FIRAT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
TÜRKİYE

A ROBUST ACTIVE-BASED FACE SPOOF DETECTION FOR FACE RECOGNITION SYSTEMS

Peter ANTHONY

Master's Thesis

DEPARTMENT OF COMPUTER ENGINEERING

FEBRUARY 2022

FIRAT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
TÜRKİYE

Department of Computer Engineering

Master's Thesis

A ROBUST ACTIVE-BASED FACE SPOOF DETECTION FOR FACE RECOGNITION SYSTEMS

Author
Peter ANTHONY

Supervisor
Assist. Prof. Dr Betul AY

FEBRUARY 2022
ELAZIG

FIRAT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
TÜRKİYE

Department of Computer Engineering

Master's Thesis

Title:	A Robust Active-based Face Spoof Detection for Face Recognition Systems
Author:	Peter ANTHONY
Submission Date:	07 January 2022
Defense Date:	07 February 2022

THESIS APPROVAL

This thesis, which was prepared according to the thesis writing rules of the Graduate School of Natural and Applied Sciences, Fırat University, was evaluated by the committee members who have signed the following signatures and was commonly approved after the defense exam made open to the academic audience.

Signature

Supervisor:	Assist. Prof. Dr Betul AY Firat University, Faculty of Engineering	Approved
Chair:	Assoc. Prof. Dr Galip AYDIN Firat University, Faculty of Engineering	Approved
Member:	Assoc. Prof. Dr Mehmet BAYGIN Ardahan University, Faculty of Engineering	Approved

This thesis was approved by the Administrative Board of the Graduate School on

..... / ..... / 20 .....

Signature

Prof. Dr. Kürşat Esat ALYAMAÇ
Director of the Graduate School

DECLARATION

I hereby declare that I wrote this Master's Thesis titled “A Robust Active-based Face Spoof Detection for Face Recognition Systems” in consistent with the thesis writing guide of the Graduate School of Natural and Applied Sciences, Firat University. I also declare that all information in it is correct, that I acted according to scientific ethics in producing and presenting the findings, cited all the references I used, express all institutions or organizations or persons who supported the thesis financially. I have never used the data and information I provide here in order to get a degree in any way.

07 February 2022

Peter ANTHONY

PREFACE

An ordinary face recognition system without a concrete and deliberate spoof detection measures (as the case is with many deployed face recognition systems) is prone to failure if there be spoof attack attempts on such a system. Hence the need for a robust anti-spoofing measure.

Drawing inspiration from the common phenomenon in photography on distortion of faces in video when the camera is close to the face, which is caused by the uneven 3D surface of the human face, this study utilizes the distortion changes of video frames of users face captured at different distance from the camera to deliver a robust active-based spoof detection, yet less intrusive, less expensive and a more generic applicability than other active-based approaches using both traditional machine learning classifiers and a deep learning model.

This thesis is sponsored by the National Information Technology Development Fund (NITDEF) scholarship scheme 2019/2020 under the Nigeria National Information Technology Development Agency (NITDA).

Peter ANTHONY
ELAZIG, 2022

TABLE OF CONTENTS

ABSTRACT

A Robust Active-based Face Spoof Detection for Face Recognition Systems

Peter ANTHONY

Master's Thesis

FIRAT UNIVERSITY
Graduate School of Natural and Applied Sciences

Department of Computer Engineering

February 2022, Page: x + 54

With increasing rate of deployment of face recognition systems in various real life cases and applications, efforts by attackers is also on the rise with variety of spoofing approaches emerging. Hence, the development of a robust spoof detection technique is crucial. Although active-based techniques have proven robust in the task of spoof detection, they are faced with the problem of intrusiveness, high cost, computational complexity, low generalization ability and most of which require extra hardware.

This study proposed an active-based robust spoof detection technique capable of detecting diverse forms of media or 2D attacks, yet, less intrusive, less expensive, low complexity, higher generalization ability than other active-based approaches. It does not require extra hardware and so can easily fit into existing systems. Spoof detection is achieved by analyzing the distortion changes of Video frames of user's face captured at different distances from the camera. Real-world facial photo and video data were also collected and used to generate both the legitimate and spoof attacks dataset. Utilizing both machine learning classifiers and deep learning model, the proposed approach achieved spoof detection with accuracy as high as 98.18%, with EER and HTER as low as 0.23 and 0.021 respectively.

Keywords: Anti-spoofing, Spoof Detection, Face Recognition, Image Distortion Analysis

ÖZET

Yüz Tanıma Sistemleri İçin Güçlü, Aktif Tabanlı Yüz Sahtekarlığı Algılama

Peter ANTHONY

Yüksek Lisans Tezi

FIRAT ÜNİVERSİTESİ
Fen Bilimleri Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı

Şubat 2022, Sayfa: x + 54

Yüz tanıma sistemlerinin çeşitli gerçek yaşam durumlarında ve uygulamalarında artan oranda konuşlandırılmasıyla birlikte, ortaya çıkan çeşitli sahtekarlık yaklaşımları ile saldırganların çabaları da artıyor. Bu nedenle, sağlam bir sahtecilik algılama tekniğinin geliştirilmesi çok önemlidir. Aktif tabanlı teknikler, sahtekarlık tespiti görevinde sağlam olduklarını kanıtlamış olsalar da, müdahalecilik, yüksek maliyet, hesaplama karmaşıklığı, düşük genelleme yeteneği ve çoğu ekstra donanım gerektiren sorunlarla karşı karşıyadırlar.

Bu çalışma, diğer aktif tabanlı yaklaşımlardan daha az müdahaleci, daha az pahalı, düşük karmaşıklık, daha yüksek genelleme kabiliyetine sahip, çeşitli ortam biçimlerini veya 2B saldırıları tespit edebilen aktif tabanlı sağlam bir sahtekarlık algılama tekniği önerdi. Ekstra donanım gerektirmez ve bu nedenle mevcut sistemlere kolayca uyum sağlar. Sahtekarlık tespiti, kameradan farklı mesafelerde yakalanan kullanıcının yüzünün Video karelerindeki bozulma değişiklikleri analiz edilerek elde edilir. Gerçek dünyadaki yüz fotoğrafı ve video verileri de toplandı ve hem meşru hem de sahte saldırı veri setini oluşturmak için kullanıldı. Hem makine öğrenimi sınıflandırıcılarını hem de derin öğrenme modelini kullanan önerilen yaklaşım, sırasıyla 0,23 ve 0,021 kadar düşük EER ve HTER ile %98,18'e kadar yüksek bir doğrulukla yanıltıcı algılama elde etti.

Anahtar Kelimeler: Sahtekarlık Önleme, Kimlik Sahtekarlığı Algılama, Yüz Tanıma, Görüntü Bozulma Analizi

LIST OF FIGURES

Page

LIST OF TABLES

Page

SYMBOLS AND ABBREVIATIONS

Symbols

μE : Micro Expression

σ : Variance

Abbreviations

FAR : False Acceptance Rate
FRR : False Rejection Rate
EER : Equal Error Rate
HTER : Half Total Error Rate

1. INTRODUCTION

Face recognition is one of the most common biometric system today, and over the past decade it has undergone a very high and rapid increase in prevalence globally. According to the International Biometric Group (IBG), facial recognition is the second most widely deployed biometric system worldwide next to fingerprint [1]**.** This growing prevalence is due to its advantage of being more convenient, contactless, and non-invasive than other biometric systems such as fingerprint and iris [2,3]. Face recognition also holds the credit of being a biometric system with higher working range and application tendencies than others [4]. Unfortunately, bulk of the existing face recognition systems are unsafe when presented with spoof attacks. Face spoofing is the activity of disguising an individual's face in order to gain an illegal access or privileges.

With the increasing rate of deployment of face recognition systems in various real life cases and applications, efforts by attackers is also on the rise with variety of spoofing approaches emerging. Hence, the development of face anti-spoof approaches has become an active and crucial area of research.

This study focused on developing an active-based robust face anti-spoof technique capable of detetcting all forms of media or 2D attacks. This study also achieved an active liveness detection approach which does not suggest to the user the livesness cue employed. Drawing inspiration from the common phenomenon in photography on distortion of faces in video when the camera is close to the face, which is caused by the uneven 3D surface of the human face, in the approach proposed in this study, distortion changes of video frames of users face captured at different distance from the camera were utilized to deliver a robust active spoof detection, yet less intrusive, less expensive and a more generic applicability than other active-based approaches using both traditional machine learning classifiers and a deep learning model.

A variety of approaches has been proposed by researchers for spoof detections, ranging from the employment of motion cues [5–7], to the current deep learning approaches [8,9]with competitive performance. Face spoof detection approaches can be broadly classified as active and passive methods. The difference between the two categories is their mode of enrollment or the approach used to capture the user's data. Active approaches have proven to be very robust with strong protection ability against spoofing. However, active approaches require user interaction which makes it intrusive and lower the range of use cases. Some active approaches such as multi-modal [10,11] and Sensor level approaches [7] can be both computationally complex and expensive as they require extra hardware. Here a semi-active approach is proposed, with reduced intrusiveness/interaction, higher application range, with no extra hardware requirement and so less expensive.

