COMBINING FACE AND IRIS FOR PRIVACY PRESERVATION

By

Achsah Junia Ledala

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Computer Science – Master of Science

2021

ABSTRACT

COMBINING FACE AND IRIS FOR PRIVACY PRESERVATION

By

Achsah Junia Ledala

With the extensive use of biometrics for authenticating users, the need to ensure privacy of biometric data is greater than ever before. Biometric authentication systems are vulnerable to attacks and the loss of biometric data will lead to loss of privacy of an individual. Multibiometrics refers to the use of multiple biometric modalities simultaneously in order to perform matching. In this work, we introduce a multibiometric fusion technique which can be used to ensure that the original raw biometric data are unlikely to be compromised and, at the same time, recognition can be performed. The face and the iris biometric modalities are fused at the feature-level to produce discriminative embeddings that can be used for recognition. The original face or the iris cannot be retrieved from the combined representation, thus preserving the privacy of the individual. We present the results of this approach, provide analysis, discuss the challenges, and list possible future directions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Biometrics

Biometrics is a field of science which deals with the recognition of individuals based on their physical or behavioral characteristics [32]. These characteristics are referred to as biometric modalities. Individuals can be recognized using these modalities which are considered to be almost distinct to every individual. Some of the biometric modalities that can be used for recognition are face, fingerprint, iris, voice, palm, hand geometry, palm veins, signature, and gait [30]. Since biometric modalities are distinct to every individual and cannot be re-generated, preserving the privacy of the biometrics is of utmost importance. Once the biometric data or template is compromised, it becomes difficult to replace it, which can lead to a loss of privacy of an individual. Therefore, it is imperative to secure the biometric data to prevent the loss of privacy. There are multiple ways in which this can be achieved, viz., by securing the database in which the biometric data is stored, by encrypting the biometric data, or by making the biometric data itself more secure. In this study, we explore the possibility of making the data itself more secure by combining two biometric modalities at the feature level such that the privacy of an individual is preserved whilst retaining the recognition accuracy.

1.2 Recognition In Biometrics

Recognition in biometrics is typically classified into identification and verification [31]. In identification, an individual is recognized from a group of individuals. In verification, the claimed identity is validated. Biometrics has a wide range of applications in various domains, such as forensics, border control, airport security, surveillance, chain of custody, financial transactions, mobile authentication, deduplication, facility access, etc. A biometric system operates in two phases — an enrollment phase and a verification phase [30]. During the enrollment phase, the

raw biometric modalities of several individuals are collected and stored in the database. During the verification stage, an individual once again presents their biometric modality either for identification or verification. Identification deals with the question “Who am I?" whereas Verification deals with the question “Am I who I claim to be?" The presented biometric modality is compared against all the stored biometric identities to determine the correct identity in case of identification so it is referred to as “one-to-many" matching, and compared against the stored claimed identity in case of verification so it is referred to as “one-to-one" matching. Therefore, a biometric recognition system is a pattern recognition system which typically has four main building blocks:

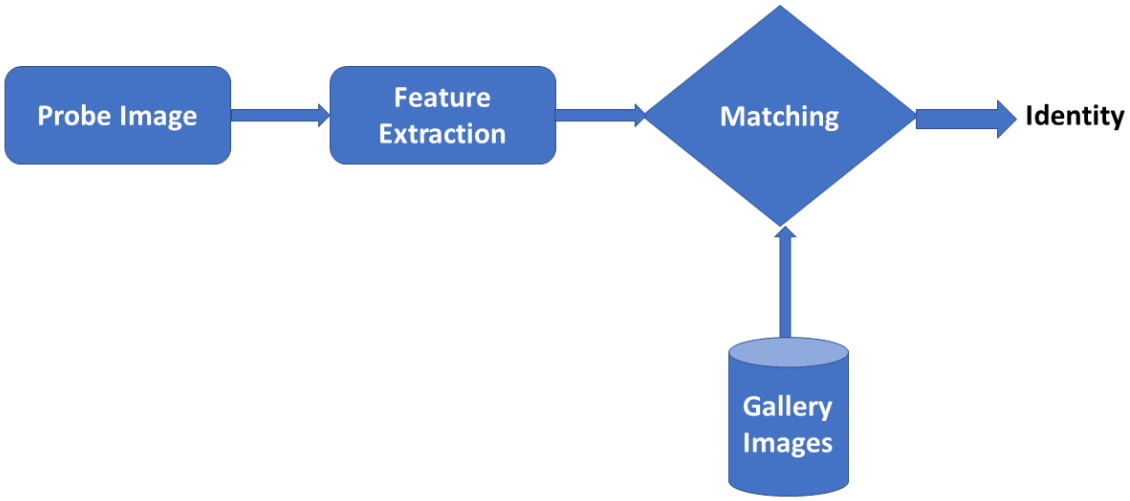

Figure 1.1: Representation of a biometric system working in identification mode

- **Data Acquisition:** A sensor is used to acquire the raw biometric data from an individual. Usually the raw data are 2D images but the data can also be in the form of video, audio, text, etc. Different sensors are designed for capturing different modalities and for different use-cases.

- **Feature Extraction:** In the feature extraction module, the raw biometric data is subjected to pre-processing to retain only the relevant distinct features useful for matching and to eliminate

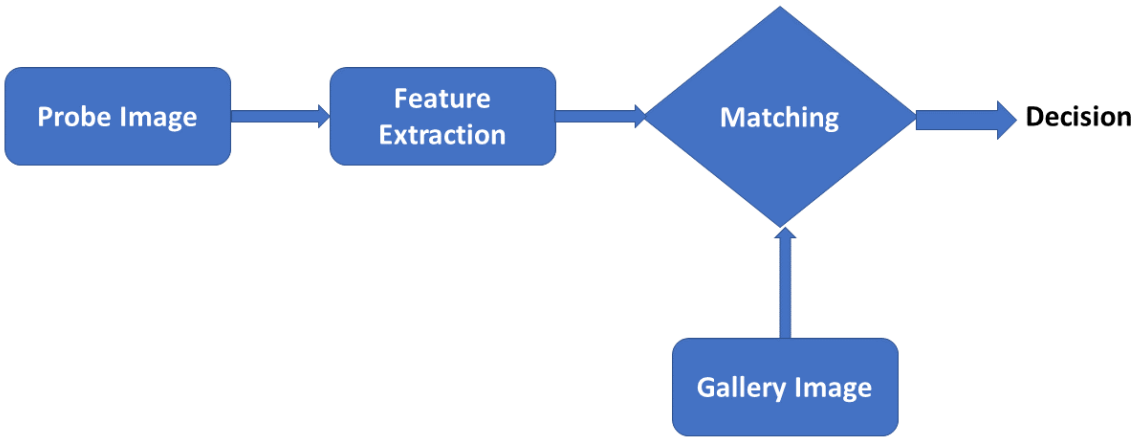

Figure 1.2: Representation of a biometric system working in verification mode

unnecessary and noisy information. For fingerprint images, minutiae points are extracted as features [28]. Iris code is extracted as the feature template for iris images [35].

- **Data Storage:** Once the relevant feature vector is extracted from the raw data, the template is stored in a database along with the identity information or with an identifier. Sometimes, the raw data is also stored along with the feature template. The set of raw data (e.g.,images) stored in the database is called the gallery.

- **Decision Module:** This is a matching stage where an individual once again presents the raw biometric data, which is called a probe, and this probe is compared with the set of gallery images in identification to determine the identity or compared with claimed identity in verification to find a match or a non-match. A match score is generated in the decision module which can be a similarity score or a dissimilarity score.

1.3 Face Recognition

Determining an identity or confirming an individual's identity based on facial features is referred to as face recognition. In face verification, two face images are compared to determine if they belong to the same person or not. This is the most common way that humans identify each other, and it is a widely used biometric modality for identity authentication. Although the exact cognitive process of human face recognition is still unknown, it is perceived that humans recognize faces based on high-level characteristics such as two eyes, nose, and mouth. However, Automatic Face Recognition (AFR) is a more difficult task as it has several challenges, such as occlusion, aging, pose variations, illumination, resolution, expression, makeup, facial hair, accessories, inter-class similarities (similarities between twins and people within the same family), large intra-class variations (variation in the images of the same identity), unconstrained environments, etc. These variations affect the performance of a face recognition system. In-spite of these challenges, state of the art deep learning networks [85] [77] [81] for face recognition achieve matching accuracies greater than 97%. Face recognition can be performed in the visible, infrared and near-infrared (NIR) spectrum [95]. However, recognition using 2D images in the visible spectrum is more common. The applications of face recognition include surveillance, image-retrieval, law enforcement, and human-computer interaction (HCI). The goal of an AFR system is to generate a match score that indicates the similarity or dissimilarity of two face images. An efficient face recognition system has the following modules: (a) Data Acquisition, (b) Face Detection, and (c) Feature Extraction and Matching

1.3.1 Data Acquisition

The Automatic Face Recognition (AFR) system first begins with acquiring the face images. Different sensors can be used to obtain different formats of data. 2D images, sequence of 2D images (videos), and 3D (depth) images are the common formats used in AFR systems. Usually frontal views of the face are considered for matching in AFR since the frontal view of a face contains more

detail of the face compared to side views of the face. Since the face is a 3D image, capturing the 2D image of the face may occlude some of the facial features. This is called self-occlusion. Face images can be captured in the visible and NIR spectrum. An NIR camera can be used to covertly capture the images of the face in a dark environment.

1.3.2 Face Detection

Face detection is a problem of locating and localizing one or more faces in an image.Face detection has been solved reasonably well by classical feature-based techniques, such as the cascade classifier. More recently deep learning methods, like MTCNN [93] have achieved state-of-the-art results on face detection datasets. Face detection can also be done based on the skin color [6] as the skin color is distinct from background objects. However, the skin color is constrained to certain ethnic groups. Deep neural networks can be trained to detect frontal and non-frontal faces, faces with different poses and rotations. Zhang [98] analyzed face detection in multispectral illuminations. Principal Component Analysis and Support Vector Machines are effective techniques used in face detection.

1.3.3 Feature Extraction And Matching

The feature extraction module involves retaining the most discriminatory features from the localized and detected face for further matching. The matching step is the final step in the AFR system, and the output of matching is a match score that indicates the similarity or dissimilarity between the compared probe and gallery face images, based on which decision is made. After the features are extracted from the face image in the feature extraction step, this feature vector is compared to the face template whose identity is known and stored in the database. Based on whether it is a verification or an identification task, algorithms such as correlation filters, convolutional neural networks(CNNs), K-means, support vector machines (SVMs), Decision trees (DTs), K-nearest neighbor (K-NN), recurrent neural networks (RNNs) are used to effectively perform the matching

process [94]. We discuss three main approaches in the feature extraction and matching module in detail - model-based approach, texture-based approach, and deep learning-based approach.

1.3.3.1 Model-based Approach

The Model-based approaches to automatic face recognition are more complex than appearance-based approaches because it involves the detection of landmark points on the face such as eyes, corners of mouth, eyes, tip of nose, and chin. Model-based approach derives a pose invariant face representation that can be used for matching face images across different poses. The Elastic Bunch Graph Matching is an example of a model based face recognition algorithm.

Elastic Face Bunch Graph Matching : In this approach, the face image is represented as a graph. The nodes of the graph correspond to landmark points on the face and the edges between the nodes represent the distance between the landmark points on the face. The Gabor coefficients, called jets, are obtained at each node by applying the Gabor wavelet transform centered at the landmark position corresponding to that node. The Face Bunch Graph is obtained by combining all the individual graphs like a stack. Thus, each node represents the variation in the landmark points across all images in the gallery. During matching, the graphs for both the probe and gallery image are generated, and the similarity between the probe graph and the gallery graph is the average similarity of the Gabor coefficients at the corresponding landmark points.

1.3.3.2 Texture-based Approach

Texture is the perceived arrangement of details (pixels) in an image. Textural analysis is a region based property and can be used to describe, segment and classify regions according to its texture content. Texture based approaches are invariant to pose and illumination conditions. Local Binary Pattern (LBP) and Scale Invariant Feature Transform (SIFT) are two techniques in object recognition that are widely used to characterize the texture of an image.

