activation functions such as Sigmoid, Tanh, Maxout, Swish, and ReLU and its types, such as leaky ReLU, ELU, and PReLU, have introduced different mathematical equations to be used in different NNs. Choosing the right activation function can speed up the learning

process because they understand nonlinear features. One of the recently considered activation functions is MISH, which has outperformed ReLU in DNNs.

B.4.5. Batch Normalization

Batch normalization or batch norm is a technique to train very deep neural networks. It is used to speed up the procedure, improve the performance and stability of artificial neural networks. Batch norm standardizes inputs into small layers according to their categories, making the training step more stable. It significantly reduces the number of cycles required for training a deep network.

B.4.6. Dropout

Dropout is a regularization technique introduced by Google Research that reduces excessive connections between neurons in the network by preventing complex coordination in the training data. This technique is a very impressive way to improve efficiency in neural networks, especially in CNNs. In neural networks, the various connections that learn a nonlinear relationship are closely fitted, causing overfitting. Dropout creates several thin network architectures by reducing some connections. A comprehensive network with small weights is selected at the end. The selected network architecture is considered for almost all other networks in the model, continuing the patch.

B.5. Convolutional neural network Models

CNNs are among the more popular neural network frameworks that are used in complex applications like DL models for computer vision and image recognition.

Over the years, CNNs have undergone a considerable amount of rework and advancement. This has left us with a plethora of CNN models. Let us now discuss the more important CNNs out of all these variants.

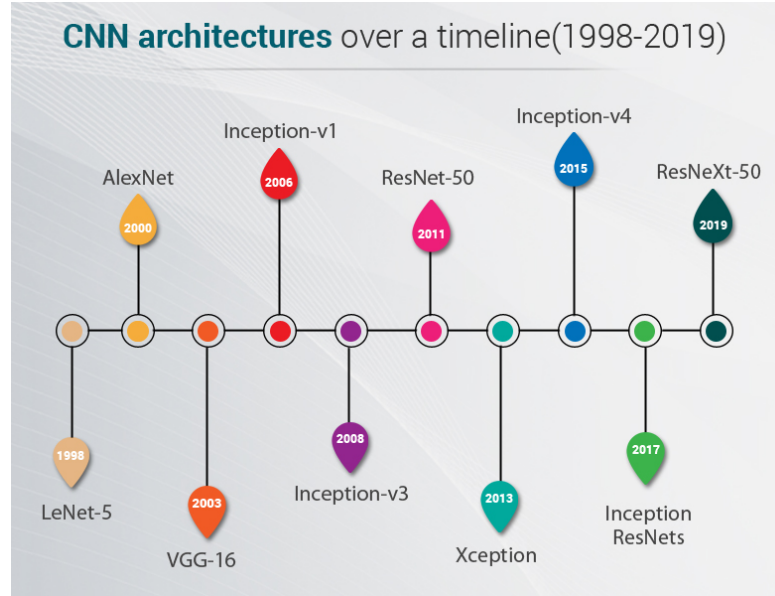


Figure B.5 CNN models

B.6. Computational Approaches

Evaluation metrics adopted within DL tasks play a crucial role in achieving the optimized classifier. They are utilized within a usual data classification procedure through two main stages: training and testing. They are utilized to optimize the classification algorithm during the training stage. This means that the evaluation metric is utilized to discriminate and select the optimized solution, e.g., as a discriminator, which can generate an extra-accurate forecast of upcoming evaluations related to a specific classifier. For the time being, the evaluation metric is utilized to measure the efficiency of the created classifier, e.g., as an evaluator, within the model testing stage using hidden data. As given in Eq. (B.1), TN and TP are defined as the number of negative and positive instances, respectively, which are successfully classified. In addition, FN and FP are defined as the number of misclassified positive and negative instances respectively. Next, some of the most well-known evaluation metrics are listed below.