In this work, the potentials of image distortion due to proximity of the face to camera in combatting face spoof attacks is explored. The basis for using the image distortion analysis (IDA) features for spoof detection is geared by three reasons:

- The common phenomenon in photography on distortion of faces in video where the camera is very near to the face, which is due to the uneven 3D-surface of the human face, as opposed to flat surface of 2D face spoof attacks.
- Image distortion has gain a lot of success in the field of biometrics and security systems, for finger print spoof detection [12], steganalysis [13,14], and face liveness detection [2,15].
- As the camera draws closer to an object, that object increases in size more quickly than objects in the background. In a study by Bryson et al [16], scatter plot of measured pixel intensity versus distance to camera reveals an exponential falloff in intensity with distance from the camera. This is because the object and the background image forms uneven 3D-surface. Whereas 2D attacks have flat surfaces, hence a different observation is expected when captured from the camera.

The employment of image distortion features as a descriptor for face spoof detection has been broadly used. Authors have come up with diverse techniques based on image distortion analysis (IDA) and have achieved promising results.

In [2], to build a technique against printed photo and replay attacks, authors combined four(4) different distortion features, including color diversity, blurriness features, specular reflectance features, and chromatic moment, to form the image distortion analysis feature vector. This features were trained using ensemble classifier made up of multiple support vector machines (SVM) for distinguishing real and fake faces. Authors in [15] also extracted the same features based on the four IDA features from the MSU-MFSD dataset which consist of printed photo attacks and replay attacks. Training using support vector machines and an artificial neural network, gave an accuracy of 94.4% and 88.9% respectively. Though they achieved a non-intrusive approach, but performance in spoof detection is not robust enough for real-application.

Several active-based approaches have also been proposed. Authors in [17] employed the use of an active near-infrared (NIR) light, they constructed NIR differential (NIRD) images. They exploited Two different characteristics based on NIRD images to achieve spoof detection. The consistency of the pixel between face and non-face regions were analyzed, and combine with context clues for discriminating spoofed faces from genuine one. Although their approach was robust against spoof detection, it is limited in generalization ability as it requires extra hardware. It is also highly intrusive. In [18] authors proposed a method based on multi-spectral using both visible (VIS) and near-infrared (NIR) spectra imaging. In [19] authors combine face and voice,

where user is requested to speak. In [5], a special hardware used for tracking pupil direction as a clue for liveness detection. In [1], user is requested to open and close his/her mouth, or blink his eyes to demonstrate liveness. All these active-based approaches are faced with either some or most of the challenges of active approaches which include use of extra hardware, high intrusiveness, expensive, computational complexity, less ability to fit into existing systems, etc.

An approach that is close to the one proposed in this study is in [20], but in their approach a single sample requires eight (8) face frames at different distance from the camera , and a the feature is represented by 7 by 2147 matrix. In contrast, the method proposed in this study, requires only 2 frames which is represented in a vector space. Also the approach does not require extra hardware which makes it possible to be used with existing systems.

The major contributions of this work is summarized as follows

- A robust active face spoof detection technique, yet less intrusive, less expensive, and a more generic applicability than other active-based approaches, with not extra hardware requirement.
- A dataset of face video frames captured at different distance from the Camera are collected from eighty (80) volunteers.
- A classical literature review based on the taxonomy of existing spoof detection approaches

2. FACE SPOOF ATTACKS

A spoofing attack is a situation whereby a fake evidence is presented to a biometric system for authentication thereby gaining false acceptance [21–24]. Face spoofing is the activity of disguising an individual's face in order to bypass a face biometric system thereby gaining illegal access or privileges. It has become very easy as face photos or videos of a person can easily be gotten from social media or captured from afar without the person's consent [25]. This is done by simply displaying photos or replaying videos, obtained illegally obtained. As efforts are put forward against the spoofing attacks, spoofing artefacts has also evolved into diverse forms of attacks. Figure 2.1 summarizes the various forms of face spoofing attacks.

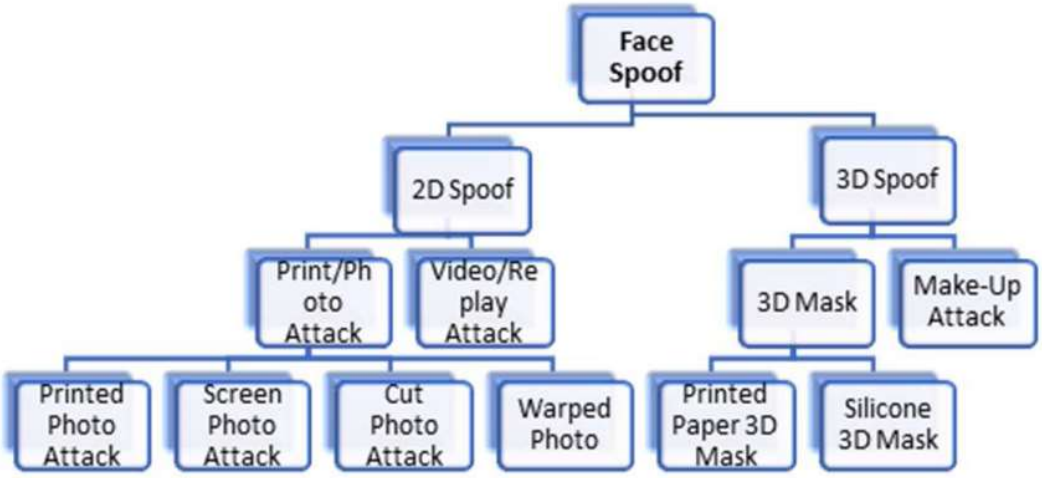

Figure 2.1. Graphical Summary of types spoof attacks

2.1. 2D Face Spoof Attacks

The 2-dimensional (2D) attacks are face presentation attacks which are based on flat surfaces. They can be subdivided into two categories which includes print attacks and replay attacks.

2.1.1. Print or Photo Attacks

Print attacks are carried out by showing a real user's photo which might be obtained from social media or captured from a distance [26]. The availability of people's photographs on the internet which are accessible without any permission, makes print attacks the most common and easiest form of face spoofing. They are also presented in diverse forms, which include:

- **Printed-photo Attack:** Here, a genuine user's printed photo is utilized in order to subvert a face biometric system.
- **Screen-photo Attack:** This is done by harnessing digital-screen of electronic/digital devices (e.g. tablet, mobile phones etc.) to display a photo containing the face image of a genuine user. Today, it has become so easy access people's photos on internet, especially through the social media such as Facebook, twitter, Instagram, etc. As such, a person's image can easily be downloaded on a mobile phone and presented to the face recognition system.
- **Cut-photo Attack:** This is achieved by using the printed copy of genuine user's photo, where some regions of the face such as mouth, nose, mouth, or eye has been cutout, while the impostor places his/her face behind the cut photo to forge the blinking of the eye and mouth movement behavior of a live face. Figure 2.2 is a typical example of a cut-photo attack with the eye and mouth region cut out.

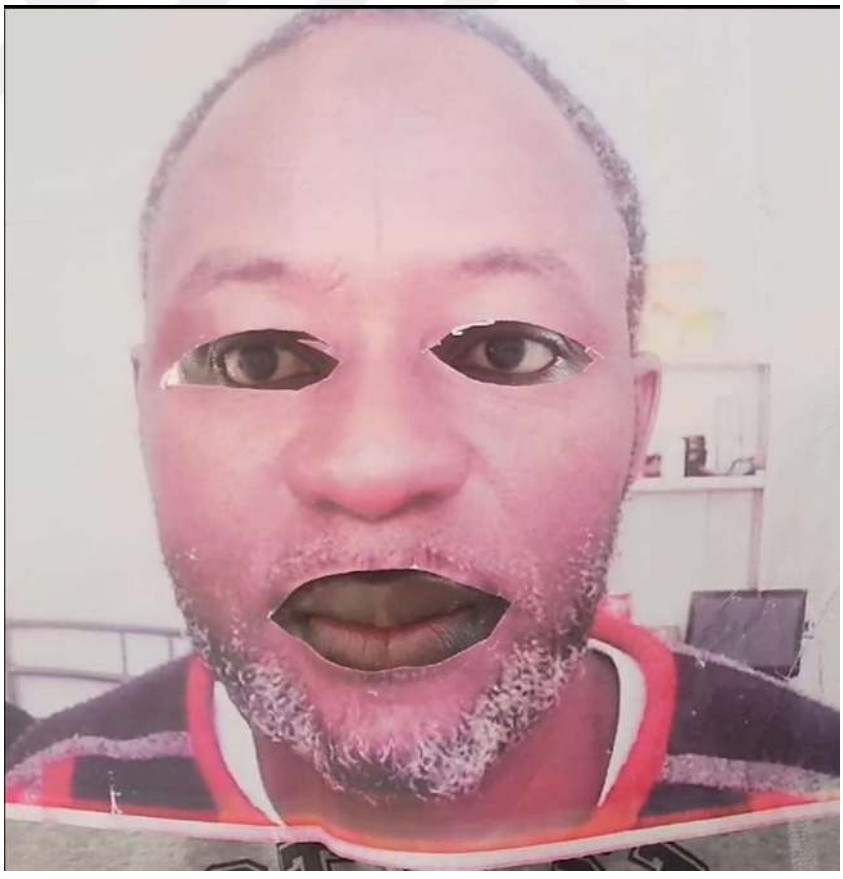

Figure 2.2. Cut Photo Attack with the eyes and mouth region cut out

- **Warped-photo Attack:** This is simply trying to bend a printed photo in some directions to mimic facial motions (See Figure 2.3). The warped photo-attacked came up to beat anti-spoofing systems where the liveness cue is based on some sort of facial motions.

