Spoof Detection on Ear Biometrics

İmren Toprak

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Computer Engineering

Eastern Mediterranean University
December 2019
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Prof. Dr. Ali Hakan Ulusoy
Acting Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Doctor of Philosophy in Computer Engineering.

Prof. Dr. Işık Aybay
Chair, Department of Computer Engineering

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Doctor of Philosophy in Computer Engineering.

Assoc. Prof. Dr. Önsen Toygar
Supervisor

Examining Committee

1. Prof. Dr. Fikret S. Gürgen
2. Prof. Dr. Cüneyt Güzeliş
3. Assoc. Prof. Dr. Duygu Çelik Ertuğrul
4. Assoc. Prof. Dr. Önsen Toygar
5. Asst. Prof. Dr. Yıltan Bitirim

ABSTRACT

Ear recognition systems are one of the biometrics-based popular person identification systems. The attacks to these biometrics identification systems become inevitable. In this context, the emerging ear recognition systems need to counter against spoof attacks. Consequently, ear anti-spoofing problem is focused in this thesis. In the biometric community, the types of attacks can be listed as; printed photo attack, display attack, replay attack, mask attack etc. In this thesis, printed photo attack is considered. Firstly, the Image Quality Assessment (IQA) methods are employed to find a solution to this problem. Therefore, 21 Full-Reference (FR) and 4 No-Reference (NR) IQA measures are implemented to extract features from ear images. In addition to this, Convolutional Neural Network (CNN) which is a deep learning method is implemented to detect impostor ear samples. Further, texture-based Binarized Statistical Image Features (BSIF) method is applied to represent the features of ear images. In this context, four different methods are proposed for distinguishing genuine ear images from impostor ones. The first one employs Decision-Level-Fusion (DLF) technique to combine FR and NR IQA measures. Secondly, three-level fusion of FR and NR IQA measures is implemented by using Score-Level-Fusion (SLF) and DLF techniques. Additionally, a CNN-based system and 5 IQA measures are combined by applying DLF technique as a third method. Finally, BSIF and CNN-based methods are fused by using DLF technique. The used databases for experiments are AMI, UBEAR, IITD, USTB Set 1, USTB Set 2 and USTB Set 3 which are publicly available. However, the spoof database for aforementioned databases is not available. Therefore, the spoof database which contains print attack is constructed in this thesis.

Keywords: Ear biometrics, Spoof detection, Image quality assessment, Deep learning, CNN-based method, Texture-based method, BSIF, Printed photo attack

ÖZ

Kulak tanıma sistemleri biyometri tabanlı popüler kişi tanıma sistemlerinden biridir. Biyometri tabanlı tanıma sistemlerine yapılan saldırılar kaçınılmaz hale gelmiştir. Bu bağlamda, ortaya çıkan kulak tanıma sistemlerinin yanıltma saldırılarına karşı mücadele etmesi gerekmektedir. Sonuç olarak, kulak yanıltma önleme problemi bu tezde ele alınmaktadır. Biyometrik sistemlere karşı saldırı türleri şöyle sıralanabilir: fotoğraf baskısı, görüntü saldırısı, tekrar görüntüleme saldırısı, maske saldırısı vb. Bu tezde, basılı fotoğraf saldırıları göz önünde bulundurulmuştur. İlk olarak, bu soruna bir çözüm bulmak için Görüntü Kalitesi Değerlendirmesi (IQA) yöntemi kullanılmıştır. Bu nedenle, kulak görüntülerinden öznitelikleri çıkarmak için 21 Tam Referanslı (FR) ve 4 Referanssız (NR) IQA yöntemi uygulanmıştır. Buna ek olarak, derin bir öğrenme yöntemi olan Konvolüsyonel Sinir Ağı (CNN), sahtekâr kulak örneklerini tespit etmek için uygulanmıştır. Ayrıca, doku bazlı yöntem olan İkili İstatistiksel Görüntü Öznitelikleri (BSIF), kulak görüntülerinin özniteliklerini temsil etmek için uygulanmıştır. Bu bağlamda, gerçek kulak görüntülerini sahtekâr olanlardan ayırmak için dört farklı yöntem önerilmiştir. İlki, FR ve NR IQA ölçümlerini birleştirmek için Karar Seviyesi Kaynaşım (DLF) tekniğini kullanır. İkinci olarak, FR ve NR IQA yöntemlerinin üç seviyeli kaynaşımı, Skor Seviyesi Kaynaşım (SLF) ve DLF teknikleri kullanılarak uygulanmıştır. Ayrıca, üçüncü bir yöntem olarak CNN ve 5 IQA yöntemi, DLF tekniği uygulanarak birleştirilmiştir. Son olarak, BSIF ve CNN tabanlı yöntemler DLF tekniği kullanılarak birleştirilmiştir. Deneyler için kullanılan veritabanları, kamuya açık olan AMI, UBEAR, IITD, USTB Set 1, USTB Set 2 ve USTB Set 3'tür. Bununla birlikte, yukarıda belirtilen veritabanları için sahte

veritabanı mevcut değildir. Bu nedenle, baskı saldırısı içeren sahte veritabanı bu tezde oluşturulmuştur.

Anahtar Kelimeler: Kulak biyometrisi, Yanıltma saptama, Görüntü kalitesi değerlendirmesi, Derin öğrenme, CNN tabanlı yöntem, Doku tabanlı yöntem, BSIF, Basılı fotoğraf saldırısı

ACKNOWLEDGEMENT

I would like to express my sincere thanks to my supervisor who is Assoc. Prof. Dr. Önsen TOYGAR for her guidance and for sharing her valuable knowledge and experiences with me in the process of completing this thesis.

I am grateful to my lovely family who encouraged me to do PhD study. It was priceless to feel the support of my family in this process.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AD	Average Difference
ALWGO	Aggregated Local Weighted Gradient Orientation
AOS	Additive Operator Splitting
BIQI	Blind Image Quality Index
BN	Batch Normalization
BSIF	Binarized Statistical Image Features
CLNF	Conditional Local Neural Fields
CNN	Convolutional Neural Network
CoALBP	Co-occurrence of Adjacent Local Binary Patterns
DBN	Deep Belief Network
DLF	Decision-Level Fusion
DOG	Difference of Gaussians
FFR	False Fake Rate
FGR	False Genuine Rate
FLF	Feature-Level Fusion
FR	Full-Reference
GLCM	Gray-Level Co-occurrence Matrix
GME	Gradient Magnitude Error
GPE	Gradient Phase Error
GW	Gabor Wavelets
HLFI	High-Low Frequency Index
HOG	Histogram of Oriented Gradients
HSV	Hue Saturation Value

HTER	Half Total Error Rate
ICA	Independent Component Analysis
IDA	Image Distortion Analysis
ISE	Image Scale Equalization
IQA	Image Quality Assessment
IQM	Image Quality Measure
JQI	JPEQ Quality Index
LBP	Local Binary Patterns
LDA	Linear Discriminant Analysis
LMSE	Laplacian Mean Squared Error
LPQ	Local Phase Quantization
MAS	Mean Angle Similarity
MAMS	Mean Angle Magnitude Similarity
M-BSIF	Multi-Scale Binarized Statistical Image Feature
MBSIF-TOP	Multiscale Dynamic Binarized Statistical Image Features
MC-CNN	Multi-Channel Convolutional Neural Network
MD	Maximum Difference
MLBP	Multi-Level Local Binary Patterns
MLPQ-TOP	Multiscale Dynamic Local Phase Quantization
MMD	Maximum Mean Discrepancy
MSE	Mean Squared Error
NAE	Normalized Absolute Error
NIQE	Natural Image Quality Evaluator
NN	Neural Networks
NR	No-Reference

NXC	Normalized Cross-Correlation
PCA	Principal Component Analysis
PSNR	Peak Signal to Noise Ratio
RAMD	R-Averaged Maximum Difference
RBM	Restricted Boltzmann Machine
ReLU	Rectified Linear Unit
RGB	Red Green Blue
RRED	Reduced Reference Entropic Difference
SC	Structural Content
SID	Scale-Invariant Descriptor
SLF	Score-Level Fusion
SME	Spectral Magnitude Error
SNR	Signal to Noise Ratio
SPE	Spectral Phase Error
SRC	Sparse Representation Classifier
SR-KDA	Spectral Regression Based Kernel Discriminant Analysis
SSIM	Structural Similarity Index Measure
SURF	Speed Up Robust Features
SVM	Support Vector Machine
TCD	Total Corner Difference
TED	Total Edge Difference
VIF	Visual Information Fidelity
WLD	Weber's Local Descriptor

Chapter 1

INTRODUCTION

In recent years, the identification systems based on biometric traits become inevitable for human life. The application of the person identification systems has wide range in the area of the biometric community. Consequently, person identification systems based on ear biometric trait attract the researchers' interest. Ear is one of the biometric traits that satisfies the requirements which are uniqueness, universality, permanence and collectability of recognition systems [1]. Since ear biometric trait has some advantages, it is also used in the implementation of recognition systems. Firstly, it has distinctive features even in the identical twins [3]. Consequently, it is understood that ear is a reliable biometric trait. The second advantage is that it does not require the cooperation of a person to be identified. Therefore, it can be used in surveillance or tracking applications. Furthermore, its structure is not affected because of the facial expressions or age variation. These advantages make the ear biometric trait useful in biometric recognition systems in terms of reliability and ease of data collection. Since ear biometric trait has the aforementioned advantages, it is a preferred biometric trait to be implemented in the person identification systems. Many studies have been done so far to propose novel and robust ear recognition systems [1–12].

Fooling the identification systems which are based on biometric by using impostor biometric trait is called spoofing. In the biometric community, the research area called spoof detection on biometric based identification systems is emerged to differentiate

genuine biometric trait from the impostor biometric trait [13]. The most common attack types are printed photo attack, digital photo attack, replay video attack, mask attack and plastic surgery attack [14-16]. Specifically, live ear image, printed photo attack of the ear image and digital photo attack of the ear image are demonstrated in Figure 1 (a), (b) and (c), respectively.

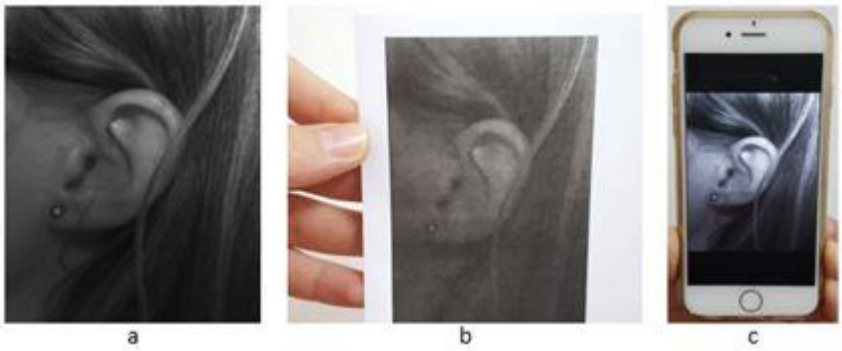

Figure 1. Type of spoof attacks (a) live ear image (b) printed photo attack (c) digital photo attack

Anti-spoofing techniques which are implemented to detect spoof attacks are divided into three categories. The first category includes sensor-level techniques that are hardware-based. In that category, specific devices are added to the sensor to measure face thermogram, blood pressure, fingerprint sweat, or reflection of eye of the biometric trait to decide whether the biometric trait is live or not. The second category includes software-based methods. In that category, applications are implemented at feature-level of biometric based identification systems. The third category includes score-level techniques. These are focused on the study of biometric systems at score-level in order to propose fusion strategies that increase their resistance against spoofing attempts [17].

Many anti-spoofing methods have been proposed by the researchers to overcome the spoof attacks against biometric recognition systems such as face, iris, fingerprint,

palmprint, finger vein, palm vein and voice. Generally, these methods are based on deep learning, texture, motion, image quality and liveness [15, 18-25].

However, there is no anti-spoofing method implemented for ear recognition systems in the biometric community. In this thesis, our aim is to propose a novel anti-spoofing method for ear identification systems [26, 27]. In this context, image quality based, deep neural network based and texture based methods are employed to find the best solution for spoofing problem on ear recognition systems.

The image quality is one of the factors that shows the differences between the real and fake images. The quality of the real image which is acquired from live individual differs from impostor image which is acquired from printed photo or digital screen. Some distortions such as dots or scratches may occur in the impostor image or color level of the image may change. IQA techniques are used to detect spoof attacks in recent studies [17, 18].

Deep learning is a form of supervised learning. There are multiple levels where different feature representations are obtained in each level in deep learning methods. The most prominent advantage of deep learning is that a feature representation is learned from data naturally by employing a learning method, instead of handcrafted learning by engineers. Deep learning is applied successfully in many applications such as medical image analysis, human action recognition, autonomous driving, text classification, face recognition, spoof detection [28]. CNN is one of the deep learning architectures that contains multiple convolutional layers to obtain different feature representation from the data. For instance, representation of edges and moles or

freckles can be obtained from face image with the first layer and second layer, respectively.

Additionally, BSIF is a feature extraction method which is based on texture of the image data. In that method, firstly, a binary code is computed for each pixel and histogram of pixels' binary code is used for feature representation of image data [29].

In this thesis, four novel methods have been proposed to detect spoof attacks against ear identification systems. In the first method, the fusion of various FR and NR IQA techniques is employed to detect fake and real ear images presented to biometric recognition systems under print attacks. In that method, the fusion is achieved by DLF [26].

Ear anti-spoofing against print attacks using three-level fusion of IQA measures has been proposed as a second method. In that method, the combination of SLF and DLF has been employed, unlike the first study [27].

As a third study, 5 IQA functions and CNN-based deep learning method have been fused to propose a novel method for ear anti-spoofing problem. In order to combine 5 IQA functions, SLF has been employed. Beside, DLF has been applied for the fusion of IQA-based and CNN-based methods.

Lastly, texture-based BSIF method and deep learning based CNN method have been exploited to propose a novel anti-spoofing technique for ear identification systems. The decision results of these methods have been fused by using the DLF.

Many experiments have been conducted for aforementioned methods on six different ear image datasets which are publicly available. The datasets are obtained from 4 ear databases and are called AMI, UBEAR, IITD, USTB Set 1, USTB Set 2 and USTB Set 3 [30-33].

This thesis provides significant contributions. The first one is that the usage of IQA technique that is applied to release the important features of the ear images. Another contribution in this thesis is CNN-based deep learning approach is firstly applied on ear anti-spoofing systems. In this context, the fusion of CNN and IQA functions for ear anti-spoofing systems is implemented the first time in this thesis. Furthermore, texture-based method BSIF is applied for feature extraction from ear images. Consequently, BSIF and CNN are combined to propose an efficient ear anti-spoofing method. Additionally, FLF, DLF and SLF strategies are presented for the detection of spoofing attacks for ear biometrics. Finally, ear anti-spoofing system results are demonstrated the first time on four ear databases namely, AMI, UBEAR, IITD and USTB in this thesis.

The rest of this thesis is arranged as follows: the literature review is explained in Chapter 2. Feature extraction methods used for ear anti-spoofing are described in Chapter 3. The details of the fusion techniques are given in Chapter 4. Chapter 5 describes the used ear databases. The methodologies and experimental results of proposed methods are explained in Chapter 6. Finally, the results are summarized in Chapter 7.