Local Binary Pattern: In the Local Binary Pattern technique, a neighbourhood patch of 3*3 is considered. The intensity values of the surrounding 8 pixels is compared with the intensity of the central pixel and a 8 bit binary string is obtained. The string is then bitwise multiplied with its weights and summed up to obtain a decimal value for the central pixel. After this is done for all the pixels, the face images are divided into smaller regions and histograms of local binary patterns for each region are computed and concatenated to generate a global feature vector which can then be normalized. During matching, the similarity or distance measure between the feature vectors of probe and gallery image is computed. Extensions of the original LBP approach include Local Ternary Patterns [87] and Dynamic Threshold Local Binary Pattern [44] which are less sensitive to noise than LBP. Local Derivative Pattern proposed in [96] can represent more information than LBP because it uses features of higher order.

Scale Invariant Feature Transform: The Scale Invariant Feature Transform consists of two main steps: (1) Determining the key points and (2) Descriptor calculation in the neighbourhood of each key point. The key points can be used to achieve tolerance against pose variation. The descriptor is a histogram of gradient orientations. Each face image is divided into multiple regions and the SIFT descriptors from each region are concatenated to generate a final descriptor which can be used for matching. This technique is invariant to scaling, illumination, rotation, and translation.

1.3.3.3 Deep Learning-based Approach

Traditional methods like model-based and texture-based approaches perform recognition based on one or two layer representations. Continuous efforts were made to improve the recognition accuracy by separately improving the pre-processing, feature extraction and matching stages of recognition. Despite these improvements, the highest recognition accuracy that could be obtained on the LFW dataset was about 95% [17]. Deep learning methods such as Convolutional Neural Networks (CNNs) outperform the traditional methods by a large margin as they learn multiple layers of representation corresponding to various layers of abstraction by using a cascade of processing

layers for feature extraction and transformation. Each layer of the neural network learns more complex features than the previous layer. The initial layers learn local features similar to the SIFT features whereas the higher layers learn more high-level abstractions. The combination of these high level features represents facial identity that is invariant to changes in the real-world. Using a deep learning approach, DeepFace model [84] achieved an accuracy of 97.35% on the LFW for the first time in 2014.

Face Processing: The performance of deep face recognition is still affected by changes in illumination, occlusion, expression, and poses. To address this problem, face processing methods such as “One-to-many augmentation" and “Many-to-one normalization" are widely used. In One-to-many augmentation, images of slight variations and poses are generated from a single source image to help the deep neural networks learn pose invariant representations. In Many-to-one normalization, a generic face image is constructed from a group of face images of non-frontal view.

Deep Feature Extraction: The deep learning algorithms are trained using large amount of data using an appropriate loss function for supervision. The loss function determines the error in the prediction. This phase is known as the training phase which involves minimization of the loss function to obtain the most appropriate parameters. The loss function is minimized by back propogating the errors to the previous layer, and an optimizer function is used to modify the weights and bias by calculating the gradient of the loss function with respect to weights. When trained on very large datasets, deep convolutional neural networks can learn highly discriminative and invariant feature representations. The effectiveness of an AFR system depends on the quality of the extracted features because facial landmarks identified by a given model determine how well the features are represented. In CNNs, a fiducial point detector is used to detect the landmarks on the face such as corners of mouth, tip of nose, center of the eyes, etc. After these landmarks are detected, the face is aligned according to normalized canonical coordinates [7]. From the aligned face, a feature vector is extracted which encodes the identity information. Some of the popular and widely used

CNN architectures in AFR are AlexNet, VGGNet, GoogleNet and ResNet [24, 38, 80, 83].

Face Matching: During the testing phase, the trained neural network is presented with a probe face image and the network generates a feature representation of this image. A distance measure such as euclidean distance or cosine distance is used to compute the similarity or dissimilarity score between the feature vector of the probe image and feature vectors of all the gallery images in case of identification and compared to the claimed gallery image's feature vector in case of verification to complete the matching process. If the score is below a predefined threshold, the two faces are considered to belong to the same identity. In some methods, the deep features are subject to post processing to ensure that the matching process is effective.

1.4 Iris Recognition

Iris can be used to recognize an individual because of the rich texture of the iris that is generally distinct to every individual. Since even the two irises of the same individual and iris of twins are unique, iris recognition is one of the most secure methods of authentication [82]. The non-intrusive nature of iris acquisition combined with its distinctiveness makes it a well sought-after biometric modality for recognition. The challenges to iris recognition are occlusion by eyelids, eyelashes, glasses, contact lens, out of focus imaging, camera angle, illumination, specular reflections, iris dilation, fake or printed contact lens, low resolution, head rotation, poor contrast, etc. John Daugman was the pioneer who designed and implemented the first Automatic Iris Recognition system (AIR). The goal of an AIR is to generate a match score that indicates the similarity or dissimilarity of two iris images. There are four steps in performing iris recognition. They are: (a) Image Acquisition, (b) Segmentation, (c) Normalization, (d) Feature Extraction, and (e) Matching

1.4.1 Image Acquisition

In image acquisition, 2D images of the eye are typically captured in the near infrared (NIR) spectrum because the rich complex texture of iris can be well captured in this spectrum. Iris

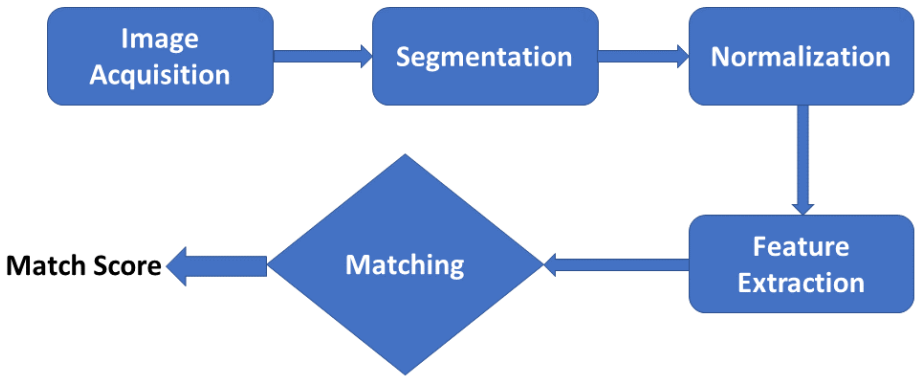

Figure 1.3: Modules of automatic iris recognition system

images taken in the NIR spectrum are non-intrusive and capture the complex textures of even dark colored irises. Although the human eye is sensitive to visible light, matching has also been performed on iris images captured in the visible spectrum [59]. However, it has been observed that the images captured in visible spectrum in unconstrained environment results in severely degraded images [27, 58].The difference in processing the iris image in visible spectrum versus NIR is that pupillary boundary is more distinct in NIR whereas limbic boundary appears to be more distinct in the visible region. These images are subjected to quality analysis to retain the images which have sufficient detail of iris texture for further pre-processing and to remove images that have occlusions and specular reflections. Some of the research in the iris image acquisition is the “stand off" subject to sensor distance of up to 3 meters [18, 92], wavefront-coded imagery [10], recognition with different wavelengths [68], using a dual-CCD camera to acquire one image in RGB and one in NIR to handle “off angle" views of the iris [15] and acquiring iris images from moving subjects [50]. The pre-processing module removes the noise and enhances the image for segmentation.

1.4.2 Segmentation

In segmentation, the aim is to detect the iris portion from the image of the eye. This module also involves the detection of eyelids and eyelashes. The inner and outer boundaries of the iris are identified from the image of the eye and isolated. The integro-differential operator has been proposed by Daugman to locate the inner and outer contours of the iris. Though initially the boundaries were assumed to be circular, later works are focused on removing that assumption [11, 26, 79] as the evaluation of results indicate improvement over methods that assume circular boundaries. Although similar techniques of the initial iris segmentation are used for finding the inner and outer boundaries of iris, different approaches have been developed to detect occlusion by eyelashes, eyelids and specular reflections in the field of iris segmentation. Research in this area include finding the pupil center of an iris [36, 41, 42] and then calculating the iris boundaries, segmenting the iris region under unconstrained conditions [13, 14], segmentation of an iris in frames of a video sequence [19, 43], detection of specular reflection in the iris image [78, 97], detection of eyelids in the iris image [48], evaluation of the quality of an iris segmentation [33, 45, 46] and segmentation on images obtained in visible spectrum [13, 14, 27, 36, 58]. Segmentation using geometric active contours and opening operators to suppress the interference from eyelashes was proposed by Roy and Bhattacharya [39, 73, 74]. Pundlik et al. [60] treat the image as a graph where the pixels are nodes and neighbouring pixels are joined using edges. Ryan et al. [76] proposed the “starburst method" for segmenting the iris. The quality of iris image that is obtained after segmentation will affect the matching performance of the AIR system.

1.4.3 Normalization

Normalization techniques ensure that the iris of different sizes are mapped to a common image domain. The difference in sizes of the iris occurs due to the dilation and contraction of the pupil, sensor distance from the subject and intrinsic variations in the size of the pupil across individuals. Normalization accounts for variations in the size of the iris. This stage is also referred to as the ‘unwrapping of the iris’. Geometric normalization transforms the cartesian coordinates to pseudo

polar coordinates using the rubber sheet model. The segmented iris is represented as a rectangular image and the rows of the rectangular image correspond to the concentric regions of the iris. Every point in the circular region between the outer and inner boundary of the iris is re-mapped to the normalized polar coordinates.

1.4.4 Feature Extraction And Matching

This step involves the extraction of the discriminating features of the iris image which is then encoded to construct a feature vector. In the Daugman approach, the unwrapped iris is convolved with a set of Gabor filters to extract a binary code called IrisCode. During matching, two iris codes are compared using hamming distance which counts the number of corresponding bits that are different between these two iris codes and a match score is generated which is used to make a decision. The matching algorithm depends on the type of encoding that has been used to construct the feature vector. Some other methods that have been proposed for feature extraction and matching are PCA [16], weighted PCA [99], 2D PCA and LDA [12], wavelet analysis [56], oriented separable wavelet transforms [88], normalized phase correlation [37], log-Gabor filters [5], rotated complex wavelet transform [9], 2D discrete wavelet transform [86]. Apart from the traditional feature extraction and matching approaches, some of the other methods that do not fit in the traditional approach are explored such as gray level co-occurrence matrices [23], local-global graph methodology [34], continuous dynamic programming [63], use of statistical measures [40], use of SIFT features [57]. However, none of these methods have shown better performance over the more traditional approaches. There are also approaches [3, 8, 20, 47, 61] that analyze the iris in individual parts and combine the results. This is done to reduce the noise and overcome the errors due to segmentation. Research has been done to improve the speed of iris matching from a large dataset. Mehrotra et al. [51] propose an indexing technique to reduce the search time whereas Roy and Bhattacharya [39, 72–75] apply feature selection to choose the most discriminating features thereby reducing the time taken for matching. Gentile et al. [21] uses a shorter iris code to index into a large iris database in order to reduce the number of comparisons to find the correct match.

Rathgeb et al. [65] proposes an "incremental" iris code matching to reduce the number of bit comparisons.

1.5 Thesis Organization

The remaining document is organized as follows:

Chapter 2 gives a detailed overview of the advantages and challenges of the multibiometric systems, fusion of biometric modalities at different levels, privacy concerns and threats in biometrics, and techniques used to ensure privacy of the biometric data and their challenges.

Chapter 3 describes the process of feature-level fusion of face and iris for the purpose of recognition and discusses the architecture of the proposed method and describes the deep learning models used.

Chapter 4 discusses the datasets that were used, the experimental setup, and presents the results of the performance on two datasets. Comparison between the individual face and iris recognition, and the fused face-iris modality recognition is done, and the results and analysis are presented.

Chapters 5 presents the conclusion and the future work that can be explored in this area.