Figure 2.3. Typical example of a warped Photo attack

The print-photo and the screen-photo attacks are the earliest forms of attacks. Other forms of attacks emerged to thwart a specific anti-spoofing algorithm. For instance, the cut-photo attack emerged in order to thwart the anti-spoofing algorithms based on motion cues such as eye-blinking and mouth movement.

2.1.2. Replay Attack

This is a more sophisticated form of attack than the earlier mentioned forms of attacks. Replay attack is achieved by displaying a looped video of a genuine user's face through the screen of a digital device (See Figure 2.4). This present a more natural form of physiological behaviors of the human face. Physiological signs such as facial expressions, eye-blinking, mouth or head movement are demonstrated in the video displayed as opposed to print attacks which are still images. In order to ease the efforts researchers in building a strong spoof detection mechanism, authors in [27] came up with a publicly available dataset known as the REPLAY-ATTACK DB. The Replay Attack Dataset consist of a thousand and three hundred (1300) video clips of print and replay attack attempts to fifty (50) clients, captured under various lighting conditions. The dataset contains both legitimate samples by having a live user trying to access the system through a built-

in webcam and those of the spoof cases by displaying a photos and video recordings to the system for at least 9 seconds.

Figure 2.4. Replay attacked performed using a mobile phone

2.2. 3D Face Spoof Attacks

This is the most sophisticated form of attack. Here, the spoofing artefact is the 3D mask of the face of a genuine user. The entire 3D structure of the human face is emulated (See Figure 2.5). This makes it more tough to overcome. The use of depth information which is often an answer for print attacks becomes frustrated when presented with 3D attacks [22]. Other forms of 3D attacks include cases of sophisticated makeup and plastic surgery.

Although this is the most sophisticated form of attack, it is also the most difficult and expensive to acquire, which makes it less common as compared to other forms of attacks. Also due to the cost of acquiring 3D masks, there are less available public dataset of 3D attacks as compared to others. Some of the publicly available datasets with 3D mask include 3DMAD [28], BUA Lock3DFace [29], and SMAD [30].

Figure 2.5. A typical example of a 3D mask attack

Figure 2.6. A typical example of a Sophisticated makeup

Figure 2.6 above is a typical example of a sophisticated makeup. The image on the left is the real image of Lucia Pittalis an Italian makeup artist as well as portrait painter and drawer. While the image on the right is an image of her transformed into an entirely different person using makeup.

2.3. Face Spoof Databases

There exist several face anti-spoofing databases made publicly available that has help progress the efforts on developing a robust face spoof detection mechanism. Some of the popular databases in literature will be discussed here.

The Print-Attack database [31] consisting of 200 videos of genuine faces and 200 videos of spoof cases collected from 50 subjects.

The Replay Attack database [27] consists of 1200 videos that include 200 real access videos, 200 print attack videos, 400 phone attack videos, and 400 tablet attack videos of 50 different subjects. Extraordinary, each video was captured in two different environments: fixed-support and hand-held.

The NUAA database [32]**,** consisting of 12,614 face images which is extracted from 143 videos of genuine and attack attempts of captured from 15 subjects .

CASIA-FAS DB [33] consist of 150 videos of live faces and 450 videos of spoof cases collected from 50 subjects. It contains warped, cut and replay attack types.

Michigan State University Mobile Face Spoofing Database (MSU MFSD) [2] consisting of 110 videos of legitimate samples and 330 videos of spoof cases collected from 35 subjects. The spoof types captured in this database is the print and replay attack types.

CASIA-SURF [34]**:** Consisting of 3000 videos of live faces and 18000 videos of spoof cases collected from 1000 subjects. The attack types captured in this database include print and cut photo.

CASIA-SURF CeFA [35]**:** This is an extension of CASIA-SURF consisting of 23538 videos collected from 1,607 subjects. This database contains almost all forms of attack types.

3DMAD [28]**:** This is a very large database consisting of 30600 videos of live faces and 45900 of spoof cases captured from 17 subjects. This consist mainly of replay and 3D mask attack types.

CelebA-Spoof [36]**:** This is a recently collected database with diverse forms of attacks. It is one of the largest face spoof database consisting of 625,537 videos captured from 10,177 subjects.

3. FACE SPOOF DETECTION

Face spoof detection also known as face anti-spoofing is simply the task of preventing false face verification. In spite of the fact that the task of face spoof detection dates back to over a decade [37], until around the year 2010 that it experienced real revolution under the European project known as TABULA RASA which was focused on biometric spoof attacks [38]. Another factor that has brought about extensive progress in developing techniques for detecting face spoofing attacks, is the collection and distribution of face spoof databases which has been made publicly available. This has relieved researchers the stress of collecting data so as to focus on building robust techniques against the diverse forms of attacks [27,28,32,33].

These key factors have brought about the publication of several approaches in combating both 2D face spoof and 3d attacks. These approaches can be classified based on mode of enrollment and based on descriptors/liveness cues. Figure 3.1 is a diagrammatical summary of the classifications of face anti-spoofing methods, with examples.

3.1. Classification Based on Mode of Enrollment

Firstly, face spoof detection can be broadly classified into two groups: Active methods and passive methods. The dichotomy between active and passive anti-spoof methods is based on their mode of enrollment/activation or the mode of capturing user data.

3.1.1. Active Spoof Detection Approaches

The active face detection methods encompass approaches that involves the user's active participation in the process of capturing data or enrollment process for anti-spoofing. Approaches considered to be active-based are interactive in nature. Users are requested to perform a few actions before the camera. For example: user may be requested to smile, demonstrate mouth movement by opening and closing his mouth, blinking his/her eye, and or nod his/her head. All requested actions must be completed successfully. Though active-based methods are robust in spoof detection, it active participation of the user which makes it intrusive and might not be appropriate for some use cases. Also most active approaches require extra hardware, which makes it more expensive and reduce the use cases and cannot be easily built into an existing system.

The approach proposed in this study is an active, yet less intrusive, less expensive, does not require extra hardware and so can easily be built into existing systems. Another disadvantage of most active-based approaches is that it reveals to users the liveness cue employed. The approach proposed in this study also holds the advantage of not suggesting to the user the liveness cue employed.

3.1.2. Passive Spoof Detection Approaches

Passive approaches are non-intrusive in nature. They do not require user interaction. Usually, a user does not need to know he/she is being tested. The system performs everything on its own – without requesting the user to carry out any extraction action or pay attention.

3.2. Classification Based on Liveness Cues or Descriptor

Secondly, based on liveness cues or descriptors, face spoof detection methods can be classified as shown below:

- Approaches based on motion cues
- Texture based approach
- Image Quality based approach
- Depth information
- Multi-modal
- Ensemble of descriptors
- Approach based on Deep learning
- Sensor level approach

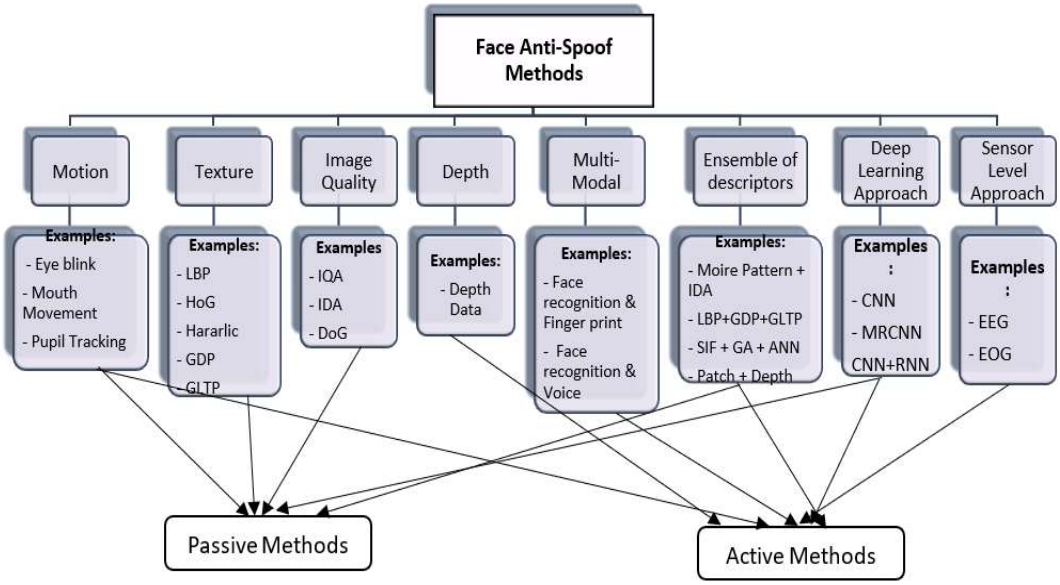

Figure 3.1. Summary of face spoof detection techniques with examples

4. LITERATURE REVIEW

Authors have revealed several approaches for spoof detection in literature. As seen in the previous chapters, both spoof attacks and spoof detection techniques has experienced tremendous evolution over the years. This chapter reviews different approaches proposed by various authors in combatting the evolving face spoofing attacks. The different approaches are discussed based on the classification of the approaches as shown in Figure 3.1.