Chapter 2

LITERATURE REVIEW

To the best of our knowledge, state-of-the-art anti-spoofing techniques are mostly based on image quality, texture, deep representation, motion and liveness [14, 34]. In biometrics community, IQA technique is a popular way to counter spoof attacks. The images are classified as genuine or impostor image by taking the quality difference between them into consideration. In [17], 25 IQA measures where 21 of them are FR and 4 of them are NR are computed and used for spoof detection of attacks based on iris, face and fingerprint biometric traits. Furthermore, Image Distortion Analysis (IDA) is used to counter spoof attacks on face identification systems in [35]. In order to construct IDA feature vector for face images, extraction of specific features is achieved. Additionally, liveness detection method is proposed in [36] and [37] for fingerprint and iris biometric traits. In these studies, IQA technique is applied with new parameterization in order to detect the liveness of a fingerprint and an iris. Two fingerprint databases (LivDet and ATVS) and one iris database (BioSec) are used in [36] and [37], respectively. Next, in [38] another recent study that is based on palmprint biometric trait is proposed. In that study, 25 IQA measures, where 21 of them are FR and 4 of them are NR, are implemented to extract quality measures. Before extracting quality measures, image enhancement is performed by using WLD. In the final step, Euclidean distance is used as the distance measure. According to the results, the proposed method is error free, less complex and cost effective. Moreover, in [39], an anti-spoofing method is proposed for iris, face and palmprint by using IQA

techniques. Iris, face and palmprint images are fused in that study to get a single image. After this step, IQM functions are applied to extract features from that image. Finally, SVM classifier is used to decide whether that image is genuine or impostor.

Texture-based methods are implemented in the feature extraction step of an anti-spoofing system. In these type of methods, in order to determine whether the biometric test image is real or fake, the comparison of the test image of the biometric trait that is acquired by the sensor or captured by a camera and the training image of that biometric trait which is enrolled in the database is performed by analyzing their texture patterns. In the study [40], M-BSIF extraction method has been applied to describe the textures of the iris images for presentation attack detection system. In [41], LBP, CoALBP, LPQ, BSIF and SID have been implemented on color images for face spoof detection system. In [42], the proposed face spoofing approach is based on learning texture features and gradient structures of face images. The LBP and GW are used for describing texture features. The HOG is used for local shape description. The SVM is employed for classification of obtained representation. Finally, the SLF is implemented to decide whether the input image belongs to a live person or not. In another texture-based work [43], the ALWGO is used to extract highly discriminant features and SRC is implemented for discriminating genuine face images from fake ones. In [44], authors proposed a novel method to detect fake face images by using LBP to analyze the micro textures of facial images. Furthermore, a new texture-based method is proposed in [45] which focused on color face images. In that work, RGB face image is converted into HSV and YCbCr color spaces. Texture features are extracted by using variations of LBP, namely LPQ, BSIF and SURF. Additionally, in [46], the authors studied palmprint spoofing to show vulnerability of biometric recognition systems which are based on palmprint and proposed a method for spoof

detection. Statistical features are extracted from palm image by using pixel intensities, wavelet coefficients and GLCM with a feature selection method and Binary Classifier to determine the result which shows either the palm image is real or fake. Another study has been implemented by using two feature descriptors that are namely MBSIF-TOP and MLPQ-TOP to detect spoof attacks against face recognition systems [47]. In order to detect fake face images, discriminative subspace is developed by using spectral regression based kernel discriminant analysis (SR-KDA). Two feature representations which are obtained from MBSIF-TOP and MLPQ-TOP feature extractors are combined by using SR-KDA method. Consequently, authors have improved the performance of the method by fusing MBSIF-TOP and MLPQ-TOP. Beside, BSIF that is texture-based feature extractor method is applied to ensure that fingerprint recognition systems are not accessible by attackers [48, 49].

In the literature, the combination of IQA and texture-based methods is performed. In this context, a new method is proposed in [50] for palmprint liveness detection. In that method, BSIF technique is implemented to extract the texture feature and get histogram of the image. Eight full-reference image quality measures are implemented and the resulting features are combined. The SVM classifier is employed to distinguish genuine palmprint and imposter palmprint. The experiment results show the high accuracy of the proposed method. Another novel approach is proposed in [18] in order to detect spoof attack against face and palmprint identification systems. In that study, LBP, DOG and HOG are used for feature extraction of an image. Next, PCA and LDA are applied for reduction of the dimensionality of an extracted feature vector. In addition to this, 7 full-reference IQM functions are applied to measure the quality of an image.

On the other hand, many spoofing studies have been performed to exploit the advantages of CNN [51-54]. In [51], the first texture features of face images are learned by CNN. Secondly, the depth information of face images is represented by extracting features with LBP. In the classification module of the aforementioned approaches, SVM classifier is employed to classify face images (real/fake). The fusion of these approaches constitute a robust face anti-spoofing method. Furthermore, authors in [52] proposed effective method by employing data-randomization approach in the training part of CNN for face anti-spoofing. In [53], another face anti-spoofing method is proposed which develops 3D CNN architecture to learn distinct information from spatial and temporal dimensions by designing data augmentation method. In order to enhance the performance of the model, they utilized generalization regularization by minimizing the MMD distance among completely different domains. Moreover, CNN model in [54] is proposed for fingerprint anti-spoofing. In this model, a new layer named as ISE is located before full connected layer to overcome the constraint of traditional CNN which needs fixed size images. Further, in [55], the authors have compared their two studies. The first study is based on handcrafted features that are classified by SVM. The second study is based on learned features using CNN. Consequently, CNN-based approach performed better results compared to the first study. Additionally, authors have proposed finger vein presentation attack detection algorithm which is based on deep CNN in [56]. Moreover, in [57], authors have proposed a speech spoof detection method based on deep NN on the dataset of BTAS2016 and overcome other methods. In the literature, a novel method that employs CNN for multiple channels that are color, depth, infrared and thermal of face images to detect spoof attacks is proposed. The proposed multi-channel CNN (MC-CNN) has been applied on their self-created face spoof database. In order to compare their method, they have implemented baseline methods on their database. As a result,

their method outperforms baseline methods [58]. Afterwards, a novel method which employs AOS to detect edges from a face image and specialized CNN architecture to extract features from diffused input images is proposed to resist spoof attacks [59]. Additionally, there is a method which is developed to detect liveness of a fingerprint biometric trait. In that method, deep learning based DBN is employed to learn distinctive features from real or fake fingerprint images. They placed the RBM at each layer in DBN [60]. Next, authors have proposed a spoof detector for fingerprint identification systems. In that proposition, MobileNet-v1 CNN model which is fed with features that are obtained from centered and aligned local patches of minutiae points of fingerprint images is employed. Authors have contributions for improvement of the performance of fingerprint anti-spoofing systems [61]. On the other hand, CNN-based method is applied for finger vein based presentation attack detection algorithms. In the study [62], the authors have designed FPNet which is based on CNN model to detect fake finger vein images. The FPNet is conducted on publicly available databases to show its effectiveness. Additionally, features are extracted by employing CNN from both local, global and entire iris region images. The obtained feature representation vectors are fused by applying FLF and SLF. Next, instead of fully-connected layer of CNN architecture, SVM is employed for classification part [63].

Besides, texture-based method and CNN-based method are combined for anti-spoofing techniques. In this context, BSIF and CNN-based methods are combined to find robust solution to iris presentation attack detection problem [64]. In that study, different representations are obtained by employing BSIF and these representations are fed to CNN model to obtain multiple classifier results. In order to find the most important and the most prominent viewpoints, meta-analysis algorithm is applied. Finally, Random Forest fusion is used to have a final decision. In the next study, MLBP

and CNN methods are employed to obtain different features from face images and resulted with two feature vectors. Further, these vectors are combined to obtain hybrid features. Finally, the SVM is employed by using hybrid features for classification of face images as real or fake [65]. Moreover, a face anti-spoofing method that uses deep CNN model where LBP method is integrated into its first layer is proposed in [66]. Their proposed method LBPnet is evaluated by conducting experiments on NUAA database. As a result, their proposed method outperforms state-of-the-art methods.

Additionally, CNN-based methods are combined with IQA-based methods to detect spoof attacks in the literature. Consequently, the authors in [67] proposed a face anti-spoofing method which is based on fusion of image quality and motion cues with the approach of CNN. Scientists state that the use of motion of the user, such as head movement, lips movement, eye blinking and expression changes, has contribution in the solution of an anti-spoofing problem. In that approach, motion of a person is tracked to detect fraudulent attempts to the identification system. In the study [68], a novel countermeasure method has been proposed to detect spoof attacks to a face recognition system. In that method, CLNF algorithm is applied for face tracking in the low-level of the proposed method. Further, Fisher vectors are used to describe the motions in the mid-level of the proposed method. According to that study, motion-based methods can be used as an extra countermeasure in anti-spoofing systems.

Besides, liveness detection is another way to cope up with the problem of spoofing. In order to detect liveness of the biometric trait, hardware-based and software-based techniques are implemented in the literature. Hardware-based techniques can be implemented in the sensor-level of an anti-spoofing system to measure the sweat of the fingerprint, facial thermogram, blood pressure or reflection of eye by integrating

extra sensing devices. On the other hand, in the study [69], a liveness detection algorithm has been implemented by using LBP for print attack of iris images on mobile devices. Further, in the study [70], eye blinking and lip movement have been considered to detect liveness of the facial images by implementing morphological operations.

Chapter 3

FEATURE EXTRACTION METHODS USED FOR EAR ANTI-SPOOFING

In this thesis, IQA-based, CNN-based and texture-based methods have been employed for feature extraction from ear images. In this chapter, the aforementioned methods are explained in details.

3.1 Image Quality Assessment

IQA plays an important key role to detect the distortions on the image. Since the IQA measures are sensitive to some kinds of image distortions, they are used to measure the quality of an image for steganalysis [71]. In addition to this, IQA measures are implemented for extracting significant features from varied biometric traits to be used in spoof detection algorithms [17, 18]. As a summary, IQA measures are able to measure the quality difference between genuine and impostor images in terms of blur, noise, contrast, change of illumination, sharpness or any other distortions.

IQA measures are classified into two groups as FR and NR. FR measures are applied when there are two images to be compared. In this group, there is a reference image which is assumed as original image. Reference image is compared with the distorted image to analyze the differences between them by using FR functions. On the other hand, NR functions are used to measure the quality of a reference image without comparison [17].

In this thesis, 21 FR and 4 NR IQA measures are implemented for extracting features of ear biometric images to develop an efficient anti-spoofing method for ear recognition systems. In the next section, the details of FR and NR IQA measures are explained in which a summary of these measures is demonstrated in Figure 2.

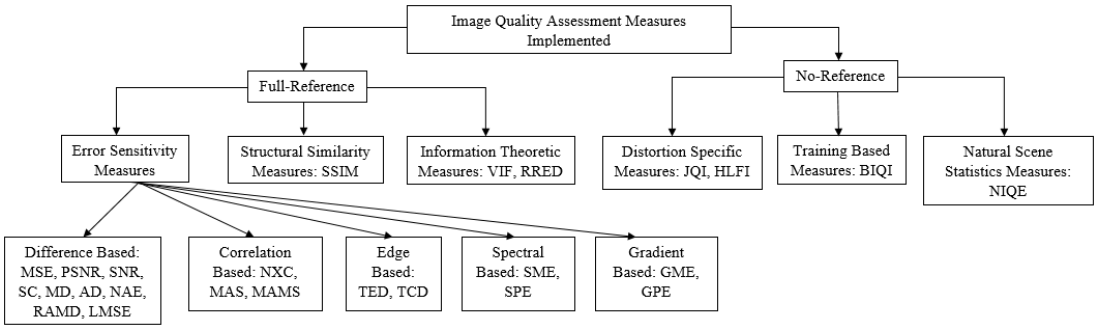

Figure 2. Classification of 25 image quality measures [17].

3.1.1 Full-Reference Measures

FR measures are collected in three different titles. Firstly, error sensitivity measures are used to measure the error between the reference image and distorted image. These measures can be introduced in five categories. The first category contains pixel difference measures which are listed as Mean Squared Error (MSE) [72], Peak Signal to Noise Ratio (PSNR) [73], Signal to Noise Ratio (SNR) [74], Structural Content (SC) [75], Maximum Difference (MD) [75], Average Difference (AD) [75], Normalized Absolute Error (NAE) [75], R-Averaged Maximum Difference (RAMD) [72] and Laplacian Mean Squared Error (LMSE) [75]. The formulas of aforementioned measures are given below.

Mean Squared Error is given as:

$$\mathrm{MSE}(\mathrm{I},\hat{\mathrm{I}}) = \frac{\mathbf{1}}{\boldsymbol{NM}} \sum_{\boldsymbol{i=1}}^{\boldsymbol{N}} \sum_{\boldsymbol{j=1}}^{\boldsymbol{M}} (\boldsymbol{I_{i,j}} - \hat{\mathrm{I}}_{\boldsymbol{i,j}})^{\mathbf{2}} \quad (1)$$

where I and Î represent original (real) and distorted (fake) images, respectively. The size of the image is represented by *NM*.

Peak Signal to Noise Ratio is calculated as:

$$\mathrm{PSNR}(\mathrm{I},\hat{\mathrm{I}}) = 10\ \log\left(\frac{\max(I^2)}{\mathrm{MSE}(\mathrm{I},\hat{\mathrm{I}})}\right) \tag{2}$$

Signal to Noise Ratio is computed as follows:

$$\mathrm{SNR}(\mathrm{I},\hat{\mathrm{I}}) = 10\ \log\frac{\sum_{i=1}^{N}\sum_{j=1}^{M}(\boldsymbol{I_{i,j}})^2}{N.M.\mathrm{MSE}(\mathrm{I},\hat{\mathrm{I}})} \tag{3}$$

Structural Content is calculated as follows:

$$\mathrm{SC}(\mathrm{I},\hat{\mathrm{I}}) = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}(\boldsymbol{I_{i,j}})^2}{\sum_{i=1}^{N}\sum_{j=1}^{M}(\hat{\mathrm{I}}_{\boldsymbol{i,j}})^2} \tag{4}$$

Maximum Difference formula is given below:

$$\mathrm{MD}(\mathrm{I},\hat{\mathrm{I}}) = \max|\boldsymbol{I_{i,j}} - \hat{\mathrm{I}}_{\boldsymbol{i,j}}| \tag{5}$$

Average Difference is calculated as follows:

$$\mathrm{AD}(\mathrm{I},\hat{\mathrm{I}}) = \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}(I_{i,j} - \hat{\mathrm{I}}_{i,j}) \tag{6}$$

Normalized Absolute Error is given as follows:

$$\mathrm{NAE}(\mathrm{I},\hat{\mathrm{I}}) = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}|\boldsymbol{I_{i,j}} - \hat{\mathrm{I}}_{i,j}|}{\sum_{i=1}^{N}\sum_{j=1}^{M}|I_{i,j}|} \tag{7}$$

R-averaged Maximum Difference is computed as follows:

$$\mathrm{RAMD}(\mathrm{I},\hat{\mathrm{I}},\mathrm{R}) = \frac{1}{R}\sum_{r=1}^{R} max_r\ \left|\boldsymbol{I_{i,j}} - \hat{\mathrm{I}}_{\boldsymbol{i,j}}\right| \tag{8}$$

where R=10 and R-highest pixel difference between two images is calculated.