CHAPTER 2

MULTIBIOMETRIC SYSTEMS AND PRIVACY

2.1 Multibiometric System

Biometric recognition systems usually use a single biometric modality to establish or verify an identity. These systems are called unibiometric systems. Some of the limitations of unibiometric systems are [71]: 1. Upper bound on identification accuracy - there is a limit beyond which we cannot improve the accuracy by fine tuning the matching or feature extraction process, 2. Spoof attacks - an impostor can try to fool the authentication system by pretending to be a genuine user, 3. Noise - the quality of the data that is presented for matching may be poor or noisy due to substandard acquisition conditions. Multibiometric systems aim to overcome the limitations posed by the unibiometric systems. Multibiometric system refers to the use of multiple biometric samples simultaneously in order to perform matching. These samples can be from different modalities; for example, fingerprint and face or images from different sensors; for example, 2D and 3D or repeated samples of the same biometric modality. These systems are expected to be more reliable because they consider multiple pieces of evidence. They can also help in anti spoofing as it is difficult to spoof multiple biometric modalities simultaneously. These systems are fault tolerant because they can continue to operate even when some modalities become unreliable due to hardware, software or intentional manipulation. Multibiometric systems enhance the matching accuracy, speed and security of the biometric systems [67]. Such a system can be implemented by the integration or fusion of multiple biometric modalities.

2.2 Fusion

Combining multiple sources of information with the goal to improve the overall matching accuracy is called fusion. Specifically, fusion in the case of verification is used to improve the accuracy and in case of identification, it is used to improve the accuracy as well as the matching

speed. Distinctive features can be extracted by fusing different biometric modalities which results in increased recognition accuracy. However, the improvement in the accuracy comes at a cost. It is more expensive and may inconvenience the user since the user may have to interact with more than one sensor during both enrollment and matching stages. Fusion can also eliminate problems due to noisy data. Many fusion techniques are available to perform identification in multibiometric systems. Much research has been done on fusing different combinations of biometric modalities to improve the identification accuracy. Fusion using multiple modalities can be performed at various levels [71]:

- **Feature-level:** In the feature-level fusion, the feature vectors of individual modalities are combined into a single feature vector. Different approaches can be used to independently extract the features from each biometric modality into separate feature vectors which can be concatenated into a single feature vector. Feature reduction techniques can be used to reduce the higher dimensionality of the new concatenated feature vector. Feature-level fusion is preferred to image-level fusion because most automatic recognition systems store the features instead of the raw images because features are more compact than images. Wang et al., Rattani and Tistarelli [66, 91] are examples of fusing face and iris information at the feature level.

- **Similarity Score-level:** This fusion can also be referred to as confidence-level fusion or matching score fusion. The match scores of individual modalities are combined to generate a new match score in the score-level fusion. Each biometric system gives a similarity score of how similar the probe image is to the gallery image independently for the two modalities. The individual scores are normalized or mapped to a common domain before combining them for a more accurate matching decision. Score-level fusion is the most common type of fusion technique and gives the best trade-off between the ease of fusion and effectiveness of fusion. Wang et al., Morizet and Gilles [53, 89, 90] are examples of score-level fusion of face and iris modalities.

- **Rank-level:** In the rank-level fusion, the ranks of individual modalities are combined to derive a consensus rank for each identity. The output of individual modalities is a set of ranked gallery identities sorted in the descending order of their probabilities. These output ranks from individual modalities is consolidated to achieve a consensus rank for each identity. Rank-level fusion can only be used in identification systems.

- **Decision-level:** This fusion can also be referred to as abstract label-level fusion. In decision-level fusion, the individual modalities are fused based on the final recognition decision. In verification, the decision is of the form Yes/No and in identification, the decision can be of the form “identified" or “not-identified". Each modality is separately classified into a match or a non-match by the biometric system and the final decision is based on the majority voting scheme [29]

- **Image-level:** In this fusion, the individual modalities are combined at the image-level to get a single image which is used for matching. Image-level fusion can also be referred to as sample-level fusion, sensor-level fusion, signal-level fusion or measurement-level fusion. The image-level fusion occurs before the matching stage where the raw images are integrated to form a single combined image which then can be used for matching.

One challenges specific to fusion is that there is no optimal procedure or rule that can be used to apply the fusion techniques. These fusion techniques can be modality and dataset specific.

2.3 Challenges Of Multibiometric Systems

The challenges in multibiometric systems are [69]:

- **Loss Of Privacy:** Privacy concerns intensify with the use of multiple biometric modalities. There is a need for the feature vectors to be securely stored while retaining the recognition accuracy.

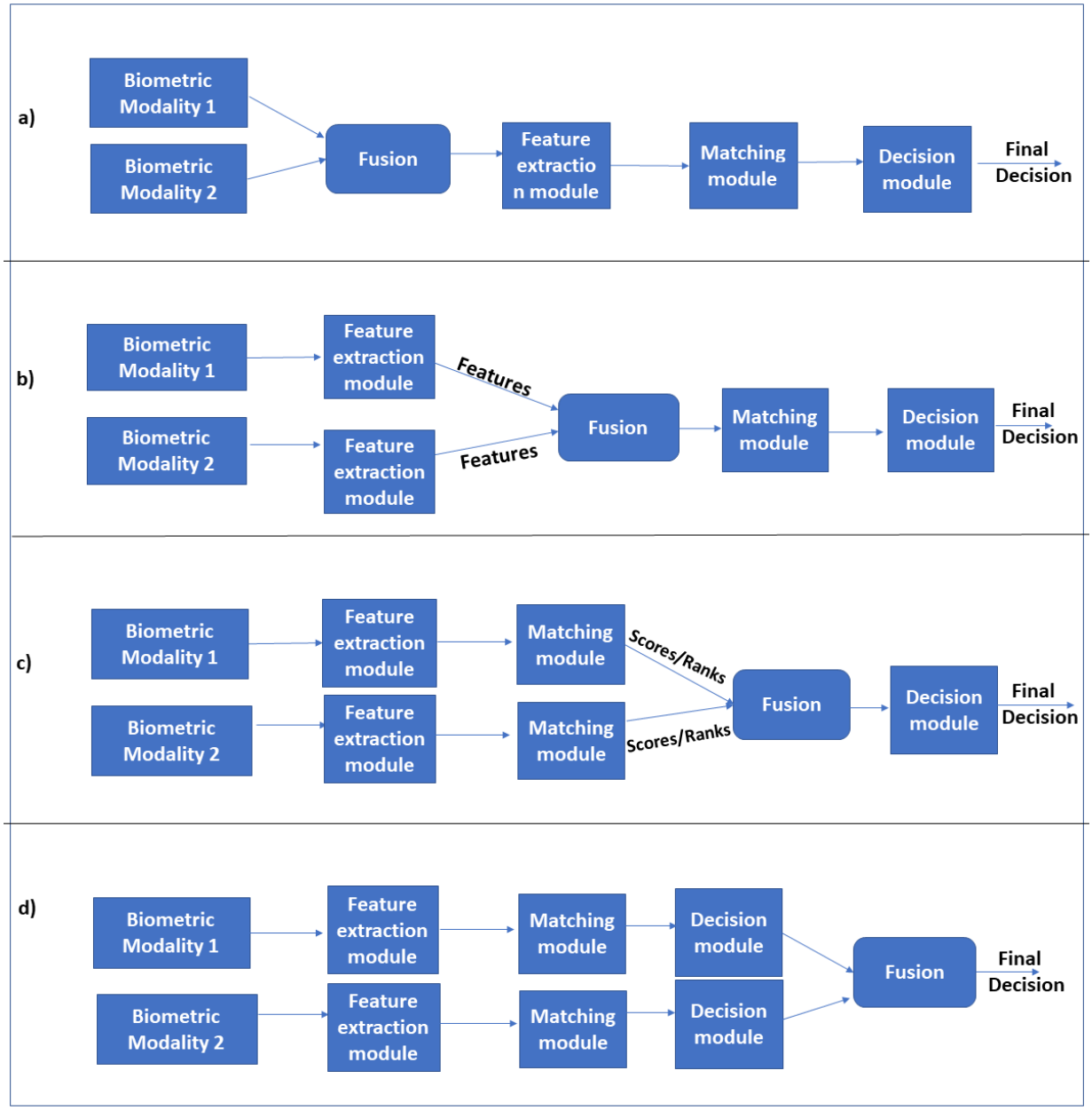

Figure 2.1: Levels of fusion: a) fusion at the image-level b) fusion at feature-level c) fusion at score-level or rank-level d) fusion at decision-level

- **Conflicts On Identity:** Sometimes individual modalities may give conflicting decisions on the identity of the user. In such cases, there should be a systematic way to make the final recognition decision.

- **Portability Issues:** Fusion algorithms have a number of tunable parameters. Transferring

these parameters across applications may be a difficult task. Robust fusion systems that can be ported across applications are necessary.

- **Changes To The Biometric Modality:** Aging and other factors can lead to modification of the biometric modality over time. This can lead to misuse of identity or spoof attack. The database should be designed so that it is adaptive to these changes and there should be an efficient way to update the feature vectors stored in the database.

- **Sensor Placement:** Since multiple sensors are being used, the sensors should be placed in such a way that it captures images of good quality for recognition and at the same time cause minimum inconvenience to the user during the image acquisition stage.

2.4 Privacy In Biometrics

2.4.1 Privacy Concerns In Biometrics

Biometric systems, like other systems, are vulnerable to attacks from hackers which can lead to compromise of biometric data. Any such compromise is catastrophic, as it leads to loss of user's identity since the biometric data cannot be easily replaced or reissued. Also, if the biometric data is compromised in one application, there is a possibility that all applications that use that biometric data can be compromised using the same method and the user can be tracked if multiple organizations share the same database.

2.4.2 Privacy Techniques In Biometrics

2.4.2.1 Biometric Template

To ensure privacy, instead of storing the raw biometric images in the database directly, most biometric systems store the data in the form of a compact digital representation called a template. Since the template contains only a compact description of the original data, the biometric template was considered as a one-way hash from which the original biometric image could not be retrieved.

Also, since the templates are encrypted, it is difficult to determine the contents of the template without the knowledge of de-crypting keys. However, studies have shown that, in spite of having an encryption and compact representation, it is possible to regenerate an estimate of raw biometric images from the stored template with a hill climbing attack [4] [70]

2.4.2.2 Hash Functions

Encryption techniques or hash functions can be used to enhance the privacy of the biometric template. Hash functions are one way algorithms and they are almost impossible to invert. However, these functions produce different hashes with even the slightest change in the input. As we have seen in the Automatic Face and Iris recognition sections, these modalities are subject to several changes due to illumination, resolution, occlusion, facial hair, aging, pose and expression variations. Because of this, it is not possible to have the exactly the same input without any variations every time. While, the hash will be the best way to ensure privacy, it may never have a positive match. Therefore, the hash functions cannot be practically used to implement privacy in biometrics.

2.4.2.3 Cancelable Biometrics

Another method to impart privacy to the biometric template is cancelable biometrics [64]. In this approach, a transformation function is used to generate a biometric template and this function is modified to produce a new variant of the template whenever one variant is compromised. Thus, the method provides revocability. The transformation function can be selected to be non-invertible. The original data cannot be retrieved even if the transformation function is known. The process of cancelable biometric template recognition is shown in figure 2.2 [55] When the key of the transformation function is exposed, the cancelable biometric template can be compromised. In such a case, if the transformation is invertible, the original biometric can be reconstructed. If the transformation is non-invertible, then the original biometric template can be approximately recovered. However, studies [54] [49] have shown that, by inverting the transformation function, the original biometric template may be restored by dictionary attacks.