4.1. Spoof Detection Based on Motion Cues

In combatting the activities of attackers, early approaches for face spoof detection were on the basis of motion cues (the blinking of the eye, rotation of the head, mouth movement, etc.). These approaches were primarily designed for print attacks [2], printed-photo and screen-photo attacks to be specific. Authors in [3] proposed an approach for spoofing detection based on the blinking of the eye using Conditional Random Fields (CRFs). The inspiration is drawn from the fact that Eye blinking could happen spontaneously. The traditional blinking rate of an individual at the state of rest is often twenty (20) times each minute, with a blink lasing typically for one quarter of a second [4]. Hence, the blinking of the eyes can be a very effective indication of liveness. The result of the experiment on samples on the CRF, Cascaded AdaBoost, and HMM (Hidden Markov Model) shows that CRF did obvious better as compared to others. Although it achieved an accuracy as high as 98.3%, it is computationally complex, the response time is also high, and has no potential of overcoming more complex spoof cases like fake motion [2].In a study by Singh, Joshi and Nandi in [1] authors proposed an anti-spoofing measure based on eye blinking and the movement of the mouth in which they incorporated to their face recognition system. Their approach divides the process into two phases: the first phase handles the detection and the second is the face recognition phase. The system will first check if the face presented is a live face or not, and if the liveness of the user is ascertained, then the face recognition phase is activated. The system first checks for the presence of the within the eye region which indicates whether the eye is opened or not. The blinking of the eyes (opening and closing) was done by checking for the existence of the eye in the eye region. A case where the eye is found, it is assumed that the eye is opened otherwise closed. Likewise, mouth movement is determined by the presence of the teeth. In the mouth area. This is done using the teeth's HSV (Hue, Saturation, Value) values. When teeth are detected in the mouth area, that connotes it is open, else it is closed. Using the few samples of attacks that the authors generated, the performance of the system gave good accuracy score, but failed in scenarios of cut-photo attacks.

Another effort based on motion cues was proposed in [5] in which the tracking of the pupil was the basis for their approach. In their study, they built a system which checks for the direction the pupil by utilizing a hardware equipment (a square frame having eight Light-Emitting-Diodes(LED), each LED for a direction). Firstly, the HaarCascade-Classifier of the OpenCV library is utilized for extracting eye area for extracting and tracing of the feature points in order to obtain a stable eye region. The eyes region is cropped-out and rotated to obtain a stable eye area. The pupil is extracted from the eyes-area. Obtaining a desirable number of frames that are stable, the Arduino will be signaled to activate a LED (Light Emitting Diode) from any direction selected randomly. After which observations are made in order to find out whether the Led position matches with the direction of the pupil. Satisfactory compliance returns data containing liveness information. Although testing of the algorithm on volunteers shows a high success ratio was achieved, the approach has low robustness and generic ability. It can with ease be bypassed by a 3D mask or cut-photo attacks. Also, being an active approach with extra hardware requirements, it cannot be easily built into existing face recognition systems. Data that contain information on liveness are return in a case where there is conformity.

4.2. Approaches Based on Image Texture Analysis

The extraction of distinct features has made texture analysis to undergo broad application n in anti-spoofing systems[39]. Several authors have utilized diverse texture-based feature descriptors as a spoof detection mechanism. These include the popular Local Binary Pattern(LBP) [27,40–43], Local Ternary Pattern (LTP) [39,40], Hararlic feature descriptors [44], Histogram of Oriented Gradient(HoG) [45–47], Gradient Directional Pattern(GDP) [46], SURF (Speeded Up Robust Features) [48], Local Phase Quantization (LPQ) [49–51], Fourier Spectra [52] and so many other forms of texture descriptors.

In [11] Lambertian reflectance was utilized in order to discriminate between 2D print attacks and live faces. The extraction of latent reflectance features was done using both variation retinex-based method and Difference-of-Gaussian (DoG) metho, which is then used for classification. Experimental result on the NUAA (Nanjing University of Aeronautics and Astronautics) dataset [33] gave an impressive result with 0.94 as the Area Under Curve.

Jukka et al. [41] proposed technique of face spoof detection using micro-texture analysis from a single image. The main idea was to emphasize the distinction of micro texture in the feature space. Authors adopted the Local Binary Patterns (LBP) which is an effective texture operator, for describing the micro-textures alongside their spatial information. The feature space vectors are used as input to for SVM classifier which classifies the micro-texture patterns as either fake or real person's image. Firstly, detected is cropped, normalized and converted into an image of 64 × 64 pixel. LBP operator is then applied on the normalized face image and then divided into 3×3

overlapping regions. Computations are then performed on the local 59-bin histograms gotten from each region, and collected into a single 531-bin histogram. Then, LBP operators are used to compute two other histograms obtained from the whole face image. Finally, a nonlinear Support Vector Machine (SVM) classifier is used to determine whether the input image is a fake or live person image. The experimental results reveals better performance with equal error rate (EER) of 2.9% than other texture operators like Gabor Wavelets and Local Phase Quantization with EER of 9.5% and 4.6% respectively as shown in [32].

In [53], RGB, Hue Saturation Values and the YCbCr color spaces were examined by authors in a bit to comprehend the which is more informative in discriminating between real faces and attack cases. Color local binary pattern descriptor was utilized in order to analyze the joint color texture from the gray scale (luminance and chrominance). They experimented their study on two public database (CASIA FASDB [33] and Replay-Attack DB [27]). Result reveals that the color texture is more discriminative than the gray scale in the various spoof cases. Also the result shows a better performance from the YCbCr channel as compared to other color spaces. Also HSV color space features was more discriminative against replay attacks than other color spaces. A combination of both HSV and YCbCr color texture representations offered the most discriminative information.

Authors in [42], further improved on their effort on color texture-based method in [53] and proposed an approach which is more robust and with more generic applicability. In their method, LPQ (Local Phase Quantization) features in YCbCr and multiscale local binary pattern features of the human face were utilized. Result of their experiment on 3 challenging publicly available datasets: CASIA-FAS DB, Replay-Attack-DB, and MSU-MFSD[2], reveals a very competitive performance. For the CASIA-FAS DB and MSU-MFSD, their color-texture representations supported the mixture of Color-local binary pattern features in HSV and LPQ features in YCbCr color spaces computed performed better than state-of-the-art, giving an impressive performance on Replay-Attack-DB. A stable performance was achieved across all the public databases used contrary to most of the other methods reviewed in their literature. The facial color-texture representation also shows a promising generalization ability in the inter-database evaluation. It therefore suggests that color-texture is more stable in unknown-conditions than that of the grayscale. It is also important to note because optimizing face normalization or the face bounding box limits shown to be essential for intra-database testing [2,54,55] were not considered in their experiments, it significantly affect their performance in the cross-database scenario.

Another approach based on texture analysis is shown in [48] authors proposed a method which describes the face's appearance using Fisher Vector encoding on SURF(Speeded Up Robust Features) features from various color spaces. In their method, SURF features were gotten from color images instead grayscale. This was achieved by application of the SURF descriptor on each

of the color band, and there after concatenate them to form a single feature vector known as the CSURF. Principal component analysis is then used to reduce the dimension of the face.

Recently, a novel approach was proposed in [45] for spoof detection based on the use of multivariate histogram of oriented gradients (MHoG) descriptor for detection of micro expression (μE) regions of face videos. Facial micro expressions (μE) are very brief facial expressions occurring spontaneously that highlights the face of humans while concealing an emotion either consciously or unconsciously. Their approach accentuates the variations in micro-expressions in spoof and original video representation by a significant amount and assert that such a variance is sufficient to discriminate against spoof attacks. The approach does automatic extraction of the region of interest (ROI) of crucial changes in μE using the MHoG parameters and thereby proposes the descriptor as more fitting for characterizing liveness. A satisfactory and statistically significant result was obtained when experimented on their custom dataset of replay attacks.