1. Accuracy: Calculates the ratio of correct predicted classes to the total number of samples evaluated [Eq. (B.1)].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (B.1)$$

2. Sensitivity or Recall: Utilized to calculate the fraction of positive patterns that are correctly classified [Eq. (B.2)].

$$Sensitivity = \frac{TP}{TP + FN} \quad (B.2)$$

3. Specificity: Utilized to calculate the fraction of negative patterns that are correctly classified [Eq. (B.3)].

$$Specificity = \frac{TN}{FP + TN} \quad (B.3)$$

4. Precision: Utilized to calculate the positive patterns that are correctly predicted by all predicted patterns in a positive class [Eq. (B.4)].

$$Precision = \frac{TP}{TP + FP} \quad (B.4)$$

5. F1-Score: Calculates the harmonic average between recall and precision rates [Eq. (B.5)].

$$F1_{score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (B.5)$$

6. J Score: This metric is also called Youdens J statistic. Eq. B.6 represents the metric.

$$J_{score} = Sensitivity + Specificity - 1 \quad (B.6)$$

7. False Positive Rate (FPR): This metric refers to the possibility of a false alarm ratio as calculated in Eq. (B.7).

$$FPR = 1 - Specificity \quad (B.7)$$

8. Area Under the ROC Curve: AUC is a common ranking type of metric. It is utilized to conduct comparisons between learning algorithms, as well as to construct an optimal learning model. In contrast to probability and threshold metrics, the AUC value exposes the entire classifier ranking performance.

The following formula is used to calculate the AUC value for a 2-class problem [Eq.(B.8)].

$$AUC = \frac{S_p - n_p(n_n + 1) / 2}{n_p n_n} \quad (B.8)$$

Here, S_p represents the sum of all positive ranked samples. The number of negative and positive samples is denoted as n_n and n_p , respectively. Compared to the accuracy metrics, the AUC value was verified empirically and theoretically, making it very helpful for identifying an optimized solution and evaluating the classifier performance through classification training.

B.7. Limitations

1. Learning only through the observations that are made.
2. There is the issue of bias.

B.8. Advantages

1. It has the best in-class performance with regard to problems.
2. It reduces the need for feature engineering.
3. It can eliminate the expenditure of unnecessary costs.
4. It can easily identify defects that are usually quite difficult to detect.

B.9. Disadvantages

1. It requires a large amount.
2. Computationally, it is expensive to train.
3. It does not have a strong theoretical foundation.

B.10. Applications

1. Automatic text generation: The corpus of text is first learned and then, the new text will be generated, either character-by-character or word-by-word.
2. This model has the ability to learn how to spell a word, punctuate a sentence, form full sentences, and it can even capture the style of the writing.

3. Healthcare: It aids in performing the diagnosis of various diseases as well as treating them.
 4. Automatic machine translation: Some specific words, sentences, or phrases in a language is translated into some other language. DL has achieved top results in areas such as text and images.
 5. Recognition of images: It can recognize and identify individuals as well as objects in that are in images and can also understand the content and the context. This is already used in gaming, tourism, retail, and so on.
 6. Prediction of earthquakes: It can teach a computer how to perform viscoelastic computations, which are necessary for use in the prediction of earthquakes.
- 

Appendix C – Generative Adversarial Networks (GAN)

A GAN is a ML model. In a GAN, 2 NNs compete against each other to be able to predict more accurately. GANs can usually run without supervision and make use of a cooperative zero-sum game framework in their process of learning.

The 2 NNs that are within a GAN are known as the generator and discriminator, as depicted in Figure C.1. The generator is a CNN, while the discriminator is a DCNN. The generator's aim is to manufacture outputs artificially, that would be easy to mistake as real data. The discriminator's aim is identifying which of the outputs that it receives was created artificially.