Laplacian Mean Squared Error is calculated as follows:

$$\mathrm{LMSE}(\mathrm{I},\hat{\mathrm{I}}) = \frac{\sum_{i=1}^{N-1}\sum_{j=2}^{M-1}\left(h(\boldsymbol{I_{i,j}}) - h(\hat{\mathrm{I}}_{\boldsymbol{i,j}})\right)^2}{\sum_{i=1}^{N-1}\sum_{j=2}^{M-1}\left(h(\boldsymbol{I_{i,j}})\right)^2} \tag{9}$$

where $h(\boldsymbol{I_{i,j}})$ [17] is computed as follows:

$$h(\boldsymbol{I_{i,j}}) = I_{i+1,j} + I_{i-1,j} + I_{i,j+1} + I_{i,j-1} - 4 * I_{i,j} \quad (10)$$

The second category is named as correlation-based measures and includes Normalized Cross-Correlation (NXC) [75], Mean Angle Similarity (MAS) [72] and Mean Angle Magnitude Similarity (MAMS) [72]. The third category is edge-based measures which include Total Edge Difference (TED) [76] and Total Corner Difference (TCD) [76]. Next category is spectral distance measures and includes Spectral Magnitude Error (SME) [78] and Spectral Phase Error (SPE) [78]. Finally, the last category is gradient-based measures which include Gradient Magnitude Error (GME) [79] and Gradient Phase Error (GPE) [79].

Normalized Cross-Correlation formula is given below:

$$\mathrm{NXC}(\mathrm{I}, \hat{\mathrm{I}}) = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}(\boldsymbol{I_{i,j}} \cdot \hat{\mathrm{I}}_{i,j})}{\sum_{i=1}^{N}\sum_{j=1}^{M}(I_{i,j})^2} \quad (11)$$

Mean Angle Similarity is computed as follows:

$$\mathrm{MAS}(\mathrm{I}, \hat{\mathrm{I}}) = 1 - \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}(\alpha_{i,j}) \quad (12)$$

where $\alpha_{i,j}$ [17] represents angle between two vectors and is computed as follows:

$$\alpha_{i,j} = \frac{2}{\pi}\cos^{-1}\frac{\langle \boldsymbol{I_{i,j}}, \hat{\mathrm{I}}_{i,j}\rangle}{\|\boldsymbol{I_{i,j}}\| . \|\hat{\mathrm{I}}_{i,j}\|} \quad (13)$$

where $\langle \boldsymbol{I_{i,j}}, \hat{\mathrm{I}}_{i,j}\rangle$ represents the scalar product of two images and length of $\boldsymbol{I_{i,j}}$ and $\hat{\mathrm{I}}_{i,j}$ are represented by $\|\boldsymbol{I_{i,j}}\|$ and $\|\hat{\mathrm{I}}_{i,j}\|$.

Mean Angle Magnitude Similarity formula is given below:

$$\mathrm{MAMS}(\mathrm{I}, \hat{\mathrm{I}}) = \frac{1}{NM}\sum_{i=1}^{N}\sum_{J=1}^{M}\left(1 - \left[1 - \alpha_{i,j}\right]\left[1 - \frac{\|\boldsymbol{I_{i,j}} - \hat{\mathrm{I}}_{i,j}\|}{255}\right]\right) \quad (14)$$

Total Edge Difference is computed as follows:

$$\mathrm{TED}(\mathrm{I}, \hat{\mathrm{I}}) = \frac{1}{NM}\sum_{i=1}^{N}\sum_{J=1}^{M}\left|\boldsymbol{I_{E_{i,j}}} - \hat{\mathrm{I}}_{\boldsymbol{E}_{i,j}}\right| \quad (15)$$

where I_E and $\hat{\mathrm{I}}_E$ represent the binary edge maps of original image and distorted image, respectively. These edge maps are obtained by Sobel operator [17].

Total Corner Difference formula is given below:

$$\mathrm{TCD}(\mathrm{I},\hat{\mathrm{I}}) = \frac{|\boldsymbol{N}_{cr}-\widehat{\boldsymbol{N}}_{cr}|}{max(\boldsymbol{N}_{cr},\widehat{\boldsymbol{N}}_{cr})} \tag{16}$$

where N_{cr} and $\widehat{N}_{cr}$ represent the number of obtained corners in the original and distorted images, respectively [17, 77].

Spectral Magnitude Error is calculated as follows:

$$\mathrm{SME}(\mathrm{I},\hat{\mathrm{I}}) = \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}\left(|F_{i,j}| - |\hat{F}_{i,j}|\right)^2 \tag{17}$$

where $F_{i,j}$ and $\hat{F}_{i,j}$ represent the Fourier transforms of original and distorted images, respectively [17].

Structural Phase Error is computed as follows:

$$\mathrm{SPE}(\mathrm{I},\hat{\mathrm{I}}) = \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}\left|\arg(F_{i,j}) - \arg(\hat{F}_{i,j})\right|^2 \tag{18}$$

where $\arg(F)$ represents the phase of the Fourier transform [17].

Gradient Magnitude Error formula is described as follows:

$$\mathrm{GME}(\mathrm{I},\hat{\mathrm{I}}) = \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}\left(|G_{i,j}| - |\hat{G}_{i,j}|\right)^2 \tag{19}$$

where $G_{i,j}$ and $\hat{G}_{i,j}$ represent gradient maps of original and distorted images and calculated as follows:

$$\mathrm{G} = G_x + iG_y \tag{20}$$

where G_x and G_y represent gradients in the x and y directions.

Gradient Phase Error formula is described as follows:

$$\mathrm{GPE}(\mathrm{I},\hat{\mathrm{I}}) = \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}\left|\arg(G_{i,j}) - \arg(\hat{G}_{i,j})\right|^2 \quad (21)$$

On the other hand, the second group of FR measures is structural similarity measures which consider image degradations as perceived changes in structural information variation. It is named as Structural Similarity Index Measure (SSIM) [80]. The formula for SSIM is given below:

$$\mathrm{SSIM}(\mathrm{I},\hat{\mathrm{I}}) = \frac{(2\mu_I\mu_{\hat{I}}+C_1)(2\sigma_{I\hat{I}}+C_2)}{(\mu_I^2+\mu_{\hat{\mathrm{I}}}^2+C_1)(\sigma_I^2+\sigma_{\hat{\mathrm{I}}}^2+C_2)} \quad (22)$$

where μ_I and $\mu_{\hat{I}}$ represent the average of the original image (I) and the distorted image Î, respectively. The variance of I and Î are represented by σ_I^2 and $\sigma_{\hat{\mathrm{I}}}^2$, respectively. Also, $\sigma_{I\hat{I}}$ represents the covariance of I and Î. Lastly, C_1 and C_2 are constants.

The last group for FR measures is information theoretic measures. Since the image information between real and impostor image differs, information theoretic measures are developed. These measures are Visual Information Fidelity (VIF) [81] and Reduced Reference Entropic Difference (RRED) [82]. The details for VIF and RRED are described in [81, 82].

3.1.2 No-Reference Measures

On the other hand, no-reference (NR) measures are divided into three categories. In the first category, quality of an image is measured according to its specific distortions. JPEQ Quality Index (JQI) [83] and High-Low Frequency Index (HLFI) [84] measures are used for this purpose. The JQI is used to measure the quality of the image in the compression algorithms like JPEG. The HLFI is used to find the sharpness of the image and it is computed as follows:

$$\mathrm{HLFI(I)} = \frac{\sum_{i=1}^{i_l}\sum_{j=1}^{j_l}|F_{i,j}| - \sum_{i=i_h+1}^{N}\sum_{j=j_h+1}^{M}|F_{i,j}|}{\sum_{i=1}^{N}\sum_{j=1}^{M}|F_{i,j}|} \quad (23)$$

where i_l, i_h, j_l and j_h represent thresholds for the lower and upper frequencies of the Fourier transform of the image. In this thesis, thresholds are $i_l = i_h = 0.15N$ $j_l = j_h = 0.15M$ [17].

The second category is based on training approaches. In this category, different kinds of distortions are analyzed to have a general quality score. Blind Image Quality Index (BIQI) [85] is applied within this category. The formula of BIQI is given in [23] as follows:

$$BIQI\ (I) = \sum_{i=1}^{5} p_i \,.\, q_i \quad (24)$$

where p_i denotes probability of specific distortion in the image. In [85], it is stated that there are set of distortions namely JPEG, JPEG2000, Fast Fading, Gaussian Blur and White Noise. Further, q_i denotes the obtained quality scores for each of the distortion.

The last category is based on natural scene statistics. Some statistical properties which alter in the distorted image are available on natural scenes. While measuring the quality of an image with this approach, Natural Image Quality Evaluator (NIQE) [86] is used for the measurement. The details of NIQE is available in [86].

3.2 Convolutional Neural Networks

Generally, Neural Networks (NN) consist of neurons that are placed in the input layer, output layer and hidden layers. These neurons are connected in a specific way in order to communicate with each other. Each neuron in the network has a specific weight. Input neurons take the signal from the environment and combine it with its weight. Afterwards, the computation result is transmitted to the subsequent layer's neurons. Finally, output layer's neurons convey the computation result to the environment [28].

The CNN is a widely used supervised deep learning model which was developed by [87]. The first component of this model consists of convolutional layers. In order to extract feature representation of an input image, convolutional operation is applied by using convolution filters in that layer. The input image is searched to detect different visual elements by convolving it with learned multiple filters such as vertical filter, horizontal filter or diagonal filter where each filter extracts different features. The convolved results are called as feature maps. The number of obtained feature maps is equal to the number of convolutional filters used. In order to transmit computed output values of neurons of current layer to the next layer, non-linear function is needed to detect non-linear features.

The most commonly used non-linear function in CNN architectures is ReLU because it works better in terms of speed. ReLU puts zero instead of negative values which represent black. Implementing the first convolutional layer provides to obtain low-level features such as edges, color, texture, gradient orientation, etc. of an image. Additionally, if more convolutional layers are added into implementation, high-level features will be obtained as well. This approach will led network to learn the input image deeper. Consequently, we can say that using multiple convolutional layers is advantageous.

The next component of a CNN model is called pooling layer which aims to reduce the computation of data and select important features. The function of this layer is subsampling the output feature representation of convolutional layer for dimension reduction of the feature map. One of the most commonly used pooling layers is Max pooling which selects maximum value from subdivided feature map. Afterwards, obtained features are flattened to construct one dimensional feature vector and this

feature vector is fed to fully-connected layers. The principle of fully-connected layer is like traditional NN model. It consists of input layer, output layer and hidden layers. In a fully-connected layer, each neuron of current layer is connected to all neurons of the next layer. In this context, the obtained feature vector will be received by neurons of input layer and pass through all hidden layers. Finally, the output of the last hidden layer will be the input to output layer to be classified as a specific class. In the output layer, Softmax classifier is applied for the classification of an image.

The pattern of input image is learned by the convolutional network and the classification provides the output. In that network, every iteration of training is achieved by applying backpropagation algorithm. In order to obtain the best classification rate, the network model will be trained over a number of epochs [88]. In the CNN model, memorizing training data problem which is also known as overfitting occurs in the case of small number of image data. This problem can be prevented by using regularization techniques where Dropout is one of these techniques. The working principle of Dropout is that some neurons are dropped in every iteration of NN model. Therefore, the network is not dependent to specific neurons [89]. Moreover, in order to optimize CNN model, some techniques are available. One of these techniques is Batch Normalization (BN) that is applied to normalize the inputs of each layer to overcome internal covariant shift problem [90].

3.3 Binarized Statistical Image Features

Binarized Statistical Image Features is a texture-based method for local image description. This method is a variant of LBP and LPQ approaches. The methodology of this approach is to compute a binary code string for each pixel of an input image by using a learnt filter which is obtained by applying ICA to natural image patches. The

linear filter is applied to the image patch that has same size with the filter. The filter response S_i is obtained for each pixel as described in the following formula:

$$S_i = \sum_{u,v} W_i(u,v) * X(u,v) \quad (25)$$

where $W(u,v)$ and $X(u,v)$ represent the linear filter and the image patch, respectively. If the obtained $S_i > 0$, the binary code for that pixel b_i is determined as 1 and 0 otherwise. The contribution of BSIF is to use the statistics of natural image patches to obtain good representation of features [91]. BSIF algorithm is implemented for the applications of identification of a person based on biometric trait and spoof detection systems [47, 92, 93].

Chapter 4

FUSION TECHNIQUES

In order to combine multiple information like features extracted from a biometric trait or scores obtained from multiple matchers or decisions made by multiple decision modules, information fusion techniques are used in the biometrics community. The most commonly used fusion levels are listed as follows: Feature-Level Fusion (FLF), Score-Level Fusion (SLF) and Decision-Level Fusion (DLF) [94].

4.1 Feature-Level Fusion

The Feature-Level Fusion techniques can be accomplished for several scenarios. One of these scenarios is that the same feature extraction algorithm is applied to obtain feature vector sets of multiple samples of the same biometric trait. In this case, features which are extracted from each sample of a biometric trait need to be concatenated to obtain the common feature vector. The second scenario is that the feature vector sets are obtained by applying multiple feature extraction algorithms or multiple samples of different biometric traits. Before fusion stage, feature normalization is applied to transform the extracted features into a common domain. In both scenarios, obtained multiple feature vector sets need to be concatenated to obtain a common feature vector set by implementing feature-level fusion techniques. In this thesis, totally 25 IQA measures are implemented for extracting the features of each ear image. Thus, the second aforementioned scenario is applicable for our study. The FLF method is implemented to combine multiple features that are obtained by using 25 IQA measures in our experiments.

4.2 Decision-Level Fusion

In the decision module of a multimodal biometric recognition system, more than one decision may occur. The DLF is implemented to make a final decision among multiple results. In the literature, some of the DLF methods are AND rule, OR rule and majority voting. The AND rule can be described as follows: at least one impostor result leads the final result to be in the impostor class. On the other hand, the OR rule can be described as follows: at least one genuine result leads the final result to be in the genuine class. In the majority voting method, the final result is obtained (genuine or impostor) according to the majority of the matchers' decisions [95].

4.3 Score-Level Fusion

In order to have efficient multimodal biometric recognition system, more than one matcher can be implemented in the matching module of the system. Computed scores of these matchers must be combined to obtain the common score. Before fusion process, score normalization is applied for transformation of computed scores of distinct matchers into a mutual space. Some of the score normalization techniques are min-max, tanh and z-score. Score-level fusion method is the most powerful fusion method in many applications [2, 3, 18].

Chapter 5

EAR DATABASES

In this thesis, the experimental studies have been conducted on 6 different ear datasets. The name of these datasets are AMI, UBEAR, IITD, USTB Set 1, USTB Set 2 and USTB Set3.

AMI database which contains 700 ear images that are taken from left and right side of 100 different subjects was created by Esther Gonzalez [30]. All ear images are taken under the same illumination conditions. The dimension of the original ear images is 492x702 pixels.

UBEAR database was created by Soft Computing and Image Analysis group in 2011 [31]. There are 4430 ear images that are taken from left and right side of 126 different subjects under different light conditions. The dimension of the ear images is 1280x960 pixels.

The IITD database was constructed in IIT Delhi campus [32]. There are right ear of 221 subjects with the size of 272x204 in the IITD dataset that is used for the experiments.

The USTB Set 1 database is constructed in 2002 and it contains right ear of 60 subjects with the resolution of 80x150. The USTB set 2 database that contains 77 subjects with

the resolution of 300x400 is constructed in 2003. The USTB set 3 dataset that contains 79 subjects with the resolution of 768x576 is constructed in 2004 [33].

In this thesis, in order to construct datasets for experimental studies, 200 ear images have been selected from AMI and UBEAR databases. Additionally, 124, 60, 76 and 78 ear images have been selected from IITD, USTB Set 1, USTB Set 2 and USTB Set 3, respectively.