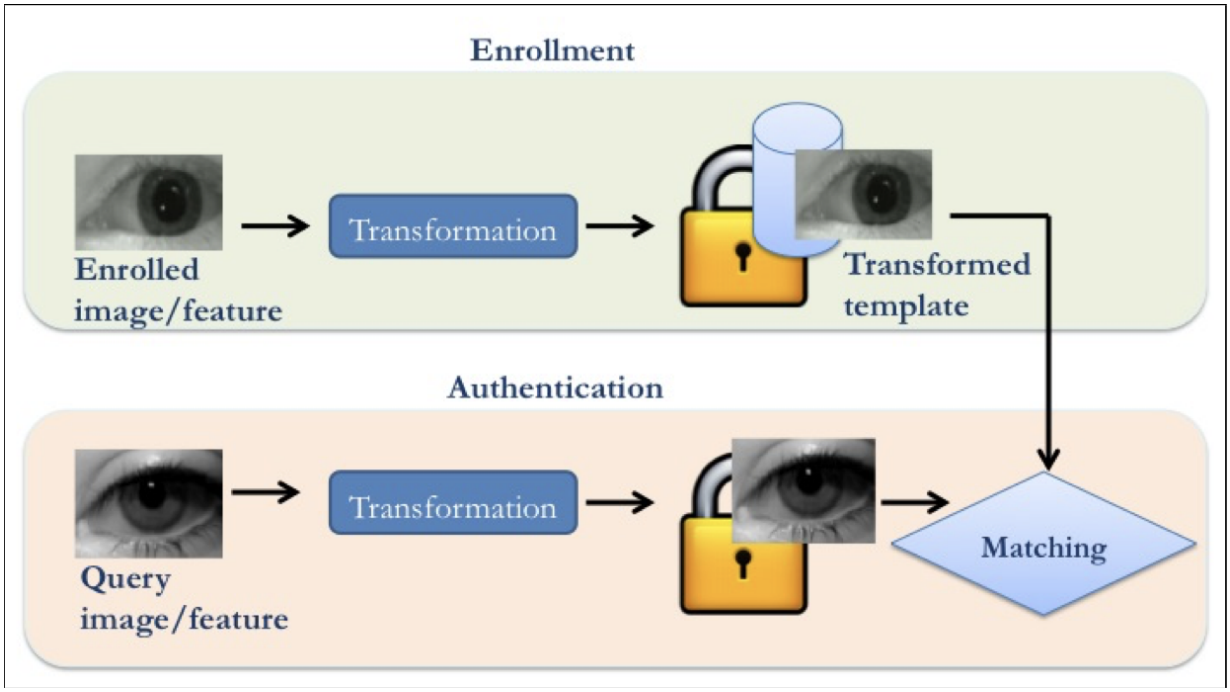

Figure 2.2: Biometric recognition using cancelable biometric template taken from [55]

CHAPTER 3

PROPOSED WORK

3.1 Method

In this work, we ensure that the privacy is preserved by combining two biometric modalities at the feature-level to generate a new biometric image that contains the features of both the modalities. This new biometric image that is generated cannot be used to match with either of the two original biometric modalities which means that, even if the database is hacked or the biometric template is compromised, attackers will not be able to retrieve the original face or iris image. In this way, the privacy is preserved.

The two biometric modalities that we consider for this study are face and iris. We extract the features from the face and iris modalities separately and fuse them into a combined representation. This combined representation is given as the input to a Generative Adversarial Network - an unsupervised neural network, to generate synthetic biometric face images. Instead of storing the original input face and iris images, the generated face image is stored in the database during the enrollment phase. During the verification phase, the user presents his face and iris biometrics and the system generates a synthetic face by the same process used in enrollment state and compares this to the stored synthetic template of the same user to generate a match score. Based on the match score and threshold, a match or a non-match is determined.

3.2 Architecture Used

We use a pre-trained ResNet [25] model to extract the features from the face and iris images. We then use the PCA technique on the feature vectors individually, to reduce its dimensionality. A DCGAN [62] network is used to generate a synthetic face image. The architecture of the DCGAN is shown in figure 3.1.

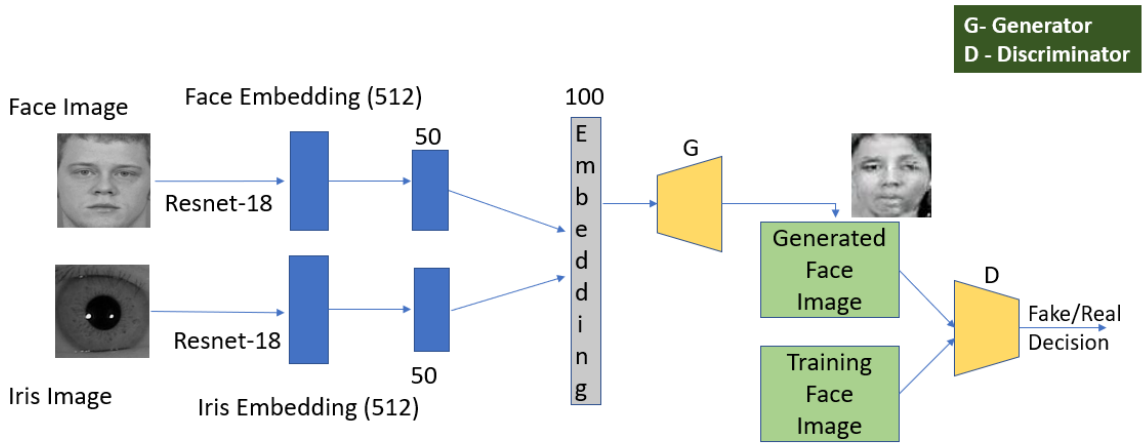

Figure 3.1: Proposed Method

3.2.1 ResNet-18

ResNet-18, also known as Residual Network, is a popular deep neural network having 18 layers with skip connections for image recognition. We extract a 512 dimensional embedding of the face and the iris image from the the last second layer of the network.

3.2.2 DCGAN

Deep Convolutional Generative Adversarial Network, or DCGAN for short, is a variant of Generative Adversarial Network (GAN) [22], which is an unsupervised learning task in machine learning which automatically learns patterns in the input data. There are two models in the GAN architecture, a generator and a discriminator which are trained simultaneously. The generator is responsible for producing synthetic or fake image and the discriminator determines whether that image is real image (from the input dataset) or a synthetic image (generated image). During training, the generator takes a random fixed-length vector as an input and generates a sample in the domain. The discriminator acts as a typical binary classification network. The two models are always competing against each other and are adversarial in that sense, and are playing a zero-sum game. DCGAN has certain architectural constraints on the network that make it an effective unsupervised learning network.

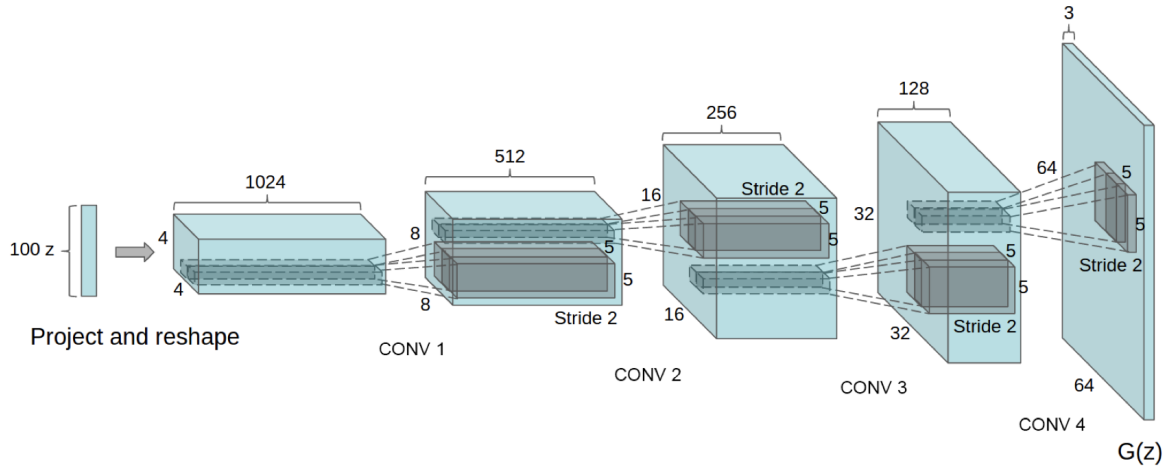

Figure 3.2: Architecture of DCGAN taken from [62]

3.3 Training Method

The face dataset is divided into a train and a test set. The DCGAN is trained only on the face images of the training set. During the testing stage, the feature vectors of the testing set of face images, and iris images are extracted using a pre-trained ResNet architecture. They are combined to form a 100 dimensional feature vector. This is given as input to the trained DCGAN model which generates new face images.

CHAPTER 4

EXPERIMENTS AND RESULTS

4.1 Datasets Used

Two multimodal datasets were used for this method. WVU and BioCOP 2008 are the multimodal datasets which have both face and iris modalities of the same user. A virtual dataset was constructed using face images from BioCOP 2008 dataset and iris images from CASIA-Iris Thousand dataset.

For the baseline experiments, the following are the statistics of the datasets that have been considered:

4.1.1 WVU Dataset

In the WVU face dataset, there are 269 subjects and the number of probe images are 1358 and the number of gallery images are 240. The WVU left iris dataset consisted of 241 subjects and 1321 probe images and 241 gallery images. The WVU right iris dataset consisted of 236 subjects and 1301 probe images and 236 gallery images.

4.1.2 BioCOP 2008 Dataset

The BioCOP 2008 face dataset consisted of 1135 subjects and 1128 probe images and 1135 gallery images.

4.1.3 CASIA-Iris Thousand Dataset

This is an iris only dataset which consisted of 1000 subjects and each subject has 10 samples for both left and right iris. The number of probe images are 2000 and the number of gallery images are 1000 for both the left and right iris.

4.1.4 Virtual Dataset

In this setup, we map one subject of the BioCOP face with one subject of the iris in a chimeric manner. 1000 subjects are considered for face images from the BioCOP 2008 dataset and each subject has 2 samples. 1000 subjects are considered for both the left and right iris and each subject has 10 samples each. 1000 face images are used for fine-tuning the DCGAN and during testing, 10000 face-iris images are used.

4.2 Experimental Setup

In these experiments, the left and right iris are considered as different biometric modalities. That is, feature-level fusion of face and left iris is different from feature-level fusion of face and right iris. All combinations of face and iris belonging to the same subject are concatenated to form combined face-iris images.

4.3 Results and Analysis

We present the matching performance for individual modalities of face and iris verification using commercial matchers and compare it with the matching performance of generated faces from fused biometric modality of face and iris.

4.3.1 Experiment 1: Baseline Experiments

The results of verification of individual face and iris modalities of WVU, BioCOP and CASIA-Iris thousand datasets are presented in table 4.1. For matching, the commercially available ROC Matcher [1] is used for face datasets and VeriEye Matcher [2] is used for iris datasets. Receiver operating characteristics of recognition of individual face and iris modalities of WVU, BioCOP and CASIA dataset are shown in figures 4.1, 4.2, 4.3, 4.4, 4.5, and 4.6. The commercially available matchers have high matching performance.

Table 4.1: True match rate of face and iris datasets at different values of false match rate for individual face and iris verification

Dataset	**TMR @ FMR=0.001**	**TMR @ FMR=0.01**
WVU Face	99.9%	100.0%
BioCOP Face	99.6%	99.6%
WVU Left Iris	99.5%	99.5%
WVU Right Iris	98.3%	98.4%
CASIA Left Iris	99.8%	99.8%
CASIA Right Iris	99.9%	100%