4.3. Approaches Based on Image Quality Analysis (IQA)

Another cue explored by many researchers is the use of Image Quality Analysis. Approaches based on Image Quality analysis has also competitive results.

In a study by Galbally and Marcel [56] a new method based on general image quality assessment (general IQA) was proposed for spoof detection. Authors derived a distortion version of each of the images. This is achieved by the application of a low pass Gaussian-kernel-filter where variance(σ) and seize are set as 0.5 and 3x3 respectively. A total of fourteen (14) image quality features were extracted from each of the images. The method was crafted such that prior to all other processing activities, the image quality features are first computed. This enhances the reduction of complexity in computation. To classify the images as either live or spoof face, a simple LDA is utilized. Evaluations performed using Replay-Attack DB and CASIA FAS-DB, shows a very competitive performance in comparison with other state-of-the-art methods evaluated on same benchmarks. The result also shows that general IQA contains high and valuable information capable of distinguishing between genuine face and spoof faces. In their experiment a decrease in the performance of spoof attacks with high image quality was observed in instances where the devices are in a fixed position. Results of evaluation on CASIA-FAS-Database reveals decreased error rate (indicating better performances) where the resolution of acquisition device is higher. Their experiments also confirmed the assumption that Gaussian filter's impact on the genuine and spoof face images samples are different.

In the study carried out in [32] a lambertian-reflectance based approach for distinguishing between 2D print-attacks from 3-dimentional live faces. The extraction of latent reflectance features was done using variational retinex, and DoG methods to obtained features that will serve

as input for classification. NUAA Database was use for the experiment and the performance indicates it is promising approach with Area AUC of 94%.

The method proposed by authors in [33] utilizes several DoG filters. This was due to the fact that there was no prior knowledge of the most discriminative frequency. They used multiple DoG filters in extracting high-frequency features from the facial images, to form redundant feature set. Then the removal of noises and low frequency details were carried out the Gaussian-variance (σ) set to specific values. Four DoG filters (σ1=0.5, σ2=1; σ1=1, σ2=1.5; σ1=1.5, σ2=2; and σ1=1, σ2=2) were considered, where σ1 and σ2 represents the inner and outer Gaussian variance respectively. The filtered images are then fused together and use as input for Support Vector Machine for the training of a final classifier. Results of their experiment using CASIA-FAS-DB shows Equal Error Rate of 17%. The accuracy was insufficient for application in a real-world system. Their result also shows that the accuracy of the approach can be affected by the quality of images and the type of attack.

Four IDA features (specular reflection, color diversity, blurriness and chromatic moment) were extracted to form a feature vector in [2]. The authors extended the work in [57], where a more generalize spoof detection method was their goal, but in their experiment the training and testing was intra-database (even though they employed cross-validation technique). The work in [2] achieved an improvement in terms of generic-ability by performing a cross database training and testing on Replay-Attack Database and CASIA FAS Database. They also collected a new spoofing dataset. The ensemble of classifiers was employed as two Support Vector Machines classifiers were used for each spoof case include the print-photo attacks and replay attacks. Experimental result of the intra-DB testing shows that their method did better than state-of-the-art methods. The inter-database test did obviously better compare to baseline approaches. There is also a reduction in the complexity of computation since only one image frame is needed.

Owing to the fact that previous efforts based on image distortion analysis (IDA), only considers the facial details without giving attention to the images patterns, Anand and Gupta [58] came up with a hybrid anti-spoofing technique that combining SIFT (Scale-Invariant Feature Transform) feature descriptor, genetic algorithm and artificial neural network to form distinct feature sets according to the categorization of images in the database. The authors used SIFT descriptor for extracting key-points and also to identify the patterns of the face region. Feature optimization is then carried out using genetic algorithm in order to separate distinct features and to obtain a normalized feature set for each of the categories of the image data. Finally, an Artificial Neural Network is designed and used to train the feature set. Authors constructed a face spoof dataset captured using camera without a form of compression applied. Although experimental result on the custom dataset shows an impressive performance with an average accuracy of 97.86% and

an error rate of 2.13%, the system was deficient in rightly identifying real faces as the result shows a False Rejection Rate (FRR) of 53.29%.

4.4. Approaches Based on Depth Cues

The depth information has more discriminating capacity against 2D spoof cases since the depth of genuine faces are not even, whereas spoof images have even depth. Atoum et al. [59] exploited the depth supervised procedure. The use of depth information definably an answer to print attacks (with plane surfaces) turns out to be frustrated in 3D attacks scenarios [22], since it emulates the entire 3D structure of human face. Approaches based on depth information can be intrusive and often require extra or special hardware such as Near Infra-Red (NIR) camera which makes it difficult to deploy in an existing system.

In order to achieve the goal of confirming the capacity of depth cue in detection of spoof attacks, authors in [7] built a system which made use of the 3D-sensor of Microsoft Kinect device to acquire depth information. Users are first enrolled by capturing the face and information into the system at first time. The system does the scanning of the image and measuring of the depth simultaneously by placing 5 points on the image that is scanned. A dot is placed at the image's centre, 30pixels on the left, upper, lower and right sides of the centre in order to mark the five points. The depth values are then compared with the expectation the will vary. Getting at least two equivalent depth values, indicates a face is planar and thus spoof. To detect and recognize face, Haar-Cascade of OpenCV library which has compatibility with Kinect Sensor is utilized. Testing their approach against five different attack types, an accuracy of 100% was attained in detecting attacks. The system was not affected by occlusions such as the use of glasses, scarfs, and makeups, likewise the size of images does not affect the performance. Irrespective of the fact that the Kinect sensor has a depth range as high as 0.7 - 6 m, it was also confirmed that, the distance from the camera particularly more than 1meter results in lower accuracy of detecting attacks. It is worth of note that basing liveness cue only on depth cue is inefficient where 3D mask attack is involved [42]

An approach that is based on depth supervised learning was proposed in [59]. Depth information were estimated from multiple RGB frames. Experimental result on CASIA-MFSD, OULUNPU [60], spoof in the wild (SiW) [61], and Replay-Attack DB reveals that their approach achieved state-of-the-art results.

4.5. Multi-Modal Approaches

Multimodal systems are said to be of more robustness in spoof detection than unimodal systems [62]. Multimodal approaches involves the fusion of two or more biometric systems, such as face and voice [19], face and finger print [10,62,63].

Authors in [19], approached spoof detection by fusion of two human traits: face and voice signal. Their report shows that they biometric system with two traits offers more security and is more reliable as compared to the unimodal system using face alone. Features from the two traits were extracted individually and classified using Gaussian mixture model. After which the traits are fused to obtained the training dataset. The model evaluated using metrics such as False Rejection Rate (FRR), False Acceptance Rate (FAR), Failure to Capture (FTC) and Equal Error Rate (EER), proves that the propose multimodal system performs better than the unimodal system and also shows good response time and accuracy requirements.

In [63] authors proposed a spoof detection approach based on fusion of biometrics. The fusion schemes which harnessed two biometric systems: Face and fingerprint recognition systems. The face recognition system was implemented by utilizing eigenfaces [64], whereas the fingerprint was implemented using the fingerprint system by *National Institute* of *Standard* and *Technology (*NIST) [65] which is available publicly.

Results of authors in [66] also shows that the multimodal approach performs better than unimodal approaches. In their approach authors fuses invariant iris features and finger vein shape features. The algorithm was based on prioritizing high dimension features reduction. They considered iris Hu moments and vein shape features of the finger in order to accomplish a robust authentication system. Features trained on SVM shows that multimodal approach outperforms unimodal system.

Although multimodal approaches have proven to be very robust and have the capacity to handle all forms of attacks, they are highly intrusive, more expensive as more hardware is required, and also has less applicability range as it cannot easily fit into an existing system.

4.6. Approaches Based on Ensemble of Descriptors

Recently, developments in the direction of material-independent static spoof detection suggests the combination of multiple features descriptors[62]. Several authors have harnessed the ensemble of descriptors to achieve promising results.

Recently, Thippeswammy, Vinutha, and Dhanapal [51]proposed an approach based on ensemble of six local appearance-based methods texture descriptors. In their approach, they combined Gradient Directional Pattern (GDP), Local Binary Pattern (LBP), Gradient Local Ternary Pattern (GLTP), Local Directional Pattern (LDiP), Local Gabor Binary Pattern Histogram Sequence (LGBPHS), and Local Phase Quantization (LPQ) feature descriptors to form an ensemble feature vector representation of an images. K-nearest neighbors (KNN) classifier is then used to classified as real or spoof images. An implementation of their algorithm on the NUAA Photo Imposter database, shows a competitive performance in identifying spoof faces. Result shows that

98.39% of spoofed faces correctly identified. Whereas, the system's performed poor in identifying genuine faces as only 51% of genuine faces were rightly identified.