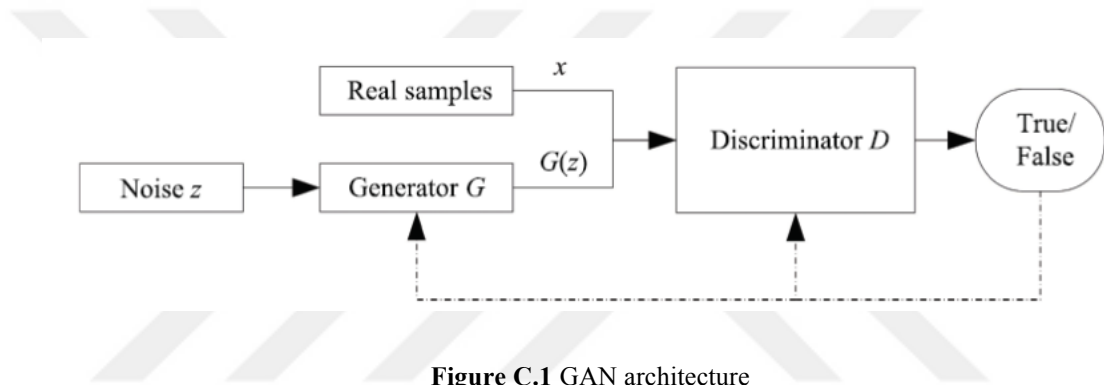


Figure C.1 GAN architecture

C.1 Discriminative vs. Generative

A GAN consists of 2 NNs. The generator generates the new data points from the random uniform distribution. The aim is producing similar false results from the inputs that are given.

The discriminator identifies the false data that was produced by the generator using real data.

The main aim in a GAN is training 2 different NNs to compete against each other using 2 different objective functions. The generator aims at fooling the discriminator into thinking that the input it was given is real. The discriminator gives a slap to the generator by identifying that this is fake. Then after taking the slap from the discriminator, the generator will learn to produce similar input training data. This

process is performed again and again or until Nash equilibrium is attained. This is known as adversarial training.

C.2 GAN Objective Function

The discriminator is a binary classifier; hence when it is fed real data, it should be able to produce a high probability for those data and a low probability for the false ones (generator output).

$z \rightarrow \text{Noise Vector}$ $G(z) \rightarrow \text{Generator is output} \rightarrow x_{fake}$

$x \rightarrow \text{training sample} \rightarrow x_{real}$

$D(x) \rightarrow \text{Discriminator is output for } x_{real} \rightarrow P(y|x_{real}) \rightarrow \{0,1\}$

$D(G(z)) \rightarrow \text{Discriminator is output for } x_{fake} \rightarrow P(y|x_{fake}) \rightarrow \{0,1\}$

Discriminator D	Generator G
$D(x) \rightarrow \text{should be maximized}$	$D(G(z)) \rightarrow \text{should be maximized}$
$D(G(z)) \rightarrow \text{should be minimized}$	

The $D(x)$, $D(G(z))$ have a score that is between 0 and 1. It is aimed to build a discriminator that can maximize the real data and minimize the false ones. $G(z)$ is in the same shape as the real input data, that is, if a 10×10 image is the genuine input, then $G(z)$ will produce the same shape, except it will be noisy. It is also aimed to build a generator that can maximize the fake data.

Table C. 1 Discriminator and generator losses equations.

Discriminator D	Generator G
$D_{loss}_{real} = \log(D(x))$	
$D_{loss}_{fake} = \log(1 - D(G(Z)))$	
$D_{loss} = D_{loss}_{real} + D_{loss}_{fake}$	$G_{loss} = \log(1 - D(G(Z)))$
$D_{loss} = \log(D(x)) + \log(1 - D(G(Z)))$	Or $-\log(D(G(z)))$
Total cost	is Total cost is.
$\frac{1}{m} \sum_{i=1}^m \log(D(x^i)) + \log(1 - D(G(z^i)))$	$\frac{1}{m} \sum_{i=1}^m \log(1 - D(G(z^i)))$
	Or $\frac{1}{m} \sum_{i=1}^m -\log(D(G(z^i)))$

As can be seen, the discriminator runs 2 times, that is 1 time for the real data, and 2 time for the fake data, prior to calculating the final loss. The generator only runs one time. After obtaining these 2 losses, the gradients are calculated in addition to their parameters, and they are then back propagated independently through their networks.