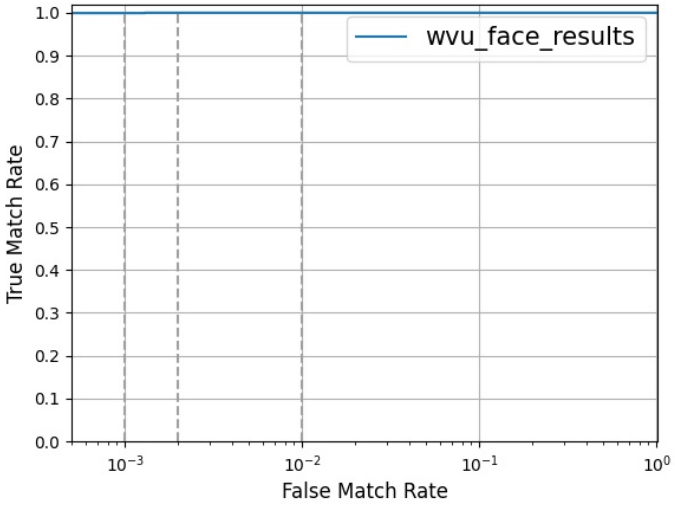

Figure 4.1: ROC of WVU face dataset

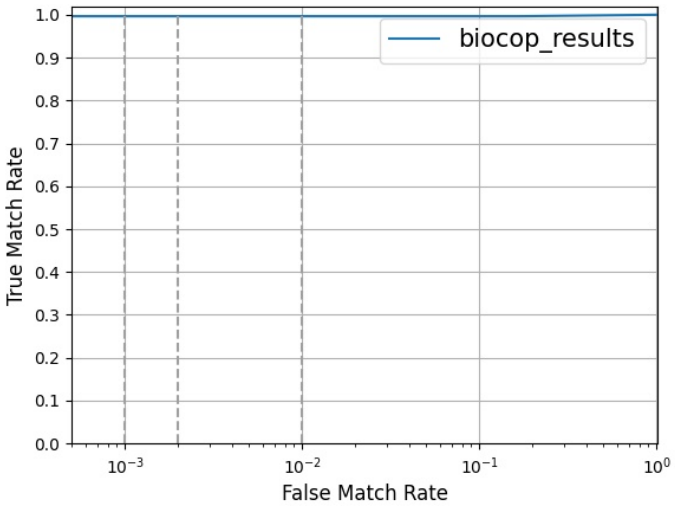

Figure 4.2: ROC of BioCOP 2008 face dataset

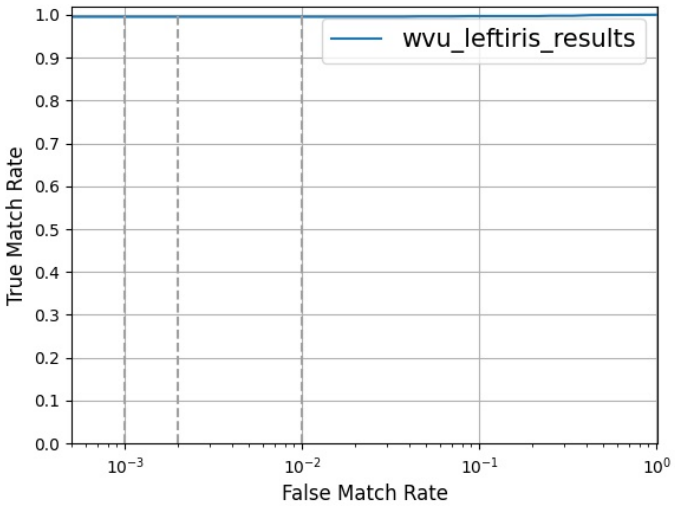

Figure 4.3: ROC of WVU left iris dataset

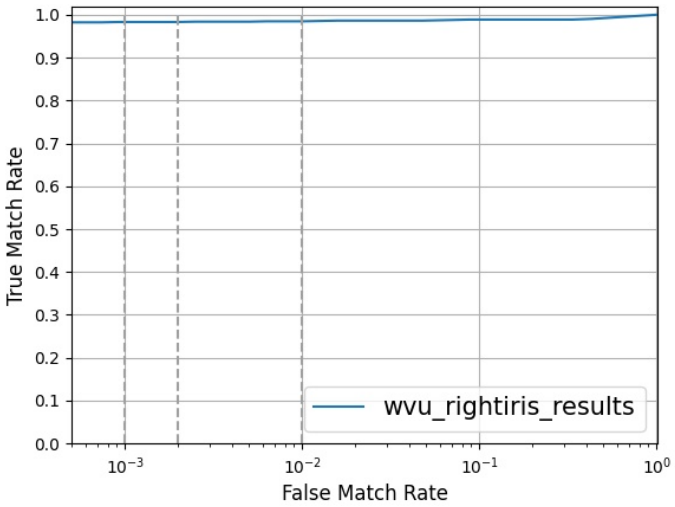

Figure 4.4: ROC of WVU right iris dataset

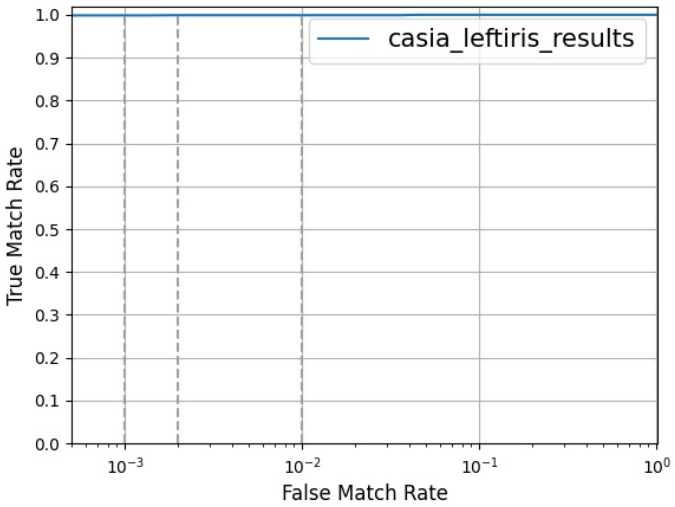

Figure 4.5: ROC of CASIA left iris dataset

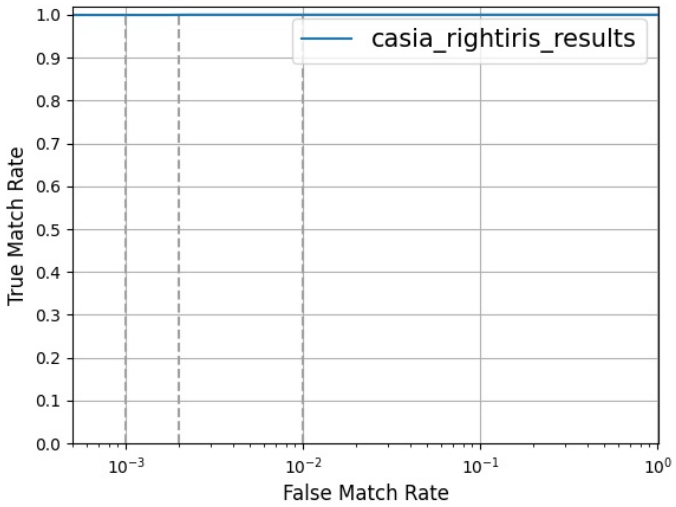

Figure 4.6: ROC of CASIA right iris dataset

Table 4.2: True match rate of face and iris datasets at different false match rate in verification mode

	BioCOP-CASIA Iris Left		**BioCOP-CASIA Iris Right**	
Experiment	**FMR=0.01**	**FMR=0.1**	**FMR=0.01**	**FMR=0.1**
Matching using fine-tuned StyleGAN	17.3%	40.9%	20.2%	47.0%
Matching using pre-trained DCGAN	48.5%	73.8%	53.7%	76.0%
Matching without PCA	28.7%	47.0%	16.7%	40.0%
Matching with PCA	55.6%	80.4%	59.1%	82.6%
Matching between generated and original face images	0.4%	9.0%	0.7%	9.0%
Matching of mismatched face-iris using DCGAN	49.4%	75.5%	53.4%	77%

4.3.2 Experiment 2: Generating Images Using Pre-trained StyleGAN Network

In this experiment, a StyleGAN network that is pre-trained on FFHQ dataset to generate images using the face-iris embedding. The images that are generated are shown in figure 4.7. The resulting images do not resemble face images and no useful information can be obtained from it so we do not perform any matching on these images.

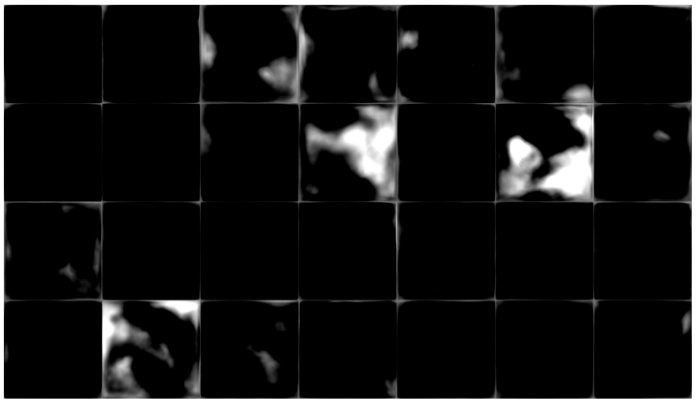

Figure 4.7: Examples of images generated by StyleGAN pre-trained on FFHQ dataset

4.3.3 Experiment 3: Generating Images Using StyleGAN Fine-tuned On BioCOP Faces

In this, the StyleGAN is fine-tuned on the BioCOP faces with the intention that it would generate images that look like faces. Some examples of generated face images are shown in figure 4.8. The generated images of all subjects are similar with some variations in the facial features. We conclude that the embedding space of the StyleGAN is larger than the embedding space of the faces. Matching is performed for the generated face images using ROC face matcher. The ROC curves for the generated face images from StyleGAN using ROC matcher for face-right iris and face-left iris are shown in figures 4.9 and 4.10, respectively.

Figure 4.8: Examples Of Images Generated By StyleGAN Fine-tuned On BioCOP Face Dataset

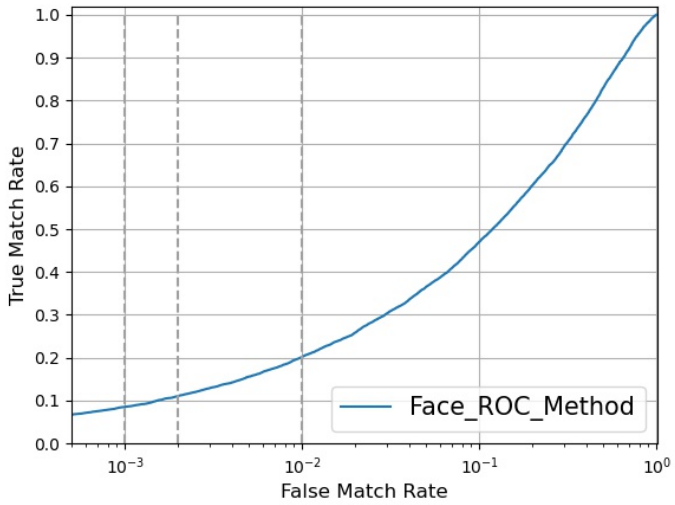

Figure 4.9: ROC of generated face images from face and right iris using StyleGAN fine-tuned on BioCOP faces

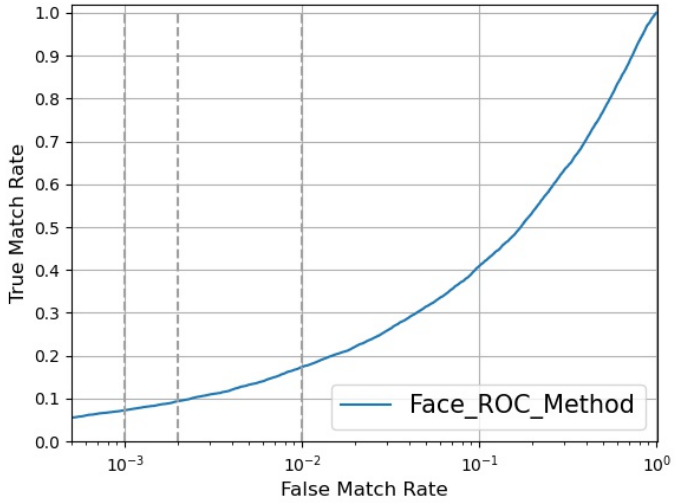

Figure 4.10: ROC of generated face images from face and left iris using StyleGAN fine-tuned on BioCOP faces

4.3.4 Experiment 4: Generating Images Using Pre-trained DCGAN Network

Next, a pre-trained DCGAN network that has been trained on Celeb-A dataset is used to generate face images. Some examples of generated face images are shown in figure 4.11. The images that

are generated appear distorted. Matching is performed for the generated face images using ROC face matcher and the matching results are higher than that of the StyleGAN matching results. The ROC curves for the generated face images from pre-trained DCGAN using the ROC matcher for face-right iris and face-left iris are shown in figures 4.12 and 4.13, respectively.