Moreover, there is no spoof ear database in the biometric community. Consequently, ear spoof databases have been created for 6 datasets. In order to construct spoof databases, each ear image is printed by using Olivetti d-colour mf223 printer. In this thesis, the original selected ear images are cropped, resized to 256x256 dimension and stored as real images for all databases. Beside of this, in order to construct spoof database for all databases, each ear image is printed by using Olivetti d-colour mf223 printer which has 1800x600 resolution. Next, the picture of printed ear images are taken from 30 cm distance by using IPhone 6S camera which is equipped with 12 megapixel. The captured ear images are cropped to get rid of the appearance of A4 paper by using Paint 3D and resized to 256x256. Their size was 1240x1780 before image alignment operation. Consequently, spoof databases are created for all of the databases separately. This type of spoof attack is called a printed photo attack. As a result, real ear images and corresponding fake ear images are stored for each aforementioned database. Samples of real and fake ear images of AMI and UBEAR databases are demonstrated in Figure 3.

Figure 3. Real (first row) and fake (second row) samples from AMI (first two columns) and UBEAR (last two columns) databases

Additionally, samples of real and fake ear images of IITD database are demonstrated in Figure 4.

Figure 4. Real (first row) and fake (second row) samples from IITD database

Further, samples of real and fake ear images of IITD, USTB Set 1, USTB Set 2 and USTB Set 3 datasets are demonstrated in Figure 5.

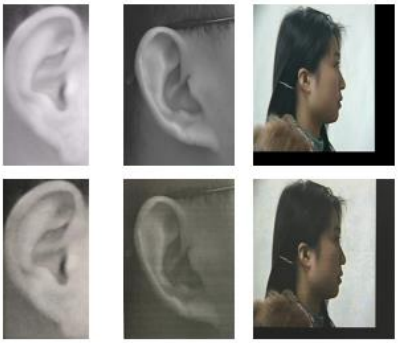

Figure 5. Real (first row) and fake (second row) samples from USTB Set 1 (first column) USTB Set 2 (second column) and USTB Set 3 (third column) databases

Chapter 6

PROPOSED METHODS FOR EAR ANTI-SPOOFING

In this thesis, ear anti-spoofing problem is studied to propose an efficient and robust method. In this context, different types of methods have been applied on various ear databases. Further, IQA and BSIF methods are implemented in MATLAB R2015b and CNN is implemented in Python. Additionally, IQA and BSIF methods are implemented on 2.20 GHz Intel (R) Pentium (R) CPU B960 machine with 4 GB RAM. Additionally, CNN is implemented on 2.00 GHz Intel (R) Xeon (R) E5-2660 v4 with machine 23 GB RAM. Results of the conducted experiments are presented as False Fake Rate (FFR) and False Genuine Rate (FGR) that represent the number of genuine images that are categorized as impostor and the number of impostor images that are categorized as genuine, respectively. Additionally, Half Total Error Rate (HTER) is computed as follows:

$$HTER = \frac{FFR+FGR}{2} \tag{26}$$

In this thesis, results which are represented with rates of FFR, FGR and HTER are presented in percentage. The details of the proposed methods are explained in the following subsections.

6.1 Fusion of Full-Reference and No-Reference Anti-Spoofing Techniques for Ear Biometrics under Print Attack

Firstly, an anti-spoofing method that employs the fusion of various FR and NR IQA techniques has been proposed to detect fake and real ear images presented to biometrics systems under print attacks. In this context, FR IQA measures such as Error

Sensitivity Measures, Pixel Difference Measures, Correlation-Based Measures, Edge-Based Measures, Spectral Distance Measures, Gradient-Based Measures, Structural Similarity Measures and Information Theoretic Measures have been used. Additionally, NR IQA measures such as Distortion Specific Measures, Training Based Measures and Natural Scene Statistics Measures have been implemented to distinguish fake and real ear images. A comparative analysis of the performance of these quality metrics and the proposed method using decision-level fusion of all aforementioned measures has been performed. The experimental results have been presented using AMI and UBEAR ear databases by creating print attack counterparts of the ear images used in these databases.

6.1.1 Methodology

In this section, the proposed method is explained in detail. Figure 6 shows the system diagram of the first proposed spoof detection method which is based on decision-level fusion. In the system diagram, firstly, I and Î are converted into gray-scale and denote original ear image and smoothed (distorted) ear image, respectively. Next, Gaussian filter (0.5 σ and 3x3 kernel size) is applied to smooth the acquired image of ear biometrics for the application of FR IQA. Additionally, 21 FR IQA measures are employed to compare the original image and the smoothed image. Moreover, 4 NR IQA measures are employed to measure the quality of reference image (I). In the matching part of the system, Nearest Neighbor classifier is applied to make a decision (real or fake) for each IQA measure. After obtaining the results for all IQA measures, the final decision is found by applying DLF using Majority Voting technique.

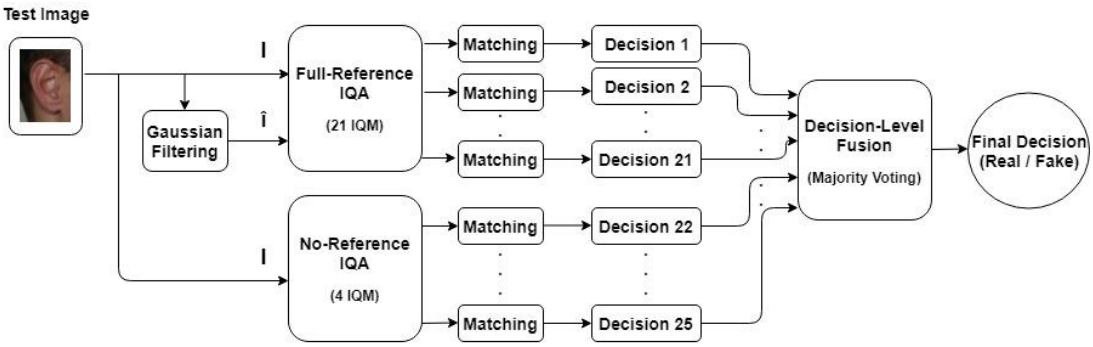

Figure 6. Block diagram of the first proposed method

6.1.2 Experimental Results

In order to show the performance of the first proposed spoof detection method, several experiments have been performed on two ear databases namely, AMI and UBEAR.

In the first experiment, 25 IQA measures are analyzed one by one. According to the results, SSIM has the minimum HTER with 4.5 for AMI database as shown in Table 1. However, as shown in Table 1 HTER of SSIM for UBEAR database is not the minimum. Therefore, we cannot make a decision according to the result of one specific IQA measure.

On the other hand, fusion of the features that are extracted by using 25 IQA measures has been analyzed to produce robust method. The experimental results using FLF method and the proposed DLF method have been demonstrated in the last two rows of Table 1. As shown in Table 1, DLF which is the first proposed method has 8.5 and 15.5 HTER values for AMI and UBEAR databases, respectively. Additionally, FLF produces 16.0 and 17.5 HTER values for AMI and UBEAR databases, respectively. Consequently, the first proposed method outperforms FLF of IQM functions.

Table 1. Results (%) for AMI and UBEAR databases.

#	Type	IQM	AMI database			UBEAR database		
			FFR	FGR	HTER	FFR	FGR	HTER
1	FR	MSE	24.0	13.0	18.5	18.0	23.0	20.5
2	FR	AD	23.0	24.0	23.5	24.0	26.0	25.0
3	FR	GME	26.0	18.0	22.0	21.0	20.0	20.5
4	FR	GPE	2.0	8.0	5.0	24.0	26.0	25.0
5	FR	LMSE	9.0	4.0	6.5	23.0	24.0	23.5
6	FR	MD	20.0	26.0	23.0	20.0	25.0	22.5
7	FR	NAE	18.0	26.0	22.0	22.0	21.0	21.5
8	FR	NCC	27.0	15.0	21.0	23.0	24.0	23.5
9	FR	PSNR	24.0	13.0	18.5	18.0	23.0	20.5
10	FR	RMD	28.0	21.0	24.5	18.0	29.0	23.5
11	FR	SNR	24.0	15.0	19.5	25.0	27.0	26.0
12	FR	SME	24.0	8.0	16.0	20.0	15.0	17.5
13	FR	SPE	23.0	15.0	19.0	14.0	29.0	21.5
14	FR	SC	24.0	15.0	19.5	25.0	27.0	26.0
15	FR	SSIM	7.0	2.0	4.5	23.0	14.0	18.5
16	FR	TCD	20.0	31.0	25.5	3.0	39.0	21.0
17	FR	TED	25.0	28.0	26.5	21.0	28.0	24.5
18	FR	VIF	27.0	33.0	30.0	24.0	2.0	13.0
19	FR	RRED	16.0	15.0	15.5	25.0	28.0	26.5
20	FR	MAMS	24.0	19.0	21.5	26.0	26.0	26.0
21	FR	MAS	29.0	24.0	26.5	12.0	23.0	17.5
22	NR	JQI	10.0	3.0	6.5	14.0	25.0	19.5
23	NR	HLFI	24.0	27.0	25.5	26.0	22.0	24.0
24	NR	BIQI	9.0	2.0	5.5	19.0	18.0	18.5
25	NR	NIQE	21.0	13.0	17.0	17.0	31.0	24.0
Our study		FLF of 25 IQMs	24.0	8.0	16.0	20.0	15.0	17.5
First Proposed Method		DLF of 25 IQMs	11.0	6.0	8.5	9.0	22.0	15.5

6.1.3 Discussion on Experimental Results

In this method, spoof detection of ear biometrics is focused under print attacks. The proposed method is based on decision-level fusion of different types of FR and NR IQA measures. According to the comparisons of the proposed method with individual IQA measures and FLF of these metrics, the proposed method achieves better results on both AMI and UBEAR databases. HTERs of the proposed DLF method on AMI

and UBEAR datasets used in the experiments are 8.50 and 15.50, respectively whereas the HTERs of the FLF approach are 16.0 and 17.5 on the aforementioned datasets under print attacks. Consequently, the proposed DLF method outperforms the FLF approach under print attacks on both ear datasets.

6.2 Ear Anti-Spoofing Against Print Attack Using Three-Level Fusion of Image Quality Measures

This proposed method is based on FR and NR IQA methods. Three-level SLF and DLF techniques are employed in the solution of the second proposed method. The experiments have been conducted by applying the printed photo attack images of AMI and UBEAR ear databases. In the second proposed method, the number of ear images have been increased to 200 which was 100 in the first proposed method. The second proposed system significantly recognizes real and fake ear images compared to the other systems implemented in this study. HTERs of the proposed system using printed photo attack images have been compared with the error rates of the other implemented systems employing various fusion techniques for ear anti-spoofing. Additionally, the proposed system have been compared with the state-of-the-art anti-spoofing techniques and CNN-based deep learning anti-spoofing systems against print attacks on other biometric traits since this is the first study presenting ear anti-spoofing systems using SLF and DLF of IQA methods.

6.2.1 Methodology

The second proposed method extracts the ear features using 25 IQA measures and then employs matching using Manhattan Distance measure [96]. The scores obtained after matching are then normalized and SLF and DLF are applied to obtain the final decision.

Figure 7 demonstrates the general block diagram of the second proposed anti-spoofing method. The proposed method can be explained in three stages.

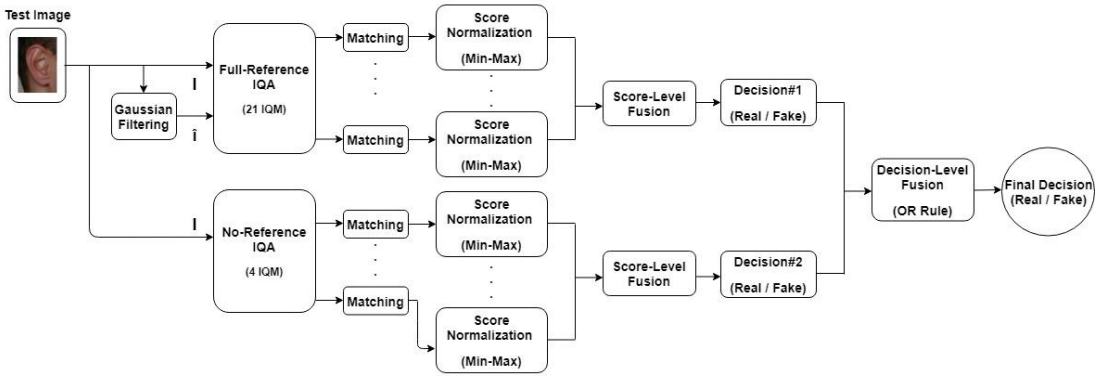

Figure 7. General block diagram of the second proposed ear anti-spoofing method

In the first stage, the test ear image is acquired and converted into grayscale before filtering step. Later, it is filtered with the Gaussian filtering method with 3x3 kernel size and 0.5 as σ value. I represents the reference image that is considered as the real image and Î is the distorted image that is considered as the fake image. Next, I and Î are used as inputs into the 21 FR IQA measure functions. In the meanwhile, I is used as an input to the 4 NR IQA measure functions. After feature extraction step, distance scores are classified by using Nearest Neighbor classifier.

In the second stage, SLF is implemented. Totally, 25 matching scores are obtained at the end of the first stage. Next, score normalization is applied to the computed scores to transform them into common domain [97]. This step is applied to the scores of 21 FR measures and 4 NR measures separately. In this context, Min-max normalization is employed. After normalization step, Sum Rule is applied for fusion of the scores of both FR measures and NR measures separately. The minimum distance among summed scores is accepted as a decision (real/fake). Two decisions are obtained after

SLF; one is obtained from FR IQA module and the second is obtained from NR IQA module.

In the third stage, DLF is implemented for the fusion of two decisions obtained in the previous stage. In the final step, OR rule is employed to decide whether the test ear sample is genuine or impostor. In the OR rule, if one of the obtained decisions is genuine, then the final decision is going to be a genuine, otherwise it is going to be an impostor.

6.2.2 Experimental Results

In this section, experiments that have been conducted to show the effectiveness of the second proposed anti-spoofing method are explained in detail. These experiments have been conducted on two databases namely AMI and UBEAR ear databases.

Many experiments have been conducted to show the effectiveness of the second proposed anti-spoofing method in this thesis.

Firstly, each IQA measure has been evaluated individually. As shown in Table 2, GPE has the minimum HTER which is 7.0 for AMI database. On the other hand, the minimum HTER does not belong to GPE for UBEAR database. Thus, if we make the final decision according to the result of one IQA measure, some inconsistencies may happen between two databases.

As a second experiment, FLF is employed to combine the features extracted from ear images. In this experiment, 25 IQA measures have been implemented for a test ear image and the extracted features of these measures have been concatenated by

implementing FLF. As shown in Table 3, 39.5 and 26.5 HTER values have been obtained for AMI and UBEAR datasets, respectively.