The discriminator and generator play the following 2-player mini-max game with value function $V(G, D)$:

$$\min_G \max_D V(D, G) = E_{x: P_{data}(x)}[\log D(x)] + E_{z: P_z(z)}[\log(1 - D(G(Z)))] \quad (C.1)$$

$$\max_D V(D) = E_{x: p_{data}(x)}[\log D(x)] + E_{z: p(z)}[\log(1 - D(G(z)))] \quad (C.2)$$

$$\max_G V(G) = E_{z: p(z)}[\log(1 - D(G(z)))] \quad (C.3)$$

C.3 Types of GANs with Math

C.3.1. Deep Convolutional GANs

The core idea is we use convolution neural networks instead of vanilla neural networks at both Discriminator D and Generator G. Discriminator D is a set of convolution layers with strided convolutions so it down samples the input image at every conv layer. Generator G is a set of convolution layers with fractional-strided

convolutions or Transpose convolutions so it up samples the input image at every conv layer.

C.3.2. Conditional GANs

Gan's can be extended to a conditional model if both the generator and discriminator are conditioned on some extra information y . y can typically be a class label or tag.

The core idea is to train a GAN with a conditioner, we can perform the conditioning by feeding y into the both the discriminator and generator as additional input layer.

$$\min_G \max_D V(D, G) = E_{x: p_{data}(x)} [\log D(x | y)] + E_{z: p_z(z)} [\log(1 - D(G(Z | y)))] \quad (C.4)$$

$$I_D^{CGAN} = E[\log(D(x, c))] + E[\log(1 - D(G(z), c))] \quad (C.5)$$

$$I_D^{CGAN} = E[\log(D(G(z), c))] \quad (C.6)$$

At Discriminator D , given input x and class $y \rightarrow$ classifies the input image is fake or real. at Generator G , given noise z , and class $y \rightarrow$ generates the image which is conditioned by y .

C.3.3. Least Square GANs (LSGAN)

As we know “Regular GAN's may lead to the vanishing gradients problem”, this slows the learning process. LSGAN attempts to overcome this problem by adopting the least squares loss function instead of the sigmoid cross entropy loss for the discriminator. The authors of the paper trained this on different datasets, and they have observed that there are two benefits of LSGANs over regular GANs. First, LSGANs can generate higher quality images than regular GANs. Second, LSGANs perform more stable during the learning process.

The objective function of LSGAN yields minimizing the Pearson χ^2 divergences.

$$\min_D V_{LSGAN}(D) = \frac{1}{2} E_{x: p_{data}(x)} [(D(x) - b)^2] + \frac{1}{2} E_{z: p_z(z)} [(D(G(Z)) - a)^2] \quad (C.7)$$

$$\min_G V_{LSGAN}(G) = \frac{1}{2} E_{z: p_z(z)} [(D(G(Z)) - c)^2] \quad (C.8)$$

where a and b are the labels for fake data and real data, respectively and c denotes the value that G wants D to believe for fake data.

C.3.4. Auxiliary Classifier GAN (ACGAN)

in ACGAN's, every generated sample has a corresponding class label, $c \sim p_c$ in addition to the noise z . G uses both (noise z and class c) to generate images.

$$X_{fake} = G(c, z) \quad (C.9)$$

The discriminator gives both a probability distribution over sources and a probability distribution over the class labels.

$$d(X) = P(S | X), P(C | X) \quad (C.10)$$

$X \rightarrow$ fake and real samples, S is the prob dist (fake or not), C is the prob dist (class label).

The objective function has two parts: the loglikelihood of the correct source, LS , and the log-likelihood of the correct class, LC .

$$L_S = E[\log P(S = real | X_{real})] + E[\log P(S = fake | X_{fake})] \quad (C.11)$$

$$L_C = E[\log P(C = c | X_{real})] + E[\log P(C = c | X_{fake})] \quad (C.12)$$

D is trained to maximize $LS + LC$ while G is trained to maximize $LC - LS$.