Figure 4.11: Examples of images generated by DCGAN pre-trained on Celeb-A dataset

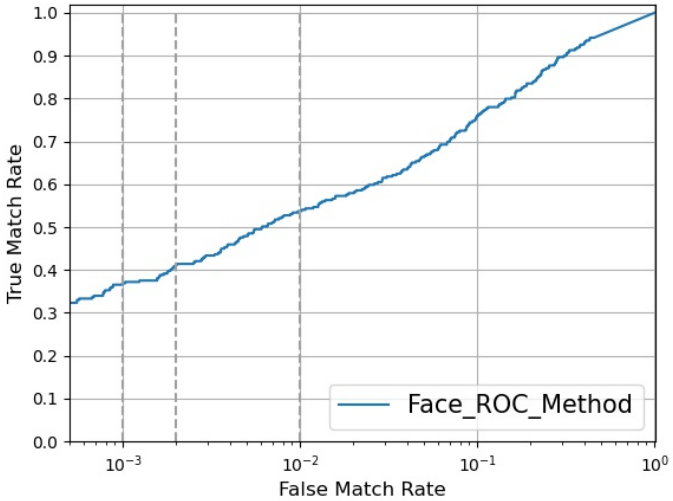

Figure 4.12: ROC of generated face images from face and right iris using pre-trained DCGAN

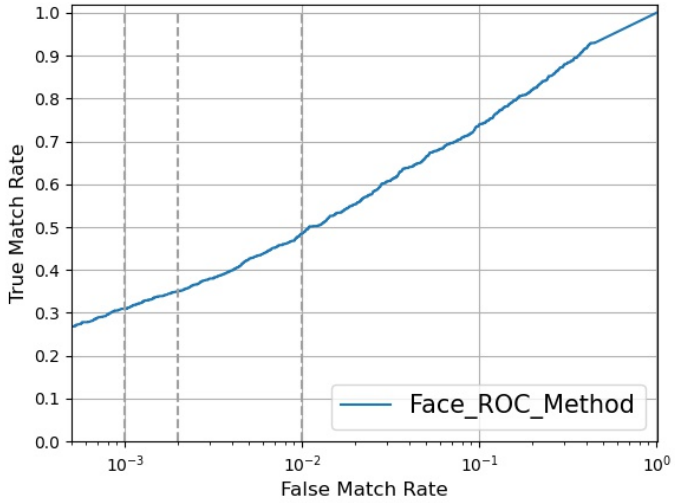

Figure 4.13: ROC of generated face images from face and left iris using pre-trained DCGAN

4.3.5 Experiment 5: Generating Images Using DCGAN Fine-tuned On BioCOP Faces

The DCGAN is fine-tuned on BioCOP faces so that it can learn to generate face images similar to that of BioCOP face. Some examples of generated face images using fine-tuned DCGAN are shown in figure 4.14. These faces are not distorted. The matching is performed for the generated face images using ROC face matcher and the matching results are higher than that of the matching results from pre-trained DCGAN. The ROC curves for the generated images using DCGAN fine-tuned on BioCOP faces for face-right iris and face-left iris are shown in figures 4.15 and 4.16, respectively.

Figure 4.14: Examples of images generated by DCGAN fine-tuned on BioCOP face dataset

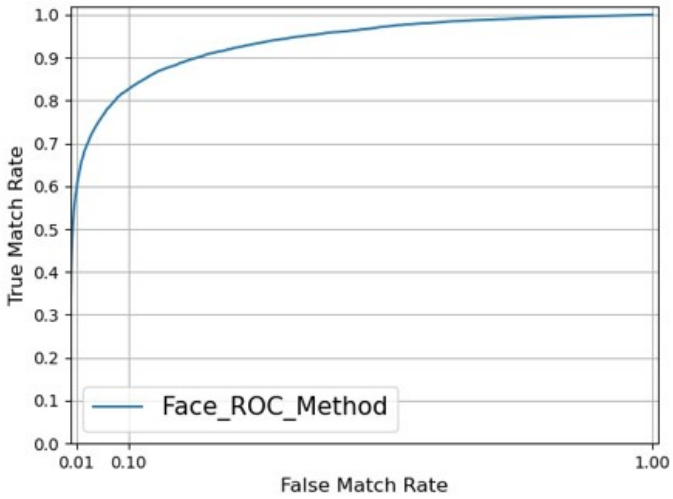

Figure 4.15: ROC of generated face images from face and right iris using DCGAN

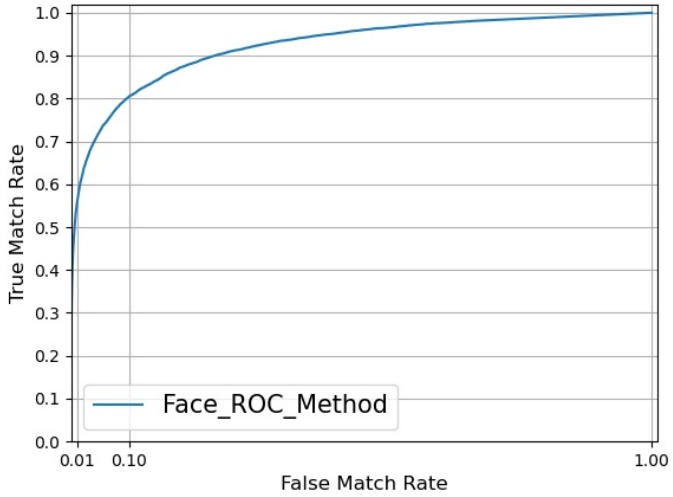

Figure 4.16: ROC of generated face images from face and left iris using DCGAN

4.3.6 Experiment 6: Matching Without Using PCA

In this experiment we reduce the dimentionality of the features in a trivial manner. The dimensionality of the feature vector of the face and iris that is obtained from the ResNet is 512 each. The dimensionality of each of these two feature vectors has to be reduced to form a 100 dimensional combined feature vector which can be given as the input to the DCGAN. The first 50 dimensions

of the face feature vector is selected to be combined with the first 50 dimensions of the iris feature vector to form a 100 dimensional face-iris feature vector. The results of the matching using the fine-tuned DCGAN when the face and iris is combined using this method is shown in table 4.2 - row 3 and ROC curves for this are shown in figures 4.15 and 4.16.

4.3.7 Experiment 7: Matching Using PCA

Principle Component Analysis (PCA), which is a popular feature extraction and dimensionality reduction technique in Machine Learning, is used to extract the top 50 dimensions of the face feature vector and the top 50 dimensions of the iris feature vector to form a 100 dimensional face-iris feature vector.The results of the matching using the fine-tuned DCGAN when the face and iris is combined after applying PCA is shown in table 4.2 - row 4. The results are much higher compared to combining the first 50 dimensions from both the modalities.

From table 4.2, we see that at FMR = 0.1, the performance significantly improves from 47% to 80% for face-left iris and from 40% to 82% for face-right iris when we apply PCA on the 512 dimensional embedding to extract a 100 dimensional embedding. We conclude that the 100 dimensional embedding of the face and iris contains the most discriminative features for the matching performance to be improved.

4.3.8 Experiment 8: Do The Original And Generated Face Images Match?

An experiment is conducted to check whether the generated face image matches with the original face image. The matching between them should be very low and the generated face image should not look like the original image, to ensure privacy. The results of this experiment are shown in table 4.2 - row 5 and the ROC curves are shown in figures 4.18 and 4.19. The matching accuracy between the original and generated face images is low indicating that original face image cannot be matched with the generated face image.

Using the proposed method, the generated face images cannot be matched with the original face images from which they are generated. In the figure 4.17, the match scores between the original

face image and generated face images are given. The match scores range from 0 to 1, where 0 means no match and 1 means they match perfectly. As we can see, the match scores are low, indicating that the original face image cannot be matched with the transformed face image, thus preserving the privacy. Also, visually, the generated face images look very different from the original face images.

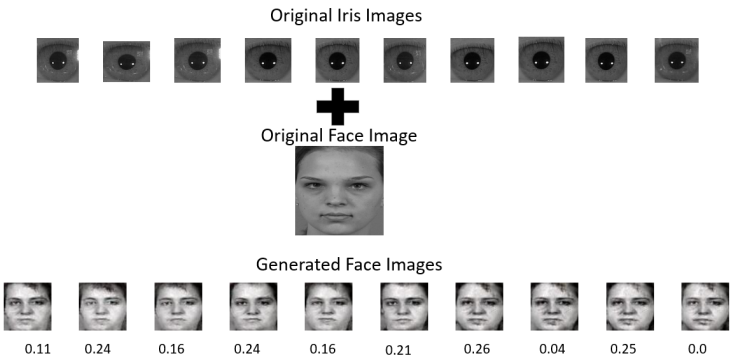

Figure 4.17: The match scores between the original face and the generated face images

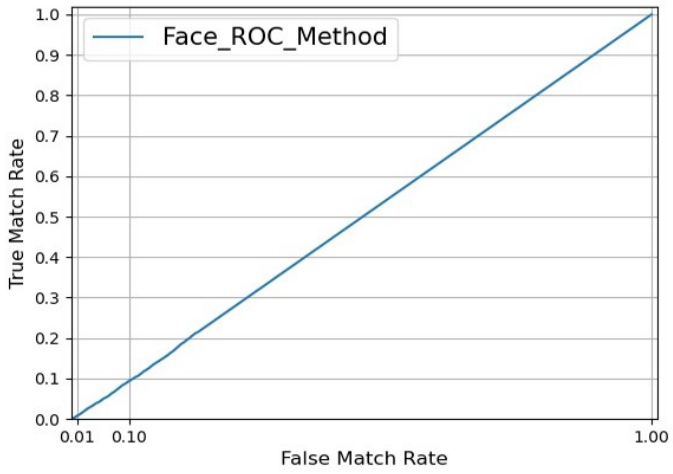

Figure 4.18: ROC curve of scores between the original face and the generated face images from face-left iris embedding

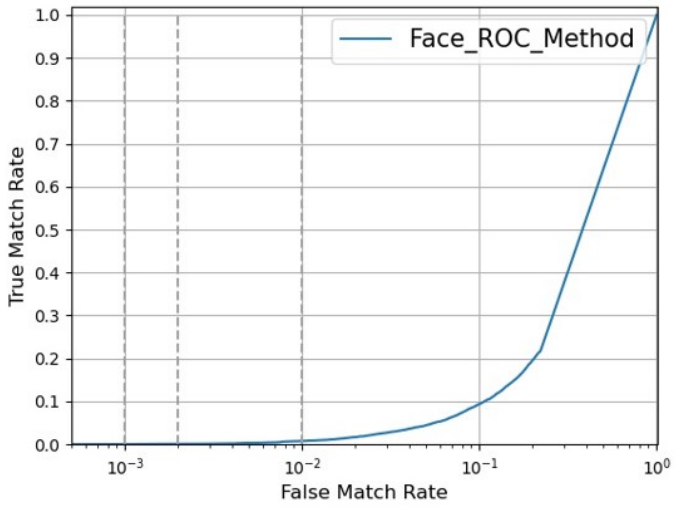

Figure 4.19: ROC curve of scores between the original face and the generated face images from face-right iris embedding

4.3.9 Experiment 9: Do The Generated Images Of Same Subject Look Similar?

In figure 4.20, the match scores between the the generated face images of the same subject are shown. The match scores are close to 1 indicating that there is high match between samples of the same subject and we can also visually notice that all the generated face images are similar to each other.

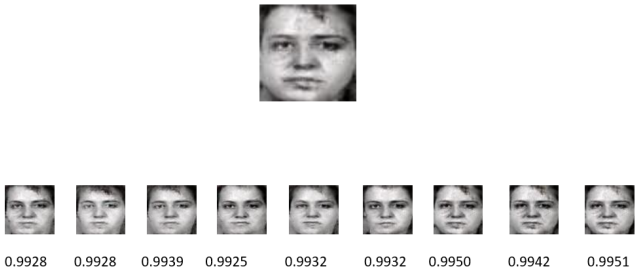

Figure 4.20: The match scores between the generated face images of the same subject

4.3.10 Experiment 10: How Is The Matching Performance When Face And Iris Of Different Subjects Are Combined?

In this experiment, we want to determine the matching performance is when the face and iris of different subjects are combined (that is the face image is paired with an iris image that is different from the previous set of experiments). Face embedding of one subject is combined with an iris embedding of another subject to get a mismatched combined face-iris embedding. These embeddings are then used to generate face images. The ROC face matcher is used to perform matching. The results are shown in table 4.2 - row 6. The performance for both left and right iris is lower than the DCGAN fine-tuned on the correct pair of subjects (refer row 4) but it is not possible to conclude that generated images will not match if the face and iris are of different identities.