Also, Dhawanpatil and Joglekar [6] in their approach combined Moiré patterns and Image Distortion Analysis (IDA) to detect spoofing cases. They divided their approach into 3 modules as shown in Figure 4.1. The input module used in capturing an individual's face which can be a real or fake face; The spoof detection module, which is achieved by checking for Moiré patterns and IDA (Image Distortion Analysis) in the face image. Overlaying two patterns on each other will produce Moiré patterns. In the case of a replay attack, light ray from the device used to display the video of the user will be casted on the capturing device's screen, hence producing Moiré patterns which will be used in detecting the replay attack. LBP (Local Binary Patterns) descriptor is used to detect Moiré patterns. This is achieved by comparing the LBP histograms of captured face and that of the database images. The result determines whether the face is spoofed or genuine [2]. While in the printed paper attacks scenario, the surface roughness of genuine and spoof faces is observed. Moiré patterns are then identified in spoof cases and the image hue and saturation parameters used to compare the color histograms of real and fake faces. Image distortion analysis is then performed and finally, the third module which is the face matcher module, does the classification of the image as either genuine or spoof using an SVM (Support Vector Machine) classifier.

Pan et al. [67] proposed a method that combines the use of eye-blinking and a texture-based scene-context matching for face spoof detection. They utilized the eye-blinking model based on CRF (Conditional Random Field) which they earlier proposed [68] to detect the blinking of the eyes. Then, the coherence of the background region and the actual background (the reference image) are checked using a texture-based approach. The reference image obtained by capturing the background, without user in the image. For genuine face image, the background region of the face for the input image and that of the reference image are theoretically identical. On the centrally, presenting replay attack or digital photo will give different background region around the face for the two images. In comparing the background of the input image and that of the reference image, Local Binary Pattern (LBP) features are first extracted from fiducial points selected by the use of DoG function, and further obtain scene matching score by calculating the χ 2 distance. Detecting spoof by either the motion cue or the texture cue, will imply denial of access to the system. This approach is limited in real life application since it requires a fixed camera and a none-monochrome background.

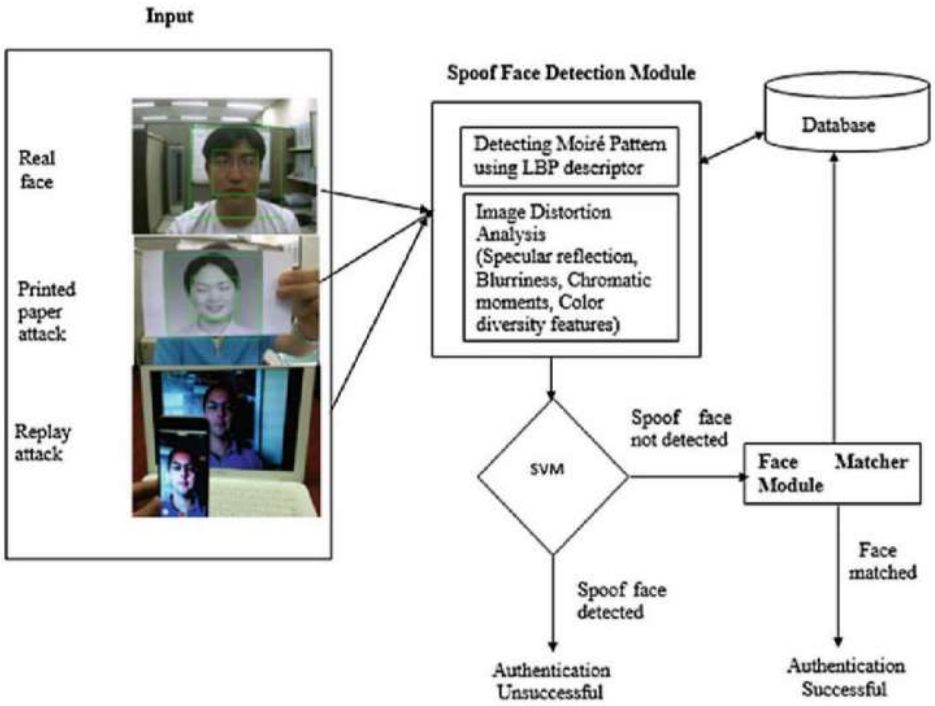

Figure 4.1. Model design of spoof detection system in [6]

Another approach based on the utilization of ensemble of descriptors was proposed in [19]. Authors combined features extracted from three (3) different feature descriptors, namely: Haralick texture, Color moment, and CLBP (Color Local Binary Pattern) feature descriptors which forms the feature space. Logistic Regression is then used to classify features obtained. An evaluation of the approach on MSU-MFSD dataset achieved competitive results as compared to state-of-the-art methods, with an F1-Score of 99.45%.

4.7. Deep Learning Approaches

In recent times, deep-learning approaches are widely employed to extract deep features [7,9,69–71]. This possesses more semantical information compared to features from traditional handcrafted techniques[69]. Hence, harnessing the deep learning for face spoof detection has experience wide usage recently. Deep learning has achieved promising results in computer vision as a field, and has shown effectiveness in tackling face spoof detection task [72]. In [73] Convolution Neural Networks(CNN) is utilized for the extraction of deep features, while Support Vector Machine is employed instead of fully connected layers for classification.

A Convolution Neural Networks - based method with two stream was proposed in [59], the utilization of full-face image and patches of the same face is employed for distinguishing genuine face from fake faces. CNN made use of the local features extracted from the image for discriminating fake patches while the holistic depth map extracted from the same image are used

to examine whether the image has a face-like depth. An experiment on CASIA-FASD, MSU-USSA, and Replay Attack datasets shows an impressive performance when compared to the performance previous state-of-the-art approaches. Although similar EER to previous approach was obtained, they got much smaller HTER (Half Total Error Rate). The fusion brought also resulted in an increase in AUC for Replay-Attack DB from 0.989 to 0.997.

A method proposed by Liu, Jourabloo, and Liu [61] is based on coherent combination of CNN and RNN (Recurrent Neural Network) architecture. Unlike other deep learning methods that also approached face spoof detection similar to most traditional method as a binary classification task, in their approach, rather than extracting any cues to separate two classes which cannot be generalized, the known-spoof patterns across spatial and temporal domains was the focus of their approach. Leveraging two auxiliary information: depth-map and rPPG (Remote Photoplethysmography) signals as supervision as shown in Figure 4.2. Depth map supervision is utilized by CNN part for finding hidden texture features for depths distinction between genuine and spoof faces. Contrary to the traditional computation of rPPG signal [48], the RNN part is allowed learn estimating the rPPG signal itself, this checks the video frames temporal variability. Experimental result of their model on OULU-NPU, CASIA-MFSD, SiW, and Replay-Attack DB attained state-of-the-art results both in the intra-database and cross-database testing scenarios. their method reduces the cross-database testing errors (HTER) by 24.6% and 8.9% on the CASIA-MFSD and Replay-Attack datasets respectively as compared to the previous leading state-of-the-art results [53].

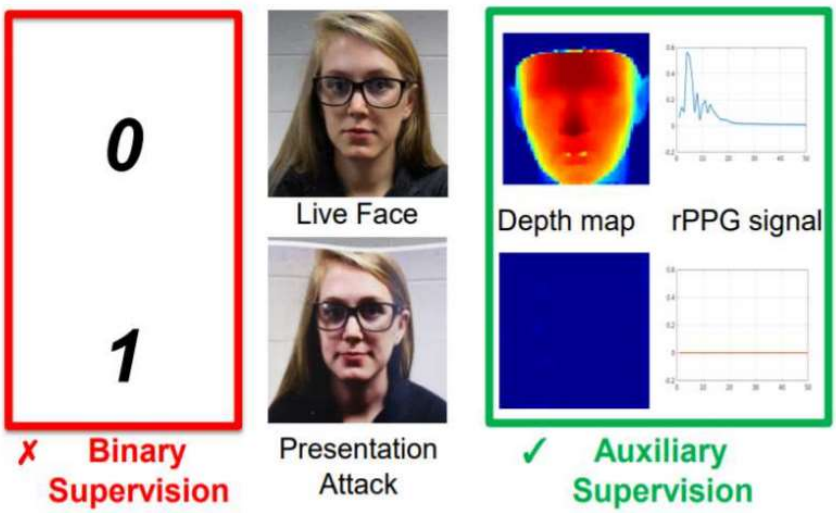

Figure 4.2. Binary supervision vs auxiliary supervision [61]

Based on earlier attempts using depth supervised learning, considering depth for auxiliary supervision was only in a single frame. In [74], a method was put forward based on the estimation of depth information in multiple RGB-frames. Authors divided their approach into two sections:

The single-frame part, mainly for finding cue for detecting attacks with static depth supervision; The other section which uses multi-frames, extracts the long-term and the short-term motions with the capacity of distinguishing genuine faces from attacks: The OFFB serves as the short-term motion module, while convolutional gated recurrent units (ConvGRU) serve to model long-term motion patterns. A state-of-the-art result was attained from experimental on OULU-NPU, SiW, Replay-Attack and CASIA-MFSD databases.