Table 2. Ear anti-spoofing results (%) based on each IQA measure

#	TYPE	IQM	AMI Database			UBEAR Database		
			FFR	FGR	HTER	FFR	FGR	HTER
1	FR	MSE	42.0	26.0	34.0	39.0	30.0	34.5
2	FR	AD	38.0	47.0	42.5	46.0	36.0	41.0
3	FR	GME	48.0	31.0	39.5	15.0	40.0	27.5
4	FR	GPE	10.0	4.0	7.0	40.0	29.0	34.5
5	FR	LMSE	10.0	20.0	15.0	37.0	52.0	44.5
6	FR	MD	46.0	36.0	41.0	40.0	54.0	47.0
7	FR	NAE	49.0	14.0	31.5	45.0	39.0	42.0
8	FR	NCC	41.0	28.0	34.5	59.0	43.0	51.0
9	FR	PSNR	42.0	26.0	34.0	38.0	30.0	34.0
10	FR	RMD	46.0	59.0	52.5	37.0	63.0	50.0
11	FR	SNR	44.0	26.0	35.0	58.0	41.0	49.5
12	FR	SME	44.0	39.0	41.5	26.0	28.0	27.0
13	FR	SPE	39.0	33.0	36.0	50.0	52.0	51.0
14	FR	SC	44.0	26.0	35.0	58.0	41.0	49.5
15	FR	SSIM	5.0	11.0	8.0	47.0	49.0	48.0
16	FR	TCD	39.0	56.0	47.5	44.0	53.0	48.5
17	FR	TED	33.0	76.0	54.5	31.0	57.0	44.0
18	FR	VIF	35.0	55.0	45.0	65.0	45.0	55.0
19	FR	RRED	24.0	28.0	26.0	55.0	44.0	49.5
20	FR	MAMS	61.0	42.0	51.5	54.0	55.0	54.5
21	FR	MAS	50.0	57.0	53.5	33.0	53.0	43.0
22	NR	JQI	12.0	15.0	13.5	32.0	44.0	38.0
23	NR	HLFI	42.0	40.0	41.0	58.0	48.0	53.0
24	NR	BIQI	10.0	14.0	12.0	30.0	45.0	37.5
25	NR	NIQE	40.0	48.0	44.0	47.0	36.0	41.5

The DLF has been employed as a third experiment. In this experiment, 25 IQA measures have been implemented separately for a test ear image and 25 decisions have been obtained. In order to make a final decision from obtained decisions, DLF has

been implemented. As shown in Table 3, 17.5 and 30.5 HTERs have been obtained for AMI and UBEAR datasets, respectively.

Table 3. Ear anti-spoofing results (%) based on the fusion methods

Method Name	AMI Database			UBEAR Database		
	FFR	**FGR**	**HTER**	**FFR**	**FGR**	**HTER**
FLF	41.0	38.0	39.5	24.0	29.0	26.5
DLF	17.0	18.0	17.5	29.0	32.0	30.5
2 Level DLF	6.0	11.0	8.5	17.0	23.0	20.0
SLF (min-max)	2.0	3.0	2.5	16.0	15.0	15.5
SLF (tanh)	2.0	11.0	6.5	41.0	16.0	28.5
SLF (z-score)	38.0	47.0	42.5	46.0	36.0	41.0
SLF + DLF (tanh)	4.0	0.0	2.0	34.0	0.0	17.0
SLF + DLF (z-score)	38.0	0.0	19.0	46.0	0.0	23.0
SLF + DLF (min-max)	2.0	0.0	1.0	20.0	0.0	10.0

The fourth experiment employs 2-level DLF. In this experiment, DLF has been implemented for FR IQA measures and NR IQA measures separately. Next, another DLF has been implemented as a second level to make a final decision from two decisions that are obtained in the first level. As shown in Table 3, 8.5 and 20.0 HTERs have been obtained for AMI and UBEAR datasets, respectively.

Afterwards, SLF is employed as the fifth experiment. Firstly, the scores are computed by using 25 IQA measures for a test ear image. Before fusion of the scores, score normalization have been applied by using min-max, tanh and z-score normalization techniques, respectively. Next, SLF has been implemented to get a common score by using Sum Rule. As shown in Table 3, 2.5 and 15.5 HTERs have been achieved by using min-max normalization for AMI and UBEAR datasets, respectively. Also, 6.5 and 28.5 HTERs have been obtained by using tanh normalization for AMI and UBEAR

datasets, respectively. Finally, 42.5 and 41.0 HTERs have been obtained by using z-score normalization for AMI and UBEAR datasets, respectively.

Moreover, SLF and DLF have been combined in the last experiment. This experiment is the second proposed anti-spoofing method. As shown in Table 3, 1.0 and 10.0 HTER values have been achieved by using min-max normalization for AMI and UBEAR datasets, respectively. On the other hand, 2.0 and 17.0 HTER values have been obtained by using tanh normalization for AMI and UBEAR datasets, respectively. Finally, 19.0 and 23.0 HTER values have been obtained by using z-score normalization for AMI and UBEAR datasets, respectively.

According to the results obtained, 2-level DLF achieves better performance for both AMI and UBEAR datasets when it is compared with the FLF and DLF. On the other hand, SLF with min-max normalization outperforms the FLF, DLF and 2-level DLF for both AMI and UBEAR datasets. Moreover, combination of SLF and DLF which is the second proposed method has the best performance over all aforementioned methods.

Anti-spoofing methods for ear biometrics are not studied in biometric community. Therefore, the proposed method is compared with state-of-the-art techniques which are implemented for other biometric traits against print attacks. Comparison results are demonstrated in Table 4. According to the results, HTER value of 7.9 is obtained for face biometric trait by implementing IQA technique against print attacks [17]. Additionally, authors in [40] focused on iris biometric trait and implemented M-BSIF and SVM to counter print attacks. HTER value of 0.29, 0.0, 1.27 and 0.0 have been obtained for VSIA, MobILive, LivDet and ATVS fake iris databases, respectively.

Table 4. Comparison with state-of-the-art methods using print attacks

Reference	Method Name	Biometric Trait	Databases Used	#Reals	#Fakes	HTER
Galbally et al. [17]	IQA-Based	Face	Replay Print Attack	100	100	7.9
Raghavendra et al. [40]	MBSIF + SVM	Iris	VSIA Print Attack	550	550	0.29
			MobILive Print Attack	-	-	0.0
			LivDet Iris Print Attack	1274	729	1.27
			ATVS Fake Iris Print Attack	800	800	0.0
Nguyen et al. [98]	FFT + Haar + Daubechies + SVM	Fingervein	Fingervein Print Attack	3300	2520	1.476
Farmanbar and Toygar [18]	HOG+NAE	Palmprint	PolyU Print Attack	500	500	5.8
[17]	IQA-Based	Ear	AMI Print Attack	200	200	39.5
			UBEAR Print Attack	200	200	26.5
[40]	MBSIF + SVM	Ear	AMI Print Attack	200	200	1.5
			UBEAR Print Attack	200	200	18.0
[18]	HOG+NAE	Ear	AMI Print Attack	200	200	31.5
			UBEAR Print Attack	200	200	39.0
Second Proposed Method	IQA-Based	Ear	AMI Print Attack	200	200	1.0
			UBEAR Print Attack	200	200	10.0

Furthermore, authors in [98] implemented Fourier Transform and wavelet transforms of Haar and Daubechies and SVM to detect fake finger vein images. In that study, they obtained 1.476 error rate for finger vein database. Additionally, authors [18] implemented HOG and NAE to detect presentation attack of palmprint images and HTER value of 5.8 has been obtained. Further, we have implemented the state-of-the-

art methods [17, 18, 40] on AMI and UBEAR databases and obtained 39.5 and 26.5 HTER values for [17], 1.5 and 18.0 HTER values for [40], 31.5 and 39.0 HTER values for [18], respectively. Finally, the proposed method which is an IQA-based method has been implemented to detect printed fake ear images and HTER value of 1.0 and 10.0 have been obtained for AMI and UBEAR databases, respectively. Therefore, the proposed method achieves encouraging results compared to state-of-the-art print attack anti-spoofing systems.

Additionally, CNN has been implemented in the same way as in [51] to compare it with the proposed method and state-of-the-art methods which are implemented for anti-spoofing on different biometric traits. According to the comparison results shown in Table 5, the CNN-based study which is implemented in this thesis achieves encouraging results. The second proposed study outperforms the implemented CNN-based study on UBEAR database. Although the CNN-based study and the second proposed method performs the same performance on AMI database, the proposed method has the advantages in terms of computational simplicity and computation time. Additionally, computation times of the proposed method and CNN-based method are 567 seconds and 18728 seconds on AMI dataset, respectively. Similarly, the computation times are 588 seconds and 19329 seconds on UBEAR dataset, respectively for the proposed and CNN-based methods.

6.2.3 Discussion on Experimental Results

In this thesis, the second proposed method that is an effective anti-spoofing method for ear biometric trait under printed photo attack. The proposed method is based on IQA technique. Three-level fusion of SLF and DLF of FR and NR IQA measures are implemented in the proposed anti-spoofing method. The performance of the proposed method is evaluated by using two databases namely AMI and UBEAR ear databases.

Table 5. Comparison with state-of-the-art methods using CNN

Reference	Method Used	Biometric Trait	Databases Used	Subject	Size of Image	HTER
[51]	CNN	Face	Collected by them	20	256x256	1.2
			CASIA	50		2.3
[52]	CNN	Face	CASIA-FASD	50	96x96	19.12
			Replay-Attack			8.39
[53]	CNN	Face	CASIA-FASD	50	128x128	1.4
			Replay-Attack	50		1.2
			MSU	35		0.0
			Rose-Youtu	20		7.0
[54]	CNN	Fingerprint	LivDet2013	-	200x200	3.7
			LivDet2011			6.45
[64]	BSIF+CNN	Iris	Notre Dame	-	260x260	3.28
			Warsaw			0.68
			Clarkson			9.45
			IITD+WVU			14.92
[51]	CNN	Ear	AMI Print Attack	100	256x256	1.0
			UBEAR Print Attack			14.5
Second Proposed Method	IQA-Based	Ear	AMI Print Attack	100	256x256	1.0
			UBEAR Print Attack			10.0

The second proposed method is compared with the fusion techniques namely FLF, DLF, 2-level DLF and SLF. As a result, it is clearly understood that the second proposed method achieves the best performance on the aforementioned datasets compared to the fusion techniques. Additionally, since ear anti-spoofing is not presented in the literature, we used state-of-the-art anti-spoofing print attack systems and CNN-based deep learning anti-spoofing systems applied on other biometric traits and compared the proposed ear anti-spoofing technique with the state-of-the-art methods. According to the comparison, the second proposed system achieved encouraging results.

6.3 Detection of Spoofing Attacks for Ear Biometrics Through Image Quality Assessment and Deep Learning

In this section, the third proposed method which detects spoof attacks that aim to have access to unauthorized accounts within ear recognition systems is explained. The proposed method employs CNN which is based on deep learning and IQM techniques to detect printed photo attacks against ear recognition systems. The experiments are conducted on publicly available ear datasets namely, AMI, UBEAR, IITD, USTB Set 1, USTB Set 2 and USTB Set 3 and the obtained results are compared with the state-of-the-art techniques that are focused on printed photo attacks as well.

6.3.1 Methodology

Firstly, IQM functions have been employed to extract distinctive features from test image in the proposed method. In the first step, the original test image I is filtered to obtain the smoothed image Î by using Gaussian 3x3 kernel filter which has 0.5 as σ value. This approach expects that the loss of quality delivered by Gaussian filtering varies between genuine and fake biometric test images [17]. Afterwards, I and Î have been used as an input to FR IQM functions. In the meanwhile, I has been used as an input to NR IQM function. In the matching step, Manhattan distance has been calculated between test image and all training images according to the obtained results in the previous step. In the next step, min-max score normalization has been applied to adjust scores obtained in different scale to a common scale. Score-level fusion has been employed as a final step to obtain final decision (real\fake). The schema of the third proposed method is illustrated in Figure 8. Moreover, deep learning based method CNN has been employed to discriminate test image either as real or fake. The architecture of conventional CNN part is illustrated in Figure 9 [87]. The CNN part includes 6 convolutional layers with size of 3x3 filter map and non-linear ReLu

activation function. After each convolutional layer, max pooling layer with 2x2 pooling size is applied. Further, BN follows each max pooling layer. Before flattening of feature maps, Dropout with 0.2 probability is applied to overcome overfitting problem. Next, concatenated feature vector is used as input layer to feed the NN. Epoch number is adjusted to 250 to train the NN. In the network, Softmax classifier is used to classify input ear image as real or fake. After implementation of aforementioned methods separately, two decisions have been obtained. Therefore, in order to obtain final decision among obtained decisions, decision-level fusion with OR rule has been applied.

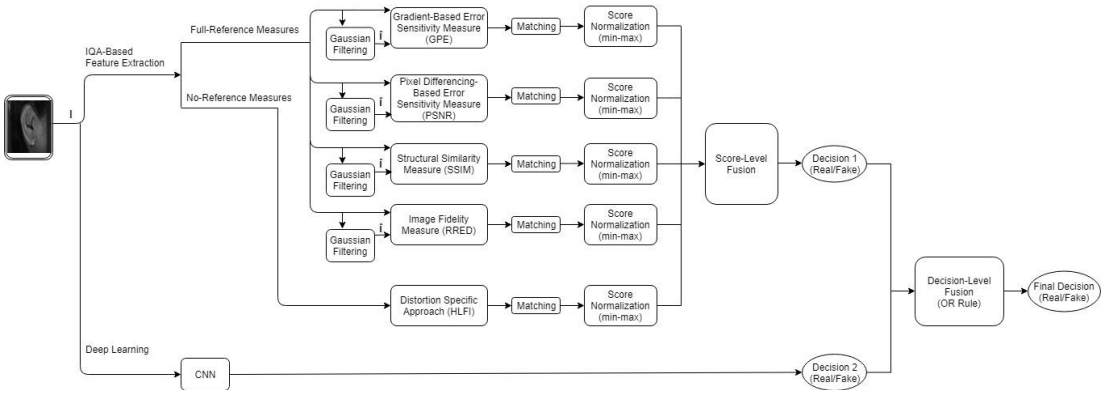

Figure 8. General block diagram of the third proposed ear anti-spoofing method

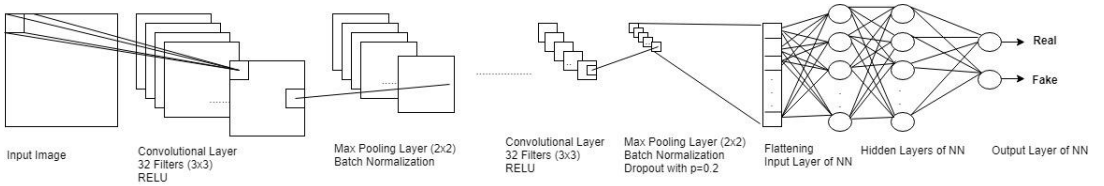

Figure 9. Architecture of the CNN part

6.3.2 Experimental Results

Several experiments have been conducted on the third proposed method. Firstly, every method have been implemented separately to observe the performance individually. Afterwards, fusion techniques have been applied to increase the performance of the

individual methods. The aforementioned experiments have been conducted on 6 different datasets namely, UBEAR, AMI, IITD, USTB Set 1, Set 2 and Set 3. The details of the conducted experiments are explained below.

Implementations of deep learning method which is CNN and IQA measures namely PSNR, GPE, SSIM, HLFI and RRED have been performed separately. The results for each method on AMI, UBEAR and IITD databases are shown in Table 6 and the error rates on USTB Set 1 and Set 2 databases are demonstrated in Table 7. Additionally, the error rates on USTB Set 3 database are demonstrated in Table 8. As shown in Table 6, Table 7 and Table 8, CNN performs better results with HTER values of 0.5, 34.5, 1.0, 4.5 and 0.0 for datasets of AMI, UBEAR, IITD, USTB Set 1 and Set 3, respectively. On the other hand, the minimum HTER value which is 0.0 is obtained by HLFI for USTB Set 2. According to the comparison of IQM functions, GPE achieves better performance with HTER values of 7.0, 36.0, 6.0 and 11.0 for AMI, UBEAR, IITD and USTB Set 1 datasets. On the other hand, HLFI performs better result with HTER values of 0.0 and 0.5 for datasets of USTB Set 2 and Set 3, respectively. Since the obtained results are not consistent for all datasets, fusion techniques have been applied to obtain more accurate results.