C.3.5. Laplacian GAN (LapGAN)

A sequential image generation framework Laplacian GAN (LapGAN) proposed by combining the CGAN model with the framework of the Laplacian pyramid (LP). LapGAN requires the multiscale generation process in which a series of the GAN generates particular levels of details of an image in an LP representation. The GAN at each generation step of the LP can be different. The LP was built from a Gaussian

pyramid (GP) using up-sampling $u(\cdot)$ and down-sampling $d(\cdot)$ functions explained as:

Let $G(I)=[I_0, I_1, \dots, I_K]$ be the GP where $I_0 = I$ and I_K are k repetitive applications of $d(\cdot)$ to I . Then, the coefficients h_k at each level k of the LP ($L(I)$) is given as:

$$h_k = L_k(I) = G_k(I) - u(G_{k+1}(I) = I_k - u(I_{k+1}))$$

(C.13)

C.3.6. Information Maximizing GAN (InfoGAN)

Information Maximizing GAN (InfoGAN) proposed an idea of a representation learning algorithm that can learn disentangled design in a wholly unsupervised way. InfoGAN is a completely unsupervised framework built on top of GAN and disentangles both discrete and continuous latent factors, scale to complicated datasets, and requires no more training time than GAN. The InfoGAN training objective is:

$$L_{InfoGAN} = V(D, G) - \lambda I(c; G(z, c)) \quad (C.14)$$

where λ refers to a hyper-parameter, z refers to un-interpretable noise, c encodes the salient latent codes, and I is for the mutual information /shared information.

C.3.7. Energy-Based GAN (EBGAN)

Energy-Based GAN (EBGAN) is a variant of the GAN architecture which combined AE and GAN frameworks where discriminator works as an energy function refers to the low energy to actual data and high energy to fake data instead of a routine GAN probability function that determines input as actual or fake. EBGAN used two distinct losses for the training of both G and D . When the generator of EBAGN was far from convergence, they got better in quality gradients and performance.

$$L_D(x, z) = D(x) + [m - D(G(z))]$$

(C.15)

$$L_G(z) = D(G(Z))$$

(C.16)

L_G , L_D and D are the correspondence with generator, discriminator, and reconstruction losses. These equations also show that minimizing generator loss L_G concerning parameters of G is like maximizing the L_D concerning parameter D has the same minimum for the positive margin m .

C.3.8. Wasserstein GAN (WGAN)

Wasserstein GAN (WGAN) proposed a substitute loss function derived through Earthmover (EM) or Wasserstein distance. Not like the standard GAN cost function where discriminator works as a binary classifier function, the discriminator in WGAN used to fit the Wasserstein distance. WGAN is the most straight-forward to train using an alternate cost function that is not suffered from the vanishing gradient problem and partially remove the MC obstacle to stabilize the GAN training and get better results in term of MC problem.

C.3.9. Boundary Equilibrium GAN (BEGAN)

Boundary equilibrium GAN (BEGAN) keep-up an equilibrium that manages the trade-off between variety and superiority. The main goal behind BEGAN is to change the loss function. The Wasserstein distance between reconstruction loss of actual and synthesized images gives the real loss. In BEGAN, the discriminator works during training as autoencoder balances the process optimizing of G and D . The idea of making the discriminator as an autoencoder first proposed in EBGAN. BEGAN cost function is:

$$L_D(x, z) = D(x) - k_t D(G(z)) \text{ for } \theta_D \quad (\text{C.17})$$

$$L_G(z) = D(G(Z)) \text{ for } \theta_{DG} \quad (\text{C.18})$$

$$L_{K+1} = k_t + \alpha(\gamma D(x) - D(G(z))) \quad (\text{C.20})$$

where L_G represents the loss of the generator, L_D represents the loss of the discriminator, $L(x)$, $L(G(z))$ represents the auto-encoder L_1 loss of real, fake data, and equilibrium hyper-parameter “ γ ” respectively.