Authors in [8] designed a Multi-Regional Convolutional Neural Networks (MRCNN) for spoof detection in their approach. In their approach, information from all patches are utilized by scattering the gradient distribution, so as to improve the robustness of the system for spoof detection. It utilizes information from the whole face region, thereby avoiding excessive emphasis of certain local areas. This through the introduction of local patches, which is the concept of local classification loss. Experiment result on Print Attack, SiW, OULU-NPU and Replay Attack datasets gave better performances in terms of robustness against adversarial attacks. Also, performance against traditional attacks was improved as compared with previous approaches.

Another recent and interesting approach is proposed in [69]. In attempt to tackle the problems of large parameters with weak generalization ability of networks, they designed a FeatherNets architecture to be a network as lite as feather. Therefore, an extremely lite network architecture they called FeatherNet A/B was designed with a streaming module that fixes the weakness of Global Average Pooling unit and uses less parameters. The training of their single FeatherNet by depth image only, achieved a higher baseline with 0.00168 ACER.

5. MATERIAL AND METHOD

Here an overview of the materials and methods employed in this work discussed including the general model design and procedures, data collection and dataset generation, models used for classification. Finally, the evaluation metrics employed for performance measurement of the proposed approach will be discussed.

5.1. Materials

For the purpose of this study, Table 5.1 provide details of materials used for the experiment of the proposed approach:

Table 5.1. List of materials used and their features

S/N	Item	Description/Features
1.	Laptop	- Brand: HP - Model: Elite book - Operating System: Windows (Windows 10 Pro) - Processor: Intel(R) Core™ i7-3667U - Processing speed: 2.00GHz and 2.50GHz - RAM: 8.00GB - Memory: 166GB SSD
2.	Mobile phone for capturing	- Brand: Huawei Mate 20 lite - Android Version: 10 - Camera: 20+2 mega-pixel rear and 24+2 mega-pixel front camera
3.	Samsung Galaxy Tab E SM-T560 Tablet used for screen photo and replay attack	display dimension of 241.9mm by 149.5mm
4.	An external hard disk for backup of captured video frames	500GB
5.	Programming Language:	Python
6.	IDE	Jupyter Notebook
7.	Iriun webcam application	For connecting phone camera with laptop wirelessly
8.	Major libraries used include:	- SciKit-Learn for Machine Learning Models Development and Evaluation - Matplotlib and Seabon for Data Visualization - Pandas and Numpy for Data Analysis - TensorFlow for Deep Learning - OpenCV for Computer Vision

5.2. The Model Design

As seen in Figure 5.1 below, the model design is divided into three (3) basic modules which include: The Frames Selection Module, Distortion Feature Extraction Module and finally the Classification Module. The frame selection module takes facial videos as input and two frames are selected each from two facial videos captured at different distance from the camera. The Feature extraction module then detects a number of facial landmarks from each frame and compute the features about the distortion changes of the faces among different frames. Finally, the classification module performs binary classification using the extracted features to distinguish fake face from real ones.

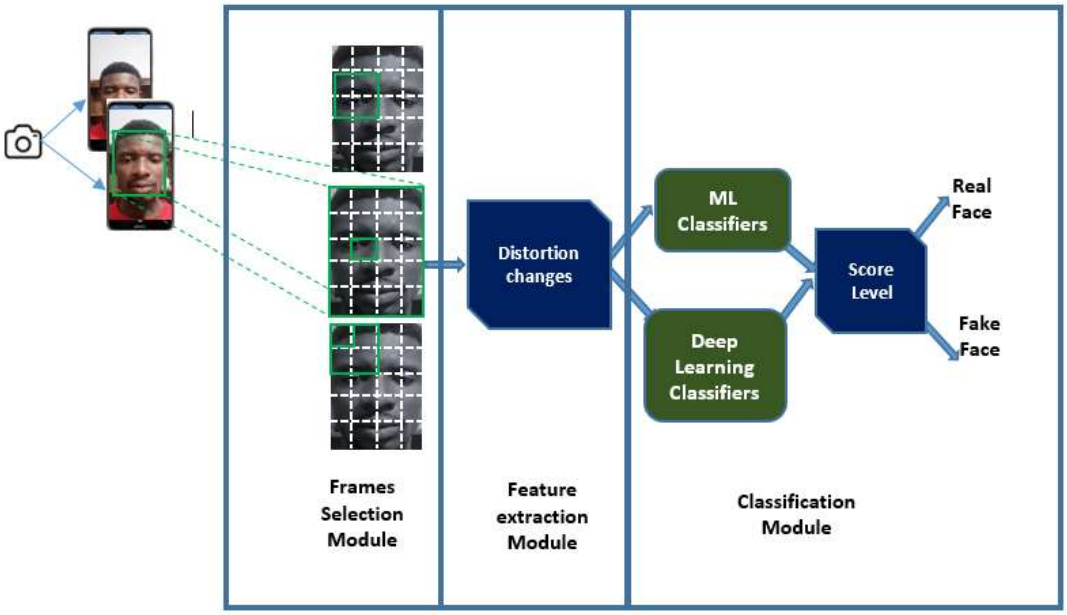

Figure 5.1. Basic Design of workflow for the proposed approach

5.2.1. Frames Selection Module

Firstly, as the user's face moves closer or farther from the mobile device, the devices camera captures two video frames taken at two different predefined distances say D_1 and D_2 between the camera and the user's face. Each video contains a number of frames of the user's face. The Frame selector module then utilizes the Dlib face detector algorithm [75] to extracts a K sequence of frames ($D_1f_1, D_1f_2, \ldots, D_1f_k$) and ($D_2f_1, D_2f_2, \ldots, D_2f_k$) from video frames captured at distances D_1 and D_2 respectively. The frames are made to form pairwise frames D_1f_i *and* D_2f_i, for $i=(1,2,\ldots, k.)$.

5.2.2. Feature Extraction Module

With the extracted frames D_1f_i *and* D_2f_i as input, the feature extraction module detects a number of facial landmarks and calculates the geometric distances (d) between the different facial landmarks in each frame and further calculate the relative distance (r) between the two frame (D_1f_i *and* D_2f_i,) which form the feature vector for detecting distortion changes in the face frames. Dlib 68 facial landmark predictor[76] is used to detect 68 facial landmarks from each of the video frames. As shown in Figure 5.2, the 68 facial landmarks are located at various regions of a face, including:

- The chin region with 17 points
- The eyebrows region with 10 points. 5 points for each eyebrow
- The nose stem region with 4 points
- Below nose with 5 points
- The eyes region with 12 point. 6 point for each eye
- And the lips or mouth region with 20 points,

The facial landmarks are denoted as ($p1$, $p2$, ..., $p68$) where $pi = (xi, yi)$ is the coordinate.

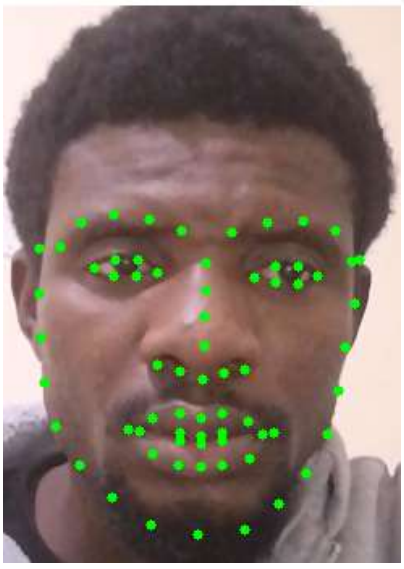

Figure 5.2. A face with detected landmarks using Dlib 68 facial landmarks detector

For each pairwise frames D_1f^i and D_2f^i, *i=(1,2,..., k.),* capturing the facial distortion is done by first calculating their geometric distance between any two face landmarks ps and pt for each frame using the formula below:

$$d = \sqrt{(xs - xt)^2\ (ys - yt)^2} \qquad 1$$

Where s, $t \in \{1, 2, ..., 68\}$ and $s \neq t$

The 68 face landmarks in each frame yielded 2278 pairwise distances $d1, d2, ..., d2278$. This 2278 pairwise distances are concatenated with the width (w) and height (h) of the detected face to form the geometric vector $(geo) = (d1, d2, ..., d2278, w, h)$. Then the relative distance for the pairwise frames D_1f_i and D_2f_i is computed as $r = (r1, r2, ..., r2278, rw, rh)$, where

$$r_j = \frac{d_{D1,i}}{d_{D2,i}} \quad 2$$

for $i = 1, 2, ..., 2278$,

$$rw = \frac{w_{D1,}}{w_{D2,}} \quad 3$$

$$rh = \frac{h_{D1,}}{h_{D2,}} \quad 4$$

Therefore, facial distortion each pairwise selected frame D_1f^i and D_2f^i is represented by a 1×2280 feature Vector (FV).

5.2.3. Classification Module

Finally, the Classification module takes the feature vector (FV) as input and utilizes the classification algorithms to determine whether the feature vector is extracted from a genuine face or spoof attack.