Table 6. Results (%) for each method on AMI, UBEAR and IITD databases

Method Name	**AMI Database**			**UBEAR Database**			**IITD Database**		
	FFR	**FGR**	**HTER**	**FFR**	**FGR**	**HTER**	**FFR**	**FGR**	**HTER**
CNN	0.0	1.0	0.5	69.0	0.0	34.5	1.0	1.0	1.0
PSNR	42.0	26.0	34.0	35.0	39.0	37.0	13.0	30.0	21.5
GPE	10.0	4.0	7.0	45.0	27.0	36.0	2.0	10.0	6.0
SSIM	4.0	11.0	7.5	46.0	39.0	42.5	0.0	39.0	19.5
RRED	24.0	28.0	26.0	58.0	54.0	56.0	34.0	31.0	32.5
HLFI	43.0	40.5	41.5	52.0	43.0	47.5	2.0	50.0	26.0

Table 7. Results (%) for each method on USTB Set 1 and Set 2 databases

Method Name	USTB Database Set 1			USTB Database Set 2		
	FFR	FGR	HTER	FFR	FGR	HTER
CNN	3.0	6.0	4.5	1.0	0.0	0.5
PSNR	18.0	9.0	13.5	3.0	0.0	1.5
GPE	14.0	8.0	11.0	7.0	7.0	7.0
SSIM	17.0	15.0	16.0	2.0	0.0	1.0
RRED	21.0	15.0	18.0	1.0	3.0	2.0
HLFI	16.0	14.0	15.0	0.0	0.0	0.0

Table 8. Results (%) for each method on USTB Set 3 database

Method Name	USTB Database Set 3		
	FFR	FGR	HTER
CNN	0.0	0.0	0.0
PSNR	18.0	3.0	10.5
GPE	3.0	0.0	1.5
SSIM	24.0	7.0	15.5
RRED	18.0	11.0	14.5
HLFI	0.0	1.0	0.5

In order to determine which fusion technique works better for the combination of the results of 5 IQM functions, SLF with min-max, tanh and z-score normalization [97], FLF and DLF with Majority Voting [94] techniques have been applied. According to the results shown in Table 9, Table 10, Table 11, Table 12 and Table 13, SLF with min-max and z-score normalizations perform same results with HTER value of 0.0 on datasets of AMI, UBEAR, USTB Set 1, Set 2 and Set 3. The results for IITD dataset is shown in Table 14. Meanwhile, DLF performs 0.0 HTER value for all datasets except UBEAR dataset. According to the results, SLF with min-max normalization technique is preferred because of its computation simplicity.

Table 9. Results (%) for fusion of IQMs on AMI database

	AMI Database		
Method Name	**FFR**	**FGR**	**HTER**
DLF (PSNR, GPE, SSIM, RRED, HLFI)	0.0	0.0	0.0
FLF (PSNR, GPE, SSIM, RRED, HLFI)	10.0	12.0	11.0
SLF-tanh (PSNR, GPE, SSIM, RRED, HLFI)	0.0	0.0	0.0
SLF-min-max (PSNR, GPE, SSIM, RRED, HLFI)	0.0	0.0	0.0
SLF-z-score (PSNR, GPE, SSIM, RRED, HLFI)	0.0	0.0	0.0

Table 10. Results (%) for fusion of IQMs on UBEAR database

	UBEAR Database		
Method Name	**FFR**	**FGR**	**HTER**
DLF (PSNR, GPE, SSIM, RRED, HLFI)	1.0	1.0	1.0
FLF (PSNR, GPE, SSIM, RRED, HLFI)	30.0	27.0	28.5
SLF-tanh (PSNR, GPE, SSIM, RRED, HLFI)	1.0	1.0	1.0
SLF-min-max (PSNR, GPE, SSIM, RRED, HLFI)	0.0	0.0	0.0
SLF-z-score (PSNR, GPE, SSIM, RRED, HLFI)	0.0	0.0	0.0

Table 11. Results (%) for fusion of IQMs on USTB Set 1 database

	USTB Set 1 Database		
Method Name	**FFR**	**FGR**	**HTER**
DLF (PSNR, GPE, SSIM, RRED, HLFI)	0.0	0.0	0.0
FLF (PSNR, GPE, SSIM, RRED, HLFI)	11.0	6.0	8.5
SLF-tanh (PSNR, GPE, SSIM, RRED, HLFI)	0.0	0.0	0.0
SLF-min-max (PSNR, GPE, SSIM, RRED, HLFI)	0.0	0.0	0.0
SLF-z-score (PSNR, GPE, SSIM, RRED, HLFI)	0.0	0.0	0.0

Table 12. Results (%) for fusion of IQMs on USTB Set 2 database

	USTB Set 2 Database		
Method Name	**FFR**	**FGR**	**HTER**
DLF (PSNR, GPE, SSIM, RRED, HLFI)	0.0	0.0	0.0
FLF (PSNR, GPE, SSIM, RRED, HLFI)	3.0	0.0	1.5
SLF-tanh (PSNR, GPE, SSIM, RRED, HLFI)	0.0	0.0	0.0
SLF-min-max (PSNR, GPE, SSIM, RRED, HLFI)	0.0	0.0	0.0
SLF-z-score (PSNR, GPE, SSIM, RRED, HLFI)	0.0	0.0	0.0

Table 13. Results (%) for fusion of IQMs on USTB Set 3 database

	USTB Set 3 Database		
Method Name	**FFR**	**FGR**	**HTER**
DLF (PSNR, GPE, SSIM, RRED, HLFI)	0.0	0.0	0.0
FLF (PSNR, GPE, SSIM, RRED, HLFI)	8.0	0.0	4.0
SLF-tanh (PSNR, GPE, SSIM, RRED, HLFI)	0.0	0.0	0.0
SLF-min-max (PSNR, GPE, SSIM, RRED, HLFI)	0.0	0.0	0.0
SLF-z-score (PSNR, GPE, SSIM, RRED, HLFI)	0.0	0.0	0.0

Table 14. Results (%) for fusion of IQMs on IITD database

	IITD Database		
Method Name	**FFR**	**FGR**	**HTER**
DLF (PSNR, GPE, SSIM, RRED, HLFI)	0.0	0.0	0.0
FLF (PSNR, GPE, SSIM, RRED, HLFI)	1.0	32.0	16.5
SLF-tanh (PSNR, GPE, SSIM, RRED, HLFI)	1.0	1.0	1.0
SLF-min-max (PSNR, GPE, SSIM, RRED, HLFI)	1.0	1.0	1.0
SLF-z-score (PSNR, GPE, SSIM, RRED, HLFI)	1.0	1.0	1.0

In order to have consistent and robust results for all datasets, fusion technique of IQA and CNN is proposed for ear anti-spoofing problem in this thesis. Results for implementation of the proposed method on AMI, UBEAR, IITD, USTB Set 1, Set 2 and Set 3 databases are shown in Table 15. As it is shown, HTER value of 0.0 has been obtained on AMI, UBEAR, IITD, USTB Set 1, Set 2 and Set 3, respectively. As a result, the third proposed method achieves the best performances for all datasets.

Table 15. Results (%) for the third proposed method (IQA + CNN)

	Proposed Method (IQA+CNN)		
Database Name	**FFR**	**FGR**	**HTER**
AMI	0.0	0.0	0.0
UBEAR	0.0	0.0	0.0
IITD	0.0	0.0	0.0
USTB Set 1	0.0	0.0	0.0
USTB Set 2	0.0	0.0	0.0
USTB Set 3	0.0	0.0	0.0

Table 16. Comparison with the state-of-the-art techniques using CNN

Methods Used	**Biometric Trait**	**Databases Used**	**#Subject**	**Size of Image**	**HTER**
CNN [51]	Face	Collected by them	20	256x256	1.2
		CASIA	50		2.3
CNN [52]	Face	CASIA-FASD	50	96x96	19.12
		Replay-Attack			8.39
CNN [53]	Face	CASIA-FASD	50	128x128	1.4
		Replay-Attack	50		1.2
		MSU	35		0.0
		Rose-Youtu	20		7.0
CNN [54]	Fingerprint	LivDet2013	-	200x200	3.7
		LivDet2011			6.45
BSIF+CNN [64]	Iris	Notre Dame	-	260x260	3.28
		Warsaw			0.68
		Clarkson			9.45
		IITD+WVU			14.92
IQA+CNN (DLF) (Third Proposed Method)	Ear	AMI	100	256x256	0.0
		UBEAR	100		0.0
		IITD	124		0.0
		USTB Set 1	60		0.0
		USTB Set 2	76		0.0
		USTB Set 3	78		0.0

The comparison results which have been made with the third proposed method and the other methods that are implemented based on deep CNN for countering against spoof attacks to person identification systems which are based on face, fingerprint and iris biometric traits are presented in Table 16. In the first study [51], the method is proposed for face anti-spoofing and the obtained error rates are 1.2 and 2.3 for their private database and CASIA, respectively. The second study [52] proposes an anti-spoofing method for face biometric trait and the obtained error rates are 19.12 and 8.39 for CASIA-FASD and Replay-Attack databases, respectively. In the third study [53], the proposed anti-spoofing method is for face biometric trait and the obtained error rates are 1.4, 1.2, 0.0 and 7.0 for CASIA-FASD, Replay-Attack, MSU and Rose-Youtu databases, respectively. Further, a novel method is implemented for fingerprint

biometric trait [54] and 3.7 and 6.45 error rates are obtained for LivDet2013 and LivDet2011 databases, respectively. Lastly, iris anti-spoofing CNN-based method is proposed in [64] and 3.28, 0.68, 9.45 and 14.92 error rates are obtained on Notre Dame, Warsaw, Clarkson and IITD+WVU databases, respectively. The comparison of the results of third method and the results of the state-of-the-art methods presents that, the third proposed method in this thesis achieves the best performances for all datasets with zero error rates.

6.3.3 Discussion on Experimental Results

In this thesis, the third proposed ear anti-spoofing method that is implemented to detect printed photo attacks by combining CNN and 5 IQM functions. Four of these IQMs are FR and one of them is NR. According to the preliminary experiments, results are not consistent for all datasets whenever deep learning and IQM methods are employed separately. Therefore, the fusion of CNN and IQA is proposed in this thesis. Comparison of the third proposed method with the state-of-the-art CNN-based studies show that the proposed method achieves the best performances for all datasets used in the experiments.

6.4 Combining Texture-Based Methods with Deep Learning for Ear Anti-Spoofing

In this section, texture-based BSIF method and deep learning based CNN model are exploited to propose the fourth novel anti-spoofing technique for ear identification systems. Currently, there have not been any study which focused on ear anti-spoofing problem by using texture-based and CNN-based methods in the literature. In this method, BSIF is employed to extract features by using ICA texture filters from ear images. Besides, CNN model is employed to learn deep representations of ear images in order to detect fake ear images. In order to propose a robust method, BSIF and CNN

methods are fused by implementing DLF technique. Corresponding experiments have been conducted on 6 different ear datasets namely, AMI, UBEAR, IITD, USTB Set 1, USTB Set 2 and USTB Set 3. The performance of the proposed method shows that it achieves to detect fake ear images. The proposed method is compared with the state-of-the-art methods that employed CNN on various biometric traits.

6.4.1 Methodology

The proposed method employs both CNN-based deep learning and texture-based BSIF method to implement a robust anti-spoofing technique for ear biometrics. General schema of the fourth proposed method is demonstrated in Figure 10. The major steps of BSIF part of the algorithm are as follows:

- Step 1: Enhance all ear images in train and test sets with histogram equalization and normalize them with mean variance normalization. Next, resize them to 256x256.
- Step 2: Convert all ear images into grayscale.
- Step 3: Divide each image into 8x8 blocks.
- Step 4: Compute BSIF code for each block with 7x7 filter. These operations are applied to all train and test ear images separately.
- Step 5: After feature extraction process, train and test sets are shuffled to obtain a realistic model.
- Step 6: Nearest Neighbor (NN) method is employed for classification part. The test ear image is compared with all train ear images and scores are obtained by calculating Manhattan distance. NN classifier is applied to determine whether ear image is genuine or impostor.

In addition to this, the detailed steps of CNN part are as follows:

- Step 1: Resize all ear images to 256x256 and convert them into grayscale.

- Step 2: Train and test sets are shuffled to obtain a realistic model.
- Step 3: In the CNN model, 5 convolutional layers are employed with 3x3 kernel map and ReLU activation function.
- Step 4: Max Pooling technique with 2x2 pooling size is applied to reduce the dimension of feature map.
- Step 5: In order to standardize the inputs of each convolutional layer and speed up training process, batch normalization is applied after pooling operation. Last obtained feature maps are flattened into vector form and it is called input layer. Further, dropout which its rate is set to 20% is applied before input layer to prevent the memorization of data.
- Step 6: The input layer is fed to hidden layer which contains 128 neurons to construct a fully-connected NN.
- Step 7: Lastly, Softmax classifier is employed in the output layer to classify ear test image as genuine or impostor.

CNN and BSIF parts have been implemented separately and two decisions have been obtained. In order to find a common decision, DLF with OR rule is applied. The logic of this rule is that if one of the method's decision is genuine, it will led final decision to be genuine, otherwise it is fake.

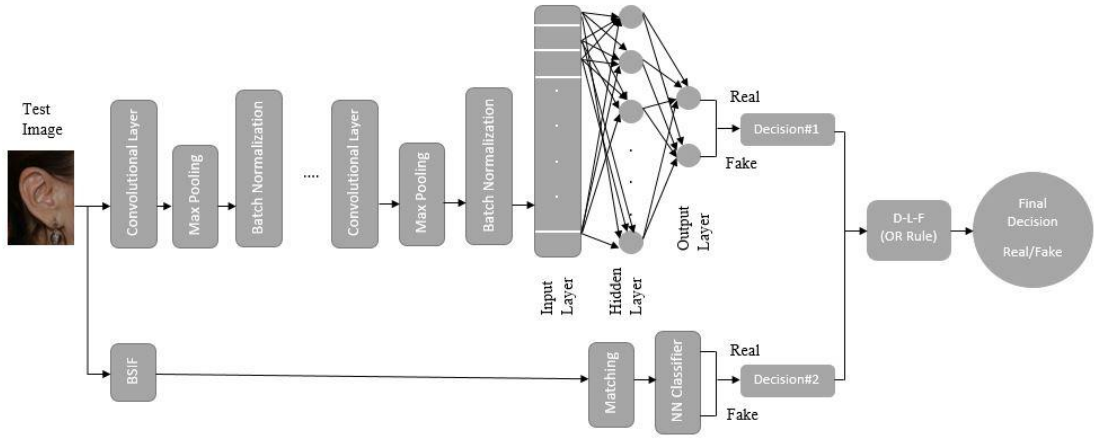

Figure 10. General block diagram of the fourth proposed ear anti-spoofing method

6.4.2 Experimental Results

Several experiments have been conducted to find the best solution to the problem of ear anti-spoofing. Firstly, BSIF has been implemented on 6 different datasets. As depicted in Table 17, Table 18 and Table 19, 1.5, 26.0, 1.0, 8.0, 1.0 and 17.0 error rates have been obtained for AMI, UBEAR, IITD, USTB Set 1, Set 2 and Set 3, respectively. Additionally, CNN-based method has been implemented and 0.5, 34.5, 1.0, 4.5, 0.5 and 0.0 error rates have been obtained for AMI, UBEAR, IITD, USTB Set 1, USTB Set 2 and USTB Set 3, respectively. CNN-based method achieves minimum error rates on AMI, USTB Set 1, Set 2 and Set 3 datasets. On the other hand, BSIF method achieves minimum error rates on UBEAR dataset. Beside on this, both methods achieve the same performance on IITD dataset. In order to achieve the best performance for all datasets, DLF of BSIF and CNN methods has been implemented. As a result, 0.0 error rates have been obtained for all datasets.