C.3.10. BigGAN

Due to its outstanding, large scale, indistinguishable, and high-quality image generation capacity, Big GAN (BigGAN) is one of the present best models. BigGAN-a DL model-can train bigger neural networks even more parameters; create a more extremely detailed image with remarkable performance. BigGAN have some essential properties such as provides exerts control over the outputs, provides interpolation phenomena between images, which means that if there are two images, it can compute the intermediate image between them and provides the best inception score (IS), i.e., the best of earlier works had an inception score (IS) around 50, but the inception score (IS) of BigGAN technique is not less than 166, which is closer to real images which would score around 233. Extension to BigGAN proposed called Bi-Directional Big GAN (BigBiGAN) improves the un-conditional image generation and representation learning capacity of the model such as increased freshet inception distance (FID) and inception score (IS) accuracy score over the baseline BigGAN model for un-conditional results.

C.3.11. Progressive-Growing GAN (PGGAN)

The Progressively Growing GAN (PGGAN) proposed a multi-scale-based GAN architecture where both G and D start its training with low-resolution image (e.g., 4×4), gradually increase the model depth by adding-up the new layers to both G and D during the training process, and end-up with the generation of large scale sharp image (e.g., 1024×1024). The basic idea behind PGGAN is to grow both G and D in synchrony, i.e., starting from a low-resolution image (e.g.42), doubles the resolution of generated image (e.g.82) with the addition of new layers to the both networks, and ends-up with the generation of high-resolution image (e.g. 1024^2) as the training progresses. The growing training approach of PGGAN enables stable learning for both the networks in big resolutions, reduced training time, and workaround GAN instability training problems.

C.3.12. Style-Based Generator Architecture for GAN (StyleGAN)

Although PGGAN generates a high-quality image, its ability to control specific features of the generated image is minimal. To curb this issue, style-based generator architecture for GAN (StyleGAN) redesigns the architecture of the generator network, makes it possible to control the image synthesis through scale-specific amendments to the styles without compromising the generated image quality but increases it significantly utilizing PGGAN. Infact, StyleGAN introduced an upgraded version of PGGAN, which only focusing on the generator network to control the co-relation between input features. StyleGAN divides the input features into three types, such as (i) coarse features-pose, hair, face, shape, (ii) medium features-facial features, eyes, and (iii) fine features-color scheme. The StyleGAN is famous for its un-conventional GAN architecture such as the use of mapping network that first transforms the input latent code into inter-mediate latent code where affine transformation then produce styles that control the layers of the synthesis network through Adaptive Instance Normalization (AdaIN) that scales the normalized input with style spatial statistics, and the PGGAN hat has been extremely flourishing in stabilizing large-resolution GAN training. StyleGAN2 comes with various improvements to image quality, efficiency, diversity, and disentanglement, and the results are incredibly improved. StyleGAN2 simply redesigns the normalization used in the generator of StyleGAN, which removes the artifacts, such as blob-shaped artifacts that resemble water droplets. The StyleGAN2 achieves excellent results in face image synthesis and quality when compared to StyleGAN.

C.4. Applications of GANs

GANs have a lot of real-life applications, some of which are:

- Generate Examples for Image Datasets
- Generating examples is very handy in medicine or material science, where there's very little data to work with.
- Generate Photographs of Human Faces
- Video game designers can use this to generate realistic human faces.
- Generate Realistic Photographs
- Very useful for photographers and videographers.

- Generate Cartoon Characters
- Artists can use this to create a new character design, or scenes in a cartoon, or even in a video game.
- Image-to-Image Translation
- Photographers can use these algorithms to convert day into night, summer into winter, etc.
- GANs can be used to simulate a worst-case scenario to optimize risk management in a business.
- Other use cases of GAN could be:
 - Text-to-Image Translation
 - Face Frontal View Generation
 - Generate New Human Poses
 - Photos to Emojis
 - Face Aging
 - Photo Inpainting
 - Clothing Translation
 - Video Prediction
 - 3D Object Generation