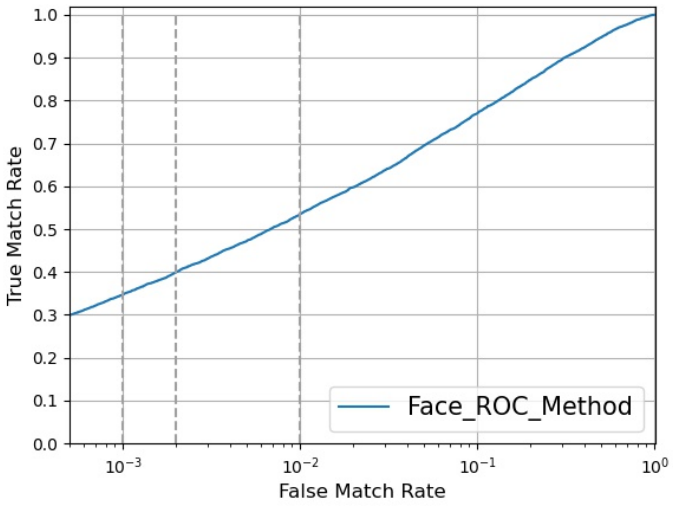

Figure 4.21: ROC curve of generated face images of mismatched face and right iris using DCGAN

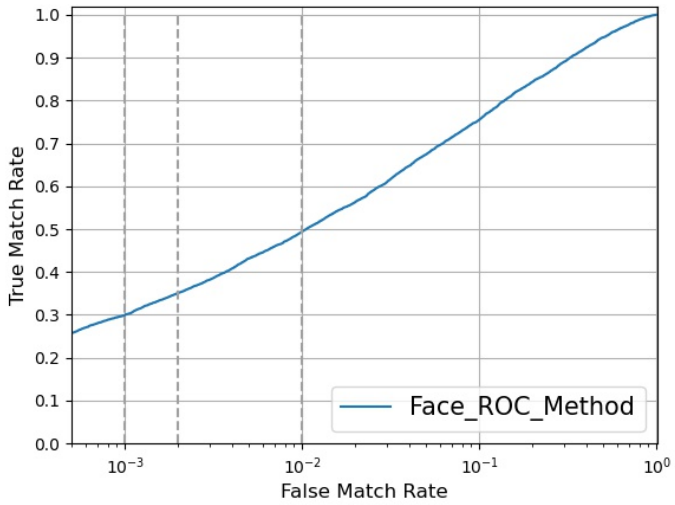

Figure 4.22: ROC curve of generated face images of mismatched face and left iris using DCGAN

4.3.11 Experiment 11: How Is The Quality Of Images Generated?

The software used to conduct the Image Quality Experiments is Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) in python [52], which is a natural scene statistic-based distortion-generic blind or no-reference (NR) image quality assessment (IQA) model that operates in the spatial domain. This method uses scene statistics of locally normalized luminance coefficients to quantify possible losses of "naturalness" in the image due to the presence of distortions, thereby leading to a holistic measure of quality. The results for the image quality are shown in figure 4.23. From the figure, we see the image quality scores in the range between 5 to 45. Using BRISQUE, the quality can be measured in the range between 1 to 100 where 1 indicates image of high quality and 100 indicates image of poor quality. We observe that the generated images from the DCGAN are of good quality however, there is a scope for improving the quality of the images significantly in the future.

Figure 4.23: Image quality score distribution

4.3.12 Experiment 12: Face-Iris Identification Experiments

The CMC curves of the face-right iris and face-left iris experiments are plotted and shown in figures 4.24 and 4.25, respectively.

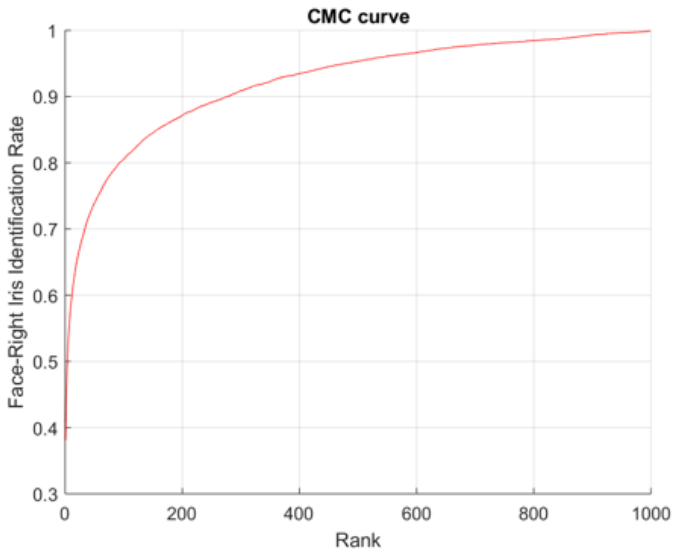

Figure 4.24: CMC curve of face-right iris

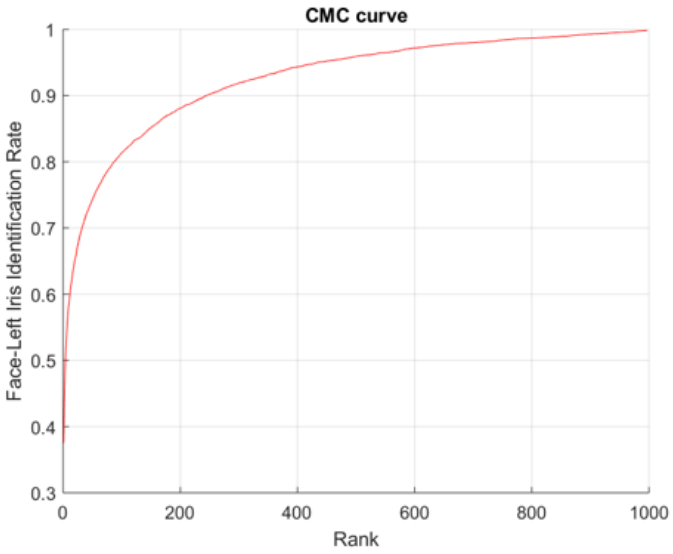

Figure 4.25: CMC curve of face-left iris

CHAPTER 5

CONCLUSION

5.1 Conclusion and Future Work

We have designed a new approach and conducted an exploratory analysis to combine face and iris modalities at the feature level to enhance the privacy in the biometric template while retaining the recognition accuracy. The results and analysis are presented. Since the biometric data is transformed by two layers of abstraction, it is difficult to retrieve the original data. The first abstraction is to combine the face and iris into a single embedding and the second layer of abstraction is to generate a face using this face-iris embedding making it difficult to extract information of the original image using the generated image. Although, to some extent, we can guarantee the privacy using this method, there is a significant drop in the recognition accuracy compared to traditional unimodal matching. In the future, there is much more to be done in investigating this approach such as other generative networks, dimensionality reduction techniques, etc. to increase this recognition accuracy to be comparable to that of individual face and iris recognition.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Roc face matcher. https://rankone.io/products/roc-sdk/. [Online: accessed 17-July-2021].

[2] VeriEye iris matcher. https://www.fulcrumbiometrics.com/Iris-Matcher-License-p/100424.htm. [Online: accessed 13-December-2018].

[3] M. Adam, F. Rossant, B. Mikovicova, and F. Amiel. Iris identification based on a local analysis of the iris texture. *Proceedings of 6th International Symposium on Image and Signal Processing and Analysis*, pages 523–528, Sept 2009.

[4] A. Adler. Images can be regenerated from quantized biometric match score data. *Canadian Conference on Electrical and Computer Engineering*, 1(6):469–472, 2004.

[5] F. S. Al-Qunaieer and L. Ghouti. Color iris recognition using hypercomplex gabor wavelets. *Symposium on Bio-inspired Learning and Intelligent Systems for Security*, pages 18–19, Aug 2009.

[6] Y. Ban, S.-K. Kim, S. Kim, and K.-A. Toh. Face detection based on skincolor likelihood. *Fourth IEEE International Conference on Biometrics: Theory Applications and Systems*, 47(4):1573–1585, 2014.

[7] Ankan Bansal, Carlos Castillo, Rajeev Ranjan, and Rama Chellappa. The do's and don'ts for cnn-based face verification. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV) Workshops*, Oct 2017.

[8] A. Bastys, J. Kranauskas, and R. Masiulis. Iris matching by local extremum points of multiscale taylor expansion. *Advances in Biometrics: Lecture Notes in Computer Science*, 5558:1070–1079, 2009.

[9] R. M. Bodade and S. N. Talbar. Shift invariant iris feature extraction using rotated complex wavelet and complex wavelet for iris recognition system. *Seventh International Conference on Advances in Pattern Recognition*, pages 449–452, Feb 2009.

[10] V. N. Boddeti and B. V. K. V. Kumar. Extended-depth-of-field iris recognition using unrestored wavefront-coded imagery. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 40(3):495–508, May 2010.

[11] R. Chen, X. Lin, T. Ding, and J. Ma. Accurate and fast iris segmentation applied to portable image capture device. *IEEE International Workshop on Imaging Systems and Techniques*, pages 80–84, May 2009.

[12] W.-S. Chen, C.-A. Chuan, S.-W. Shih, and S.-H. Chang. Iris recognition using 2d-lda + 2d-pca. *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 869–872, April 2009.

[13] Y. Chen, M. Adjouadi, A. Barreto, N. Rishe, and J. Andrian. A computational efficient iris extraction approach in unconstrained environments. *IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems*, pages 1–7, Sept 2009.

[14] Y. Chen, M. Adjouadi, C. Han, and A. Barreto. A new unconstrained iris image analysis and segmentation method in biometrics. *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 13–16, June 2009.

[15] C. Chou, S. Shih, W. Chen, V. W. Cheng, and D. Chen. Non-orthogonal view iris recognition system. *IEEE Transactions on Circuits and Systems for Video Technology*, 20(3):417–430, March 2010.

[16] S. S. Chowhan and G. N. Shinde. Evaluation of statistical feature encoding techniques on iris images. *WRI World Congress on Computer Science and Information Engineering*, 7:71–75, March 2009.

[17] D D. Chen, X Cao, F Wen, and J Sun. Blessing of dimensionality: High-dimensional feature and its efficient compression for face verification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3025–3032, Oct 2013.

[18] W. Dong, Z. Sun, and T. Tan. A design of iris recognition system at a distance. *Chinese Conference on Pattern Recognition*, pages 1–5, Nov 2009.

[19] Y. Du, N. L. Thomas, and E. Arslanturk. Multi-level iris video image thresholding. *IEEE Workshop on Computational Intelligence in Biometrics: Theory, Algorithms, and Applications*, pages 38–45, March 2009.

[20] R. Garg, V. Gupta, and V. Agrawal. Efficient iris recognition method for identification. *International Conference on Ultra Modern Telecommunications and Workshops*, pages 1–6, Oct 2009.

[21] J.E. Gentile, N. Ratha, and J. Connell. An efficient, two-stage iris recognition system. *IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems*, Sept 2009.

[22] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.

[23] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 3(6):610–621, Nov 1973.

[24] K He, X Zhang, S Ren, and J Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[25] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[26] Z. He, T. Tan, Z. Sun, and X. Qiu. Toward accurate and fast iris segmentation for iris biometrics. *IEEE transactions on pattern analysis and machine intelligence*, 31(9):1670–1684, September 2009.

[27] H.Proença. Iris recognition: On the segmentation of degraded images acquired in the visible wavelength. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(8):1502–1516, August 2009.

[28] A. Jain, A. Ross, and S. Prabhakar. Fingerprint matching using minutiae and texture features. In *Proceedings 2001 International Conference on Image Processing (Cat. No.01CH37205)*, volume 3, pages 282–285 vol.3, 2001.

[29] A. K. Jain, S. Prabhakar, and S. Chen. Combining multiple matchers for a high security fingerprint verification system. pages 1371–1379, 1999.

[30] A. K. Jain, A. Ross, and K. Nandakumar. Introduction to biometrics. *Springer*.

[31] A. K. Jain, A. Ross, and K. Nandakumar. Introduction to biometrics. *Springer*.

[32] A. K. Jain, A. Ross, and S. Prabhakar. An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1):4–20, Jan 2004.

[33] N. Kalka, N. Bartlow, and B. Cukic. An automated method for predicting iris segmentation failures. *IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems*, pages 1–8, Sept 2009.

[34] R. Kannavara and N. Bourbakis. Iris biometric authentication based on local global graphs: An fpga implementation. *IEEE Symposium on Computational Intelligence for Security and Defense Applications*, pages 1–7, July 2009.

[35] Mohd. Tariq Khan, Deepak Arora, and Shashwat Shukla. Feature extraction through iris images using 1-d gabor filter on different iris datasets. In *2013 Sixth International Conference on Contemporary Computing (IC3)*, pages 445–450, 2013.