Table 17. Results (%) for each method on AMI and UBEAR databases

	AMI Database			**UBEAR Database**		
Method Name	**FFR**	**FGR**	**HTER**	**FFR**	**FGR**	**HTER**
BSIF	3.0	0.0	1.5	51.0	1.0	26.0
CNN	0.0	1.0	0.5	69.0	0.0	34.5
DLF (BSIF+CNN)	0.0	0.0	0.0	0.0	0.0	0.0

Table 18. Results (%) for each method on IITD and USTB Set 1 databases

	IITD Database			USTB Set 1 Database		
Method Name	**FFR**	**FGR**	**HTER**	**FFR**	**FGR**	**HTER**
BSIF	2.0	0.0	1.0	10.0	6.0	8.0
CNN	1.0	1.0	1.0	3.0	6.0	4.5
DLF (BSIF+CNN)	0.0	0.0	0.0	0.0	0.0	0.0

Table 19. Results (%) for each method on USTB Set 2 and USTB Set 3 databases

	USTB Set 2 Database			USTB Set 3 Database		
Method Name	**FFR**	**FGR**	**HTER**	**FFR**	**FGR**	**HTER**
BSIF	2.0	0.0	1.0	23.0	11.0	17.0
CNN	1.0	0.0	0.5	0.0	0.0	0.0
DLF (BSIF+CNN)	0.0	0.0	0.0	0.0	0.0	0.0

Additionally, state-of-the-art CNN-based anti-spoofing methods which are performed on iris, face and fingerprint biometric traits are compared with the proposed method since there are only two ear anti-spoofing studies in the literature [26, 27]. The summary of the comparison is shown in Table 20. The first study [59] proposes a CNN-based spoof detection algorithm which is conducted on Replay attack and NUAA databases and HTER values of 10.0 and 0.98 are obtained, respectively. Further, CNN-based study [52] on face biometric trait is conducted on CASIA-FASD and Replay attack databases and 19.12 and 8.39 HTER values are obtained, respectively. Additionally, a face anti-spoofing method [53] is proposed by using CNN model and conducted on CASIA-FASD, Replay attack, MSU and Rose-Youtu databases. In that method, HTER values are 1.4, 1.2, 0.0 and 7.0, respectively. On the other hand, a CNN-based method [54] is implemented for fake fingerprint detection algorithm. It is conducted on LivDet 2013 and 2011 databases and 3.7 and 6.45 HTER values are obtained. In the next study [58], a multi-channel CNN-based method is proposed and conducted on WMCA face spoof database. In that study, HTER value is 0.3. Further, in [64], BSIF and CNN are fused for iris presentation attack detection on

Notre Dame, Warsaw, Clarkson and IITD+WVU databases and 3.28, 0.68, 9.45 and 14.92 error rates are obtained. In [63], authors propose an iris presentation attack detection method that is based on CNN + SVM. In that study, 0.016 and 0.292 error rates are obtained for Warsaw-2017 and NDCLD-2015 databases, respectively. Finally, the fourth proposed method achieves superior results compared to the state-of-the-art anti-spoofing methods with all HTER values equal to 0.0 for all datasets.

Table 20. Comparison with the state-of-the-art studies using CNN

Methods Used	**Biometric Trait**	**Databases Used**	**#Subject**	**Size of Image**	**HTER**
CNN [59]	Face	Replay Attack	20	64x64	10.0
		NUAA	50		0.98
CNN [52]	Face	CASIA-FASD	50	96x96	19.12
		Replay-Attack			8.39
CNN [53]	Face	CASIA-FASD	50	128x128	1.4
		Replay-Attack	50		1.2
		MSU	35		0.0
		Rose-Youtu	20		7.0
CNN [54]	Fingerprint	LivDet2013	-	200x200	3.7
		LivDet2011			6.45
MC-CNN [58]	Face	WMCA	72	128x128	0.3
BSIF+CNN [64]	Iris	Notre Dame	-	260x260	3.28
		Warsaw			0.68
		Clarkson			9.45
		IITD+WVU			14.92
CNN+SVM [63]	Iris	Warsaw-2017	-	-	0.016
		NDCLD-2015			0.292
BSIF+CNN (Fourth Proposed Method)	Ear	AMI	100	256x256	0.0
		UBEAR	100		0.0
		IITD	124		0.0
		USTB Set 1	60		0.0
		USTB Set 2	76		0.0
		USTB Set 3	78		0.0

6.4.3 Discussion on Experimental Results

A robust ear anti-spoofing method is proposed by combining deep learning based CNN and texture-based BSIF methods. CNN and BSIF are implemented separately for feature extraction and representation of ear images. Experiments have been conducted on 6 different ear datasets which are namely AMI, UBEAR, IITD, USTB Set 1, USTB Set 2 and USTB Set 3. Combination of BSIF and CNN provides superior error rates on all aforementioned datasets. In the biometric community, ear anti-spoofing is not studied with CNN and BSIF methods. Consequently, the proposed method has a contribution on the improvement of ear anti-spoofing applications in the biometric community. Further, the proposed method is compared with the state-of-the-art anti-spoofing studies which are implemented for face, iris and fingerprint biometric traits by using CNN-based method and the results are still superior compared to the other biometric anti-spoofing methods.

6.5 Comparison of All Proposed Methods

The comparison of all proposed methods in terms of error rates and execution times is demonstrated in Table 21. The first and the second proposed method employ IQA technique basically. The first study [26], has been implemented on AMI and UBEAR databases which contain 50 subjects. As shown in Table 21, 8.5 and 15.5 HTER values have been obtained for AMI and UBEAR databases, respectively. The number of subjects for both databases have been incremented to 100 in the second study [27]. In that method which combines SLF and DLF techniques, 1.0 and 10.0 HTER values have been obtained for AMI and UBEAR databases, respectively. The datasets that are used for the first proposed method contain less number of subjects compared to the datasets used in the experiments of the second proposed method. Therefore, the first proposed method takes less execution times. Consequently, as it is observed in Table

21, the second method outperforms the first method. Additionally, the third and the fourth proposed method are conducted on 6 different ear databases which makes it different from previous experimental studies and instead of using IQA method only, texture-based and deep learning based methods are employed as well in the third and the fourth methods. The disadvantage of combining the IQA technique with BSIF or CNN is that it causes more execution time. On the other hand, advantage of combining IQA technique with BSIF or CNN is that it provides 0.0 error rates for all datasets for both of the third and the fourth proposed methods. Besides, the third proposed method takes less execution time compared to the fourth proposed method for all datasets. Therefore, it can be stated that the third proposed method is the best proposed method according to the error rates and execution time.

Table 21. Comparison of all proposed methods in terms of execution times and error rates

Method Name	**Databases Used**	**# of Subject**	**Size of Image**	**Execution Time (sec)**	**HTER**
IQA (DLF) [26] (First Proposed Method)	AMI	50	256x256	271	8.5
	UBEAR	50		270	15.5
IQA (SLF+DLF) [27] (Second Proposed Method)	AMI	100	256x256	567	1.0
	UBEAR	100		588	10.0
IQA+CNN (DLF) (Third Proposed Method)	AMI	100	256x256	18999	0.0
	UBEAR	100		19341	0.0
	IITD	124		1405	0.0
	USTB Set 1	60		574	0.0
	USTB Set 2	76		842	0.0
	USTB Set 3	78		1126	0.0
BSIF+CNN (DLF) (Fourth Proposed Method)	AMI	100	256x256	21518	0.0
	UBEAR	100		21795	0.0
	IITD	124		2420	0.0
	USTB Set 1	60		882	0.0
	USTB Set 2	76		1295	0.0
	USTB Set 3	78		1343	0.0

Chapter 7

CONCLUSION

In this thesis, the aim is to propose a novel method for ear anti-spoofing problem. In this context, four ear anti-spoofing methods are proposed. The experiments for these methods are conducted on six different datasets namely, AMI, UBEAR, IITD, USTB Set1, USTB Set 2 and USTB Set 3.

In the first method, FR and NR IQA measures are combined by using DLF technique for ear anti-spoofing problem. In that method, 8.5 and 15.5 error rates have been obtained on AMI and UBEAR databases, respectively.

In order to enhance the performance of the first proposed method, the second method is proposed. The second method employs SLF to combine the matching scores obtained from FR and NR IQA measures. Afterwards, DLF is employed make a final decision. In that method, the experiments have been conducted on databases that are used in the first method as well. The difference is that the number of used subjects is increased from 50 to 100. The obtained error rates are 1.0 and 10.0 for AMI and UBEAR databases, respectively. Consequently, it can easily be observed that the performance of the second proposed method is getting better than the first proposed method. Additionally, SLF and DLF have been combined to propose robust ear anti-spoofing method. Last but not least, CNN-based deep learning method has been firstly applied on ear anti-spoofing systems and compared with the second proposed method.

According to the comparison, the second proposed method achieves better performance.

Further, the third proposed method employs DLF to combine IQA-based and CNN-based methods to detect fake ear images. In this method, 4 FR IQA measures that are PSNR, GPE, SSIM and RRED and 1 NR IQA measure that is HLFI are employed to extract features from ear images. Additionally, CNN-based method is employed to find deeper representation of ear images. Afterwards, the results of these methods are combined by using DLF with OR rule. The experiments have been conducted on five different datasets which are AMI, UBEAR, IITD, USTB Set 1, USTB Set 2 and USTB Set 3. The obtained error rates are 0.0 for all aforementioned databases. The first contribution of this method is that the combination of CNN-based and IQA-based methods have been implemented the first time in this method. Additionally, the fusion of 5 image quality metrics for the detection of spoofing attacks for ear biometrics through FLF, DLF and SLF strategies are presented. Finally, ear anti-spoofing system results are demonstrated the first time on six ear datasets in this study.

The last proposed method combines the texture-based BSIF and CNN-based methods to counter against print attack on ear recognition systems. The fusion of these methods is achieved by DLF technique. The combination of BSIF and CNN is implemented the first time for ear anti-spoofing systems in this method. Additionally, this method has been applied on six different datasets namely, AMI, UBEAR, IITD, USTB Set 1, USTB set 2 and USTB Set 3. According to the comparison of this method with the first and the second proposed methods, it can obviously be observed that the last proposed method that is DLF of BSIF and CNN achieves better performances on AMI and UBEAR databases. However, the comparison of the third and the fourth proposed

methods in terms of error rates shows that the error rates of both of these methods are superior, but the third proposed method outperforms the fourth proposed method in terms of execution times. Consequently, the third proposed method is the best method proposed in this thesis.

As a future work, different types of attacks such as replay video attack and digital photo attack can be investigated on ear biometrics. Further, ear anti-spoofing system can be integrated to face anti-spoofing system to develop a more robust anti-spoofing system. Beside face biometric trait, ear images can be fused with other biometric traits such as iris or fingerprint as well.

REFERENCES

[1] Emeršič, Ž., Štruc, V., and Peer, P. (2017). Ear recognition: More than a survey. Neurocomputing, 255, pp. 26-39.

[2] Alqaralleh, E., and Toygar, Ö. (2018). Ear recognition based on fusion of ear and tragus under different challenges. *International Journal of Pattern Recognition and Artificial Intelligence,* 32 (09), p. 1856009.

[3] Toygar, Ö., Alqaralleh, E., and Afaneh, A. (2018). Symmetric ear and profile face fusion for identical twins and non-twins recognition. *Signal, Image and Video Processing,* 12 (6), pp. 1157-1164.

[4] Hassaballah, M., Alshazly, H. A., and Ali, A. A. (2019). Ear recognition using local binary patterns: A comparative experimental study. *Expert Systems with Applications,* 118, pp. 182-200.

[5] Islam, S. M., Davies, R., Bennamoun, M., and Mian, A. S. (2011). Efficient detection and recognition of 3D ears. *International Journal of Computer Vision,* 95 (1), pp. 52-73.

[6] Alagarsamy, S. B., and Kondappan, S. (2018). Ear recognition system using adaptive approach Runge–Kutta (AARK) threshold segmentation with ANFIS classification. *Neural Computing and Applications,* pp. 1-12.

[7] Ganapati, I. I., Ali, S. S., and Prakash, S. (2018). Geometric statistics-based descriptor for 3D ear recognition. *The Visual Computer*, pp. 1-13.

[8] Hourali, F., and Gharravi, S. (2017). An Ear Recognition Method Based on Rotation Invariant Transformed DCT. *International Journal of Electrical and Computer Engineering,* 7 (5), p. 2895.

[9] Sarangi, P. P., Mishra, B. S. P., and Dehuri, S. (2019). Fusion of PHOG and LDP local descriptors for kernel-based ear biometric recognition. *Multimedia Tools and Applications,* 78 (8), pp. 9595-9623.

[10] Yuan, L., and chun Mu, Z. (2012). Ear recognition based on local information fusion. *Pattern Recognition Letters,* 33 (2), pp. 182-190.

[11] Omara, I., Li, F., Zhang, H., and Zuo, W. (2016). A novel geometric feature extraction method for ear recognition. *Expert Systems with Applications,* 65, pp. 127-135.

[12] Toygar, Ö., Alqaralleh, E., and Afaneh, A. (2019). On the use of ear and profile faces for distinguishing identical twins and nontwins. *Expert Systems,* p. e12389.

[13] Nixon, K. A., Aimale, V., and Rowe, R. K. (2008). Spoof detection schemes. *In Handbook of biometrics* (pp. 403-423). Springer, Boston, MA.

[14] Rakshit, R. D. and Kisku, D. R. (2017). Face Spoofing and Counter-Spoofing: A Survey of State-of-the-art Algorithms. *Transaction on Machine Learning and Artificial Intelligence,* 5 (2), pp. 31-73.

[15] Galbally, J., Marcel, S., and Fierrez, J. (2014). Biometric antispoofing methods: A survey in face recognition. *IEEE Access*, 2, pp. 1530-1552.

[16] Ren, Y., Fang, Z., Liu, D., and Chen, C. (2019). Replay attack detection based on distortion by loudspeaker for voice authentication. *Multimedia Tools and Applications*, 78 (7), pp. 8383-8396.

[17] Galbally, J., Marcel, S., and Fierrez, J. (2014). Image quality assessment for fake biometric detection: Application to iris, fingerprint, and face recognition. *IEEE transactions on image processing*, 23 (2), pp. 710-724.

[18] Farmanbar, M., and Toygar, Ö. (2017). Spoof detection on face and palmprint biometrics. *Signal, Image and Video Processing*, 11 (7), pp. 1253-1260.

[19] Sousedik, C., and Busch, C. (2014). Presentation attack detection methods for fingerprint recognition systems: a survey. *Iet Biometrics*, 3 (4), 219-233.

[20] Ramachandram, D., and Taylor, G. W. (2017). Deep multimodal learning: A survey on recent advances and trends. *IEEE Signal Processing Magazine*, 34 (6), pp. 96-108.

[21] Komogortsev, O. V., Karpov, A., and Holland, C. D. (2015). Attack of mechanical replicas: Liveness detection with eye movements. *IEEE Transactions on Information Forensics and Security*, 10 (4), pp. 716-725.

[22] Qiu, X., Kang, W., Tian, S., Jia, W., and Huang, Z. (2017). Finger vein presentation attack detection using total variation decomposition. *IEEE Transactions on Information Forensics and Security,* 13 (2), pp. 465-477.

[23] Tome, P., and Marcel, S. (2015, May). On the vulnerability of palm vein recognition to spoofing attacks. *In 2015 International Conference on Biometrics (ICB)* (pp. 319-325). IEEE.