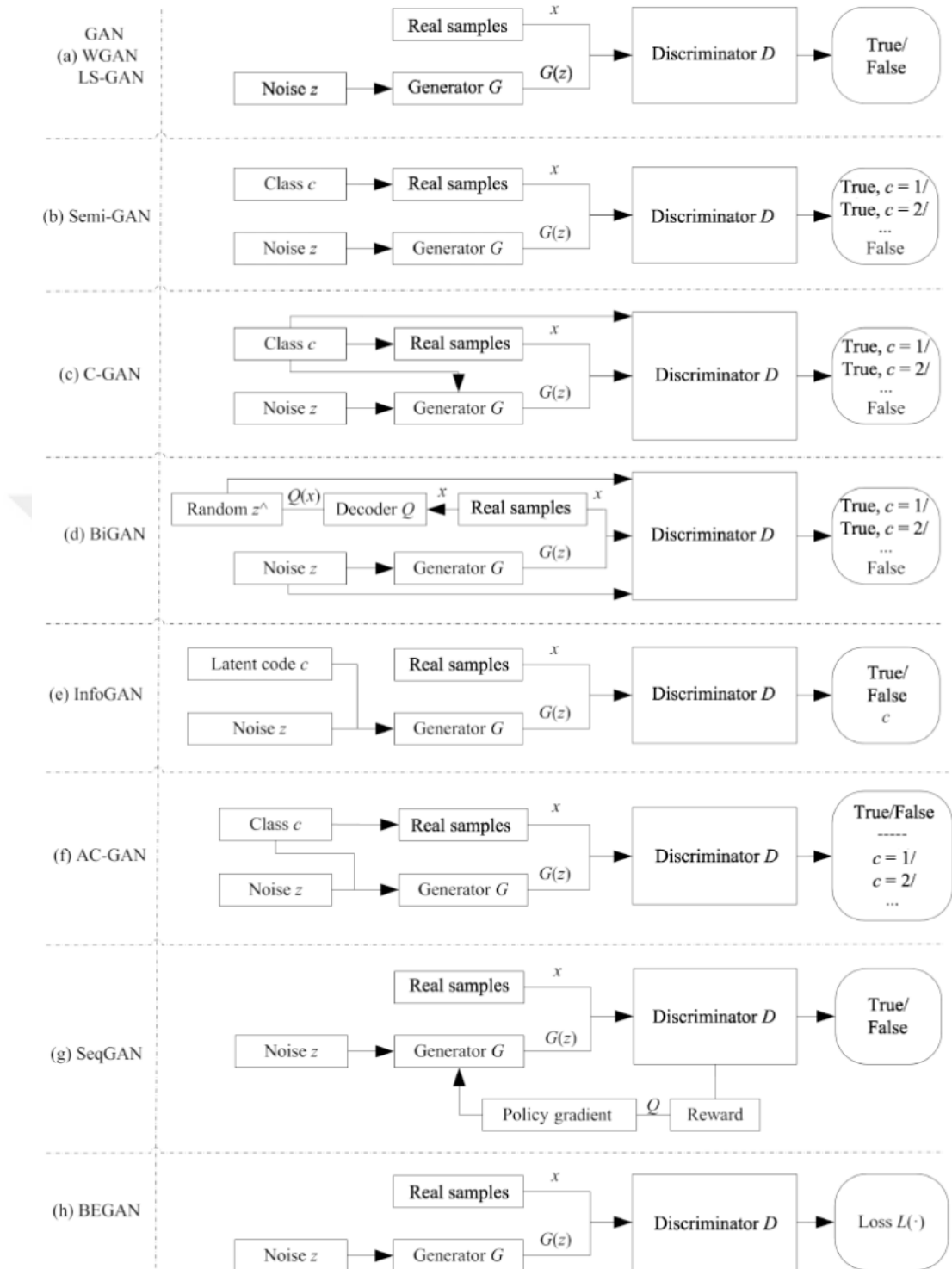


Figure C. 2 Some of variants of the GAN.





Khaled Alashik

Department of Electrical and Computer Engineering

July 2021 ANKARA