[36] R. Kheirolahy, H. Ebrahimnezhad, and M. H. Sedaaghi. Robust pupil boundary detection by optimized color mapping for iris recognition. *14th International CSI Computer Conference*, pages 170–175, Oct 2009.

[37] E. Krichen, S. Garcia-Salicetti, and B. Dorizzi. A new phase-correlation-based iris matching for degraded images. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 39(4):924–934, Aug 2009.

[38] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, pages 1097–1105, 2012.

[39] K.Roy and P.Bhattacharya. Iris recognition in nonideal situations. *Information Security: Lecture Notes in Computer Science*, 5735:143–150, 2009.

[40] K. S. S. Kyaw. Iris recognition system using statistical features for biometric identification. *International Conference on Electronic Computer Technology*, pages 554–556, Feb 2009.

[41] R. D. Labati, V. Piuri, and F.Scotti. Agent-based image iris segmentation and multiple views boundary refining. *IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems*, pages 1–7, Sept 2009.

[42] R. D. Labati, V. Piuri, and F. Scotti. Neural-based iterative approach for iris detection in iris recognition systems. *IEEE Symposium on Computational Intelligence for Security and Defense Applications*, pages 1–6, July 2009.

[43] Y. Lee, P.J. Phillips, and R.J. Micheals. An automated video-based system for iris recognition. *Advances in Biometrics : Lecture Notes in Computer Science*, 5558:1160–1169, 2009.

[44] W. Li, P. Fu, and L. Zhou. Face recognition method based on dynamic threshold local binary pattern. *In Proceedings of the 4th International Conference on Internet Multimedia Computing and Service*, 2012.

[45] Y. Li and M.Savvides. A pixel-wise, learning-based approach for occlusion estimation of iris images in polar domain. *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 1357–1360, April 2009.

[46] Y. Li and M. Savvides. Automatic iris mask refinement for high performance iris recognition. *IEEE Workshop on Computational Intelligence in Biometrics: Theory, Algorithms, and Applications*, pages 52–58, March 2009.

[47] J. Lin, J.-P. Li, H. Lin, and J. Ming. Robust person identification with face and iris by modified pum method. *International Conference on Apperceiving Computing and Intelligence Analysis*, pages 321–324, Oct 2009.

[48] X. Liu, P. Li, and Q. Song. Eyelid localization in iris images captured in less constrained environment. *Advances in Biometrics :Lecture Notes in Computer Science*, 5558:1140–1149, 2009.

[49] Andy Luong, Michael Gerbush, Brent Waters, and Kristen Grauman. Reconstructing a fragmented face from a cryptographic identification protocol. In *2013 IEEE Workshop on Applications of Computer Vision (WACV)*, pages 238–245, 2013.

[50] S. McCloskey, W. Au, and J. Jelinek. Iris capture from moving subjects using a fluttering shutter. *Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems*, pages 1–6, Sept 2010.

[51] H. Mehrotra, B.G. Srinivas, B. Majhi, and P. Gupta. Indexing iris biometric database using energy histogram of dct subbands. *Contemporary Computing*, 40(4):194–204, 2009.

[52] Anish Mittal, Anush Krishna Moorthy, and Alan Conrad Bovik. No-reference image quality assessment in the spatial domain. *IEEE Transactions on Image Processing*, 21(12):4695–4708, 2012.

[53] N. Morizet and J. Gilles. A new adaptive combination approach to score level fusion for face and iris biometrics combining wavelets and statistical moments. *Advances in Visual Computing: Lecture Notes in Computer Science*, 5359:661–671, 2008.

[54] Abhishek Nagar, Karthik Nandakumar, and Anil K. Jain. Biometric template transformation: a security analysis. In Nasir D. Memon, Jana Dittmann, Adnan M. Alattar, and Edward J. Delp III, editors, *Media Forensics and Security II*, volume 7541, pages 237 – 251. International Society for Optics and Photonics, SPIE, 2010.

[55] Vishal M. Patel, Nalini K. Ratha, and Rama Chellappa. Cancelable biometrics: A review. *IEEE Signal Processing Magazine*, 32(5):54–65, 2015.

[56] C. M. Patil and S. Patilkulkarani. Iris feature extraction for personal identification using lifting wavelet transform. *International Conference on Advances in Computing, Control, and Telecommunication Technologies*, pages 764–766, Dec 2009.

[57] C.M. Patil and S. Patilkulkarni. An approach to enhance security environment based on sift feature extraction and matching to iris recognition. *Information Processing and Management*, 70:527–530, 2010.

[58] H. Proença. Iris recognition: A method to segment visible wavelength iris images acquired on-the-move and at-a-distance. *Advances in Visual Computing: Lecture Notes in Computer Science*, 5358(8):731–742, 2008.

[59] H. Proença. Quality assessment of degraded iris images acquired in the visible wavelength. *IEEE Transactions on Information Forensics and Security*, 6(1):82–95, March 2011.

[60] S. J. Pundlik, D. L. Woodard, and S. T. Birchfield. Non-ideal iris segmentation using graph cuts. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pages 1–6, June 2008.

[61] M. Rachubinski. Iris identification using geometrical wavelets. *Computer Vision and Graphics*, 2009.

[62] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks, 2016.

[63] K.R. Radhika, S.V. Sheela, M.K. Venkatesha, and G.N. Sekhar. Multi-modal authentication using continuous dynamic programming. *Biometric ID Management and Multimodal Communication: Lecture Notes in Computer Science*, 5707:228–235, July 2009.

[64] N. Ratha, J. Connell, and R. Bolle. Enhancing security and privacy in biometrics-based authentication systems. *IBM Syst. J.*, 40:614–634, 2001.

[65] C. Rathgeb, A. Uhl, and P. Wild. Incremental iris recognition: A single-algorithm serial fusion strategy to optimize time complexity. *Fourth IEEE International Conference on Biometrics: Theory Applications and Systems*, Sept 2010.

[66] A. Rattani and M. Tistarelli. Robust multi-modal and multi-unit feature level fusion of face and iris biometrics. *Advances in Biometrics: Lecture Notes in Computer Science*, 5558:960–969, 2009.

[67] A. Ross, K. Nandakumar, and A. K. Jain. Handbook of multibiometrics. *Springer*.

[68] A. Ross, R. Pasula, and L. Hornak. Exploring multispectral iris recognition beyond 900nm. *IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems*, pages 1–8, Sept 2009.

[69] Arun Ross and Norman Poh. *Multibiometric Systems: Overview, Case Studies, and Open Issues*, pages 273–292. Springer London, London, 2009.

[70] Arun A. Ross, Jidnya Shah, and Anil K. Jain. Toward reconstructing fingerprints from minutiae points. *Proc. SPIE 5779, Biometric Technology for Human Identification II*, 2005.

[71] A. Rossn. An introduction to multibiometrics. In *Proceedings of the 15th European Signal Processing Conference*, September 2007.

[72] K. Roy and P. Bhattacharya. Improving features subset selection using genetic algorithms for iris recognition. *Improving Features Subset Selection Using Genetic Algorithms for Iris Recognition*, 5064:292–304, 2008.

[73] K. Roy and P. Bhattacharya. Level set approaches and adaptive asymmetrical svms applied for nonideal iris recognition. *Image Analysis and Recognition: Lecture Notes in Computer Science*, 5627:418–428, 2009.

[74] K. Roy and P. Bhattacharya. Nonideal iris recognition using level set approach and coalitional game theory. *Computer Vision Systems: Lecture Notes in Computer Science*, 5815:394–402, 2009.

[75] K. Roy and P.Bhattacharya. Optimal features subset selection using genetic algorithms for iris recognition. *Image Analysis and Recognition: Lecture Notes in Computer Science*, 5112:894–904, 2008.

[76] W. J. Ryan, D. L. Woodard, A. T. Duchowski, and S. T. Birchfield. Adapting starburst for elliptical iris segmentation. *IEEE Second International Conference on Biometrics: Theory, Applications and Systems*, pages 1–7, Sept 2008.

[77] Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.

[78] F. Scotti and V. Piuri. Adaptive reflection detection and location in iris biometric images by using computational intelligence techniques. *IEEE Transactions on Instrumentation and Measurement*, 59(7):1825–1833, July 2010.

[79] S. Shah and A. Ross. Iris segmentation using geodesic active contours. *IEEE Transactions on Information Forensics and Security*, 4(4):824–836, Dec 2009.

[80] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[81] Yi Sun, Ding Liang, Xiaogang Wang, and Xiaoou Tang. Deepid3: Face recognition with very deep neural networks. *CoRR*, abs/1502.00873, 2015.

[82] Zhenan Sun, Alessandra A. Paulino, Jianjiang Feng, Zhenhua Chai, Tieniu Tan, and Anil K. Jain. A study of multibiometric traits of identical twins. In B. V. K. Vijaya Kumar, Salil Prabhakar, and Arun A. Ross, editors, *Biometric Technology for Human Identification VII*, volume 7667, pages 283 – 294. International Society for Optics and Photonics, SPIE, 2010.

[83] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9, 2015.

[84] Y Taigman, M Yang, M Ranzato, and L Wolf. Deepface: Closing the gap to human-level performance in face verification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1701–1708, Oct 2014.

[85] Yaniv Taigman, Ming Yang, Marc'Aurelio Ranzato, and Lior Wolf. Deepface: Closing the gap to human-level performance in face verification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2014.

[86] N. Tajbakhsh and K. Misaghian N.M. Bandari. A region-based iris feature extraction method based on 2d-wavelet transform. *Biometric ID Management and Multimodal Communication. BioID 2009. Lecture Notes in Computer Science*, 5707:301–307, 2009.

[87] X. Tan and B. Triggs. Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE transactions on image processing*, 19(6):1635–1650, 2010.

[88] V. Velisavljevic. Low-complexity iris coding and recognition based on directionlets. *IEEE Transactions on Information Forensics and Security*, 4(3):410–417, Sept 2009.

[89] F. Wang and J. Han. Multimodal biometric authentication based on score level fusion using support vector machine. *Opto-Electronics Review*, 17:59–64, Mar 2009.

[90] J. Wang, Y. Li, X. Ao, C. Wang, and J. Zhou. Multi-modal biometric authentication fusing iris and palmprint based on gmm. *IEEE 15th Workshop on Statistical Signal Processing*, pages 349–352, Aug 2009.

[91] Z. Wang, Q. Han, X. Niu, and C. Busch. Feature-level fusion of iris and face for personal identification. *Advances in Neural Networks: Lecture Notes in Computer Science*, 5553:356–364, 2009.

[92] F. W. Wheeler, A. G. A. Perera, G. Abramovich, B. Yu, and P. H. Tu. Stand-off iris recognition system. *IEEE Second International Conference on Biometrics: Theory, Applications and Systems*, pages 1–7, Sept 2008.

[93] Jia Xiang and Gengming Zhu. Joint face detection and facial expression recognition with mtcnn. In *2017 4th International Conference on Information Science and Control Engineering (ICISCE)*, pages 424–427, 2017.

[94] Yu-Xin Yang, Chang Wen, Kai Xie, Fang-Qing Wen, Guan-Qun Sheng, and Xin-Gong Tang. Face recognition using the sr-cnn model. *Sensors*, 18(12), 2018.

[95] Dong Yi, Rong Liu, RuFeng Chu, Zhen Lei, and Stan Z. Li. Face matching between near infrared and visible light images. In Seong-Whan Lee and Stan Z. Li, editors, *Advances in Biometrics*, pages 523–530, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg.

[96] B. Zhang, Y. Gao, S. Zhao, and J. Liu. Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor. *IEEE transactions on image processing*, 19(2):533–544, 2009.

[97] X. Zhang, Q. Wang, H. Zhu, C. Yao, L. Gao, and X. Liu. Noise detection of iris image based on texture analysis. *Chinese Control and Decision Conference*, pages 2366–2370, June 2009.

[98] Z. Zhang, D. Yi, Z. Lei, and S. Z. Li. Regularized transfer boosting for face detection across spectrum. *IEEE Signal Processing Letters*, 19(3):131–134, March 2012.

[99] Z. Zhiping, H. Maomao, and S. Ziwen. An iris recognition method based on 2dwpca and neural network. *Chinese Control and Decision Conference*, pages 2357–2360, June 2009.