[24] Nikam, S. B., and Agarwal, S. (2009). Ridgelet-based fake fingerprint detection. *Neurocomputing,* 72 (10-12), pp. 2491-2506.

[25] Bhilare, S., Kanhangad, V., and Chaudhari, N. (2018). A study on vulnerability and presentation attack detection in palmprint verification system. *Pattern Analysis and Applications,* 21 (3), pp. 769-782.

[26] Toprak, İ., and Toygar, Ö. (2018, May). Fusion of Full-Reference and No-Reference Anti-Spoofing Techniques for Ear Biometrics under Print Attacks. International Conference on Advanced Technologies, Computer Engineering and Science (ICATCES) (pp. 538-542).

[27] Toprak, İ., and Toygar, Ö. (2019). Ear anti-spoofing against print attacks using three-level fusion of image quality measures. *Signal, Image and Video Processing*, pp. 1-8. DOI: https://doi.org/10.1007/s11760-019-01570-w

[28] LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *nature,* 521 (7553), p. 436.

[29] Kannala, J., and Rahtu, E. (2012, November). Bsif: Binarized statistical image features. *In Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)* (pp. 1363-1366). IEEE.

[30] Gonzalez, E., Alvarez, L. and Mazorra, L. AMI Ear Database. Retrieved Jan 10, 2018 from http://www.ctim.es/research_works/ami_ear_database

[31] Raposo, R., Hoyle, E., Peixinho, A., and Proença, H. (2011, April). Ubear: A dataset of ear images captured on-the-move in uncontrolled conditions. *In 2011 IEEE Workshop on Computational Intelligence in Biometrics and Identity Management (CIBIM)* (pp. 84-90). IEEE.

[32] Kumar, A., and Wu, C. (2012). Automated human identification using ear imaging. *Pattern Recognition,* 45 (3), pp. 956-968.

[33] Mu, Z. and Yuan, L. Ear Recognition Laboratory at USTB. Retrieved Dec 15, 2018 from http://www1.ustb.edu.cn/resb/en/index.htm

[34] Czajka, A., and Bowyer, K. W. (2018). Presentation attack detection for iris recognition: An assessment of the state-of-the-art. *ACM Computing Surveys (CSUR),* 51 (4), p. 86.

[35] Wen, D., Han, H., and Jain, A. K. (2015). Face spoof detection with image distortion analysis. *IEEE Transactions on Information Forensics and Security,* 10 (4), pp. 746-761.

[36] Galbally, J., Alonso-Fernandez, F., Fierrez, J., and Ortega-Garcia, J. (2012). A high performance fingerprint liveness detection method based on quality related features. *Future Generation Computer Systems,* 28 (1), pp. 311-321.

[37] Galbally, J., Ortiz-Lopez, J., Fierrez, J., and Ortega-Garcia, J. (2012, March). Iris liveness detection based on quality related features. *In 2012 5th IAPR International Conference on Biometrics (ICB)* (pp. 271-276). IEEE.

[38] Aishwarya, D., Gowri, M., and Saranya, R. K. (2016, March). Palm print recognition using liveness detection technique. *In 2016 Second International Conference on Science Technology Engineering and Management (ICONSTEM)* (pp. 109-114). IEEE.

[39] Pravallika, P., and Prasad, K. S. (2016, August). SVM classification for fake biometric detection using image quality assessment: Application to iris, face and palm print. *In 2016 International Conference on Inventive Computation Technologies (ICICT)* (Vol. 1, pp. 1-6). IEEE.

[40] Raghavendra, R., and Busch, C. (2015). Robust scheme for iris presentation attack detection using multiscale binarized statistical image features. *IEEE Transactions on Information Forensics and Security,* 10 (4), pp. 703-715.

[41] Boulkenafet, Z., Komulainen, J., and Hadid, A. (2016). Face spoofing detection using colour texture analysis. *IEEE Transactions on Information Forensics and Security,* 11 (8), pp. 1818-1830.

[42] Määttä, J., Hadid, A., and Pietikäinen, M. (2012). Face spoofing detection from single images using texture and local shape analysis. *IET biometrics,* 1 (1), pp. 3-10.

[43] Beham, M. P., and Roomi, S. M. M. (2018). Anti-spoofing enabled face recognition based on aggregated local weighted gradient orientation. *Signal, Image and Video Processing,* 12 (3), pp. 531-538.

[44] Freitas Pereira, T. d., Komulainen, J., Anjos, A., De Martino, J. M., Hadid, A., Pietikäinen, M., and Marcel, S. (2014). Face liveness detection using dynamic texture. *EURASIP Journal on Image and Video Processing,* 2014 (1), 2.

[45] Boulkenafet, Z., Komulainen, J., and Hadid, A. (2018). On the generalization of color texture-based face anti-spoofing. *Image and Vision Computing,* 77, pp. 1-9.

[46] Bhilare, S., Kanhangad, V., and Chaudhari, N. (2018). A study on vulnerability and presentation attack detection in palmprint verification system. *Pattern Analysis and Applications,* 21 (3), pp. 769-782.

[47] Arashloo, S. R., Kittler, J., and Christmas, W. (2015). Face spoofing detection based on multiple descriptor fusion using multiscale dynamic binarized statistical image features. *IEEE Transactions on Information Forensics and Security,* 10 (11), pp. 2396-2407.

[48] Ghiani, L., Hadid, A., Marcialis, G. L., and Roli, F. (2013, September). Fingerprint liveness detection using binarized statistical image features. *In 2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS)* (pp. 1-6). IEEE.

[49] Li, Q., and Chan, P. P. (2014, July). Fingerprint liveness detection based on binarized statistical image feature with sampling from Gaussian distribution. *In 2014 International Conference on Wavelet Analysis and Pattern Recognition* (pp. 13-17). IEEE.

[50] Li, X., Bu, W., and Wu, X. (2015, November). Palmprint liveness detection by combining binarized statistical image features and image quality assessment. *In Chinese Conference on Biometric Recognition* (pp. 275-283). Springer, Cham.

[51] Wang, Y., Nian, F., Li, T., Meng, Z., and Wang, K. (2017). Robust face anti-spoofing with depth information. *Journal of Visual Communication and Image Representation*, 49, pp. 332-337.

[52] Rehman, Y. A. U., Po, L. M., and Liu, M. (2018). LiveNet: Improving features generalization for face liveness detection using convolution neural networks. *Expert Systems with Applications*, 108, pp. 159-169.

[53] Li, H., He, P., Wang, S., Rocha, A., Jiang, X., and Kot, A. C. (2018). Learning generalized deep feature representation for face anti-spoofing. *IEEE Transactions on Information Forensics and Security*, 13 (10), pp. 2639-2652.

[54] Yuan, C., Xia, Z., Jiang, L., Cao, Y., Wu, Q. J., and Sun, X. (2019). Fingerprint Liveness Detection Using an Improved CNN With Image Scale Equalization. *IEEE Access,* 7, pp. 26953-26966.

[55] Czajka, A., Bowyer, K. W., Krumdick, M., and VidalMata, R. G. (2017). Recognition of image-orientation-based iris spoofing. *IEEE Transactions on Information Forensics and Security,* 12 (9), pp. 2184-2196.

[56] Raghavendra, R., Venkatesh, S., Raja, K. B., and Busch, C. (2017, April). Transferable deep convolutional neural network features for fingervein presentation attack detection. *In 2017 5th International Workshop on Biometrics and Forensics (IWBF)* (pp. 1-5). IEEE.

[57] Dinkel, H., Qian, Y., and Yu, K. (2018). Investigating Raw Wave Deep Neural Networks for End-to-End Speaker Spoofing Detection. *IEEE/ACM Transactions on Audio, Speech, and Language Processing,* 26 (11), pp. 2002-2014.

[58] George, A., Mostaani, Z., Geissenbuhler, D., Nikisins, O., Anjos, A., and Marcel, S. (2020). Biometric Face Presentation Attack Detection with Multi-Channel Convolutional Neural Network. *IEEE Transactions on Information Forensics and Security,* 15, pp. 42-55.

[59] Alotaibi, A., and Mahmood, A. (2017). Deep face liveness detection based on nonlinear diffusion using convolution neural network. *Signal, Image and Video Processing,* 11 (4), pp. 713-720.

[60] Kim, S., Park, B., Song, B. S., and Yang, S. (2016). Deep belief network based statistical feature learning for fingerprint liveness detection. *Pattern Recognition Letters,* 77, pp. 58-65.

[61] Chugh, T., Cao, K., and Jain, A. K. (2018). Fingerprint spoof buster: Use of minutiae-centered patches. *IEEE Transactions on Information Forensics and Security,* 13 (9), pp. 2190-2202.

[62] Qiu, X., Tian, S., Kang, W., Jia, W., and Wu, Q. (2017, October). Finger Vein Presentation Attack Detection Using Convolutional Neural Networks. *In Chinese Conference on Biometric Recognition* (pp. 296-305). Springer, Cham.

[63] Nguyen, D., Pham, T., Lee, Y., and Park, K. (2018). Deep learning-based enhanced presentation attack detection for iris recognition by combining features from local and global regions based on NIR camera sensor. *Sensors,* 18 (8), p. 2601.

[64] Kuehlkamp, A., Pinto, A., Rocha, A., Bowyer, K. W., and Czajka, A. (2018). Ensemble of multi-view learning classifiers for cross-domain iris presentation attack detection. *IEEE Transactions on Information Forensics and Security,* 14 (6), pp. 1419-1431.

[65] Nguyen, D. T., Pham, T. D., Baek, N. R., and Park, K. R. (2018). Combining deep and handcrafted image features for presentation attack detection in face recognition systems using visible-light camera sensors. *Sensors,* 18 (3), p. 699.

[66] De Souza, G. B., da Silva Santos, D. F., Pires, R. G., Marana, A. N., and Papa, J. P. (2017). Deep texture features for robust face spoofing detection. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 64 (12), pp. 1397-1401.

[67] Feng, L., Po, L. M., Li, Y., Xu, X., Yuan, F., Cheung, T. C. H., and Cheung, K. W. (2016). Integration of image quality and motion cues for face anti-spoofing: A neural network approach. *Journal of Visual Communication and Image Representation*, 38, pp. 451-460.

[68] Edmunds, T., and Caplier, A. (2018). Motion-based countermeasure against photo and video spoofing attacks in face recognition. *Journal of Visual Communication and Image Representation,* 50, pp. 314-332.

[69] Gragnaniello, D., Sansone, C., and Verdoliva, L. (2015). Iris liveness detection for mobile devices based on local descriptors. *Pattern Recognition Letters*, 57, pp. 81-87.

[70] Singh, M., and Arora, A. S. (2017). A robust anti-spoofing technique for face liveness detection with morphological operations. *Optik,* 139, pp. 347-354.

[71] Avcıbaş, İ. (2001). Image Quality Statistics and their use in Steganalysis and Compression. Institute for Graduate Studies in Science and Engineering in partial fulfillment of the requirements for the degree of Doctor of Philosophy, Boğaziçi University.

[72] Avcibas, I., Sankur, B., and Sayood, K. (2002). Statistical evaluation of image quality measures. *Journal of Electronic imaging,* 11 (2), pp. 206-224.

[73] Huynh-Thu, Q., and Ghanbari, M. (2008). Scope of validity of PSNR in image/video quality assessment. *Electronics letters,* 44 (13), pp. 800-801.

[74] Yao, S., Lin, W., Ong, E., and Lu, Z. (2005, September). Contrast signal-to-noise ratio for image quality assessment. *In IEEE International Conference on Image Processing 2005* (Vol. 1, pp. I-397). IEEE.

[75] Eskicioglu, A. M., and Fisher, P. S. (1995). Image quality measures and their performance. *IEEE Transactions on Communications,* 43 (12), pp. 2959-2965.

[76] Martini, M. G., Hewage, C. T., and Villarini, B. (2012). Image quality assessment based on edge preservation. *Signal Processing: Image Communication,* 27 (8), pp. 875-882.

[77] Harris, C. G., and Stephens, M. (1988, August). A combined corner and edge detector. *In Alvey Vision Conference* (Vol. 15, No. 50, pp. 10-5244).

[78] Nill, N. B., and Bouzas, B. (1992). Objective image quality measure derived from digital image power spectra. *Optical engineering,* 31 (4), pp. 813-826.

[79] Liu, A., Lin, W., and Narwaria, M. (2011). Image quality assessment based on gradient similarity. *IEEE Transactions on Image Processing,* 21 (4), pp. 1500-1512.

[80] Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing,* 13 (4), pp. 600-612.

[81] Sheikh, H. R., and Bovik, A. C. (2006). Image information and visual quality. *IEEE Transactions on Image Processing,* 15 (2), pp. 430-444.

[82] Soundararajan, R., and Bovik, A. C. (2011). RRED indices: Reduced reference entropic differencing for image quality assessment. *IEEE Transactions on Image Processing,* 21 (2), pp. 517-526.

[83] Wang, Z., Sheikh, H. R., and Bovik, A. C. (2002, September). No-reference perceptual quality assessment of JPEG compressed images. *In Proceedings. International Conference on Image Processing* (Vol. 1, pp. I-I). IEEE.

[84] Zhu, X., and Milanfar, P. (2009, July). A no-reference sharpness metric sensitive to blur and noise. *In 2009 International Workshop on Quality of Multimedia Experience* (pp. 64-69). IEEE.

[85] Moorthy, A. K., and Bovik, A. C. (2010). A two-step framework for constructing blind image quality indices. *IEEE Signal Processing Letters,* 17 (5), pp. 513-516.

[86] Mittal, A., Soundararajan, R., and Bovik, A. C. (2012). Making a "completely blind" image quality analyzer. *IEEE Signal Processing Letters,* 20 (3), pp. 209-212.

[87] LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE,* 86 (11), pp. 2278-2324.

[88] Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J. and Chen, T. (2018). Recent advances in convolutional neural networks. *Pattern Recognition,* 77, pp. 354-377.

[89] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research,* 15 (1), pp. 1929-1958.

[90] Ioffe, S., and Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:*1502.03167.

[91] Kannala, J., and Rahtu, E. (2012, November). Bsif: Binarized statistical image features. *In Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)* (pp. 1363-1366). IEEE.

[92] Ouamane, A., Benakcha, A., Belahcene, M., and Taleb-Ahmed, A. (2015). Multimodal depth and intensity face verification approach using LBP, SLF, BSIF, and LPQ local features fusion. *Pattern Recognition and Image Analysis,* 25 (4), pp. 603-620.

[93] Raghavendra, R., and Busch, C. (2015). Robust scheme for iris presentation attack detection using multiscale binarized statistical image features. *IEEE Transactions on Information Forensics and Security,* 10 (4), pp. 703-715.

[94] Ross, A., and Jain, A. (2003). Information fusion in biometrics. *Pattern Recognition Letters,* 24 (13), pp. 2115-2125.

[95] Ross, A., Nandakumar, K., and Jain, A. Level of Fusion in Biometrics, *Handbook of Multibiometrics.* Springer, Boston, MA (2006)

[96] Parveen, S., Ahmad, S. M. S., Hanafi, M. and Adnan, W. A. W. (2015). Face anti-spoofing methods. *Current Science,* 108 (8), pp. 1491-1500.

[97] Jain, A., Nandakumar, K., and Ross, A. (2005). Score normalization in multimodal biometric systems. *Pattern Recognition,* 38 (12), pp. 2270-2285.

[98] Nguyen, D. T., Park, Y. H., Shin, K. Y., Kwon, S. Y., Lee, H. C., and Park, K. R. (2013). Fake finger-vein image detection based on Fourier and wavelet transforms. *Digital Signal Processing,* 23 (5), pp. 1401-1413.