Different classification algorithms are exploited in this module for comparison including both machine learning models (Support Vector Machine, Linear Discriminant Analysis and K-Nearest Neighbor) and deep learning model (Convolutional Neural Network). These models are discussed in section 5.3 of this chapter.

5.3. Data Collection and Dataset Generation

This section describe entire process employed for data collection and also the techniques for generating the datasets are discussed. The four different datasets generated which will be discussed include: the legitimate dataset, the printed photo attack dataset, the screen photo dataset, and the replay attack dataset.

5.3.1. Data Collection

For the purpose of this study, eighty (80) volunteers were used for the data capturing, consisting of 56 men and 22 females. All volunteers are between the ages of 15 to 40 years.

For the data collection, a simple interface program called “Capture Control” was written using python which aided the data collection. The control capture program helps to control the facial video captures at two predefined distances between the face and the camera. While the mobile phone does the capturing, the video files are saved directly on the laptop. The mobile phone and the laptop are connected using a webcam application called “iriun” which allows the mobile phone camera to serve as a web camera for the laptop.

In the second part, we collect the participants’ selfie facial videos at two different device positions with the face covering the middle area of the image. Each participant is requested to take 2 selfie frontal face video clips of 1080HD at 24fps (frames per second) using the mobile phone at controlled distances D1 and D2 between the mobile phone used and his/her face. The video clips for each distance are set to last for 3 seconds where each frame is 1920 × 1080 pixels in size with the face covering the middle of the frame. A program with simple interface (see Figure 5.3) is designed to aid the control and capturing of the face videos at two (2) different distance (20cm and 50cm) from the device camera. The participant first holds the mobile phone and tries to fit his face in the ellipse shape on the interface (see Figure 5.3). As shown in Figure 5.4, once the face is fits to the ellipse shape (implying 50cm distance between the face and the camera), the program automatically captures the facial video frames for 3 seconds and adjust the parameters for the next distance (see Figure 5.5). The ellipse shape automatically increases so as to capture the face at a lower distance(20cm). So in total, for each participant, 2 facial video data (at D1=20cm and D2=50cm) are obtained.

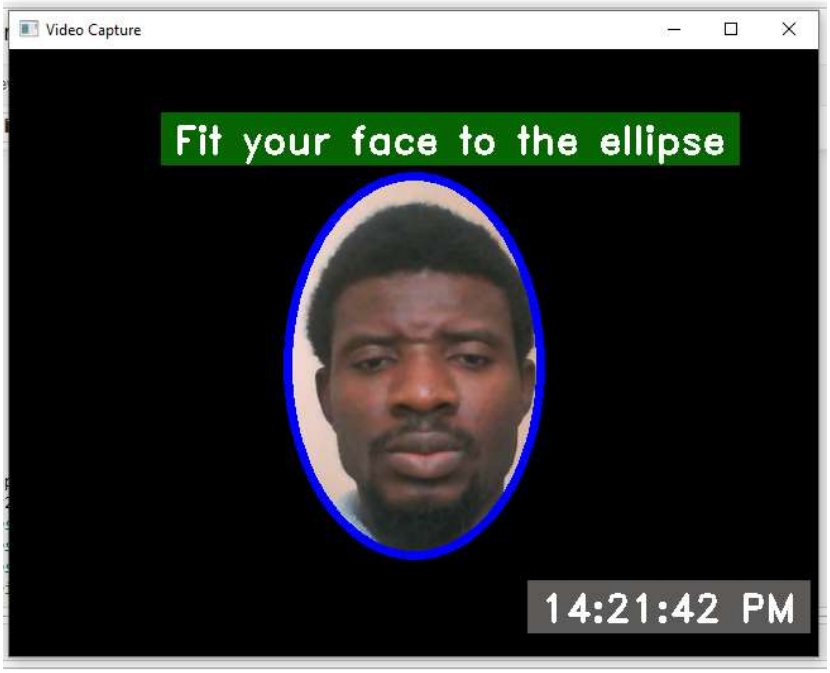

Figure 5.3. Interface with ellipse set to capture face at 50cm away from the camera

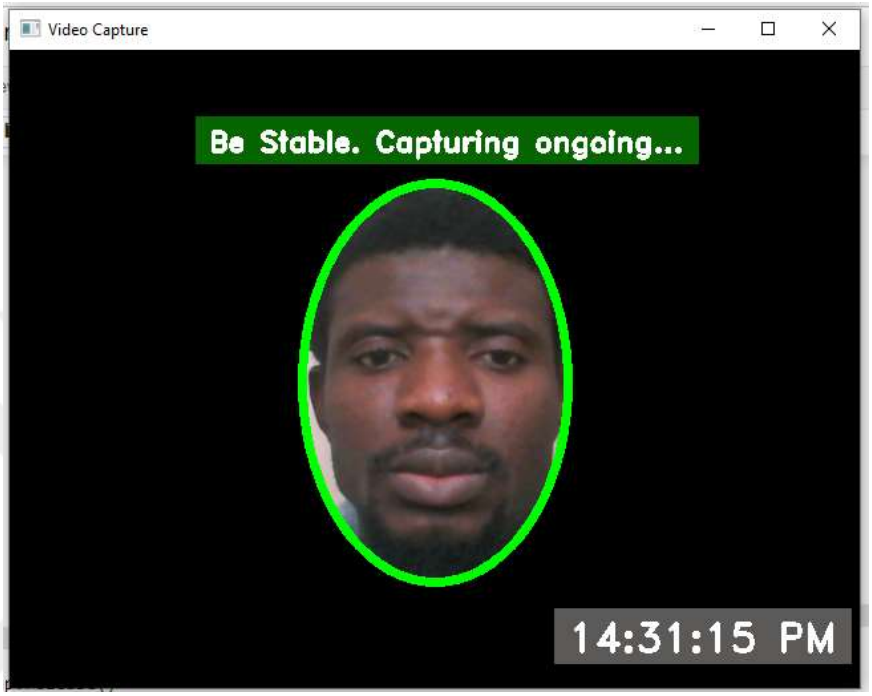

Figure 5.4. Interface capturing detected face at 50cm away from the camera

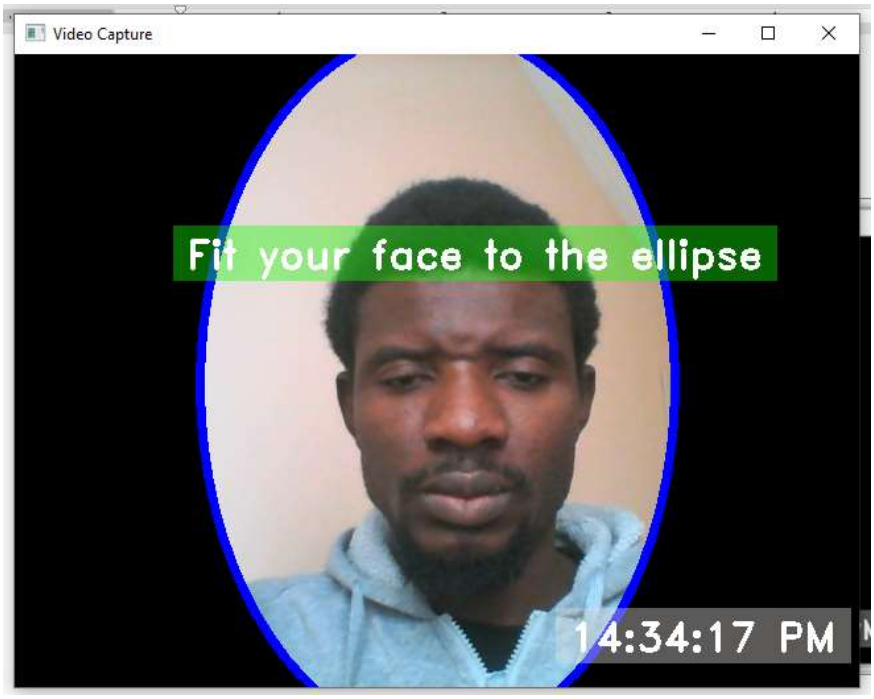

Figure 5.5. Interface with ellipse set to capture face at 20cm away from the camera

Figure 5.6. Sample face images from the data collected

The above Figure 5.6 shows image samples from the data collected. Images on the left side depict face images captured at 20cm distance between the face and the camera. Whereas images on the right side of the table depict face images captured at a distance of 50cm between the face and the camera.

5.3.2. Dataset Generation

As mentioned in section 5.1 of this chapter, four different dataset classes were generated which includes the genuine and spoof cases: Legitimate or genuine Face dataset, printed-photo attack dataset, screen-photo attack dataset and replay attack dataset were generated from the facial videos and photos collected.

Legitimate Dataset

The legitimate dataset consists of the facial videos captured during using the mobile phone with the aid of the control program with simple interface from the distance D1 = $20cm$ to the distance $D2 = 50cm$ as discussed in section 4.1. Therefore, the legitimate dataset consists of eighty (80) videos at D1=20cm and eighty for D1=50cm. That is 80 pairwise videos.

Printed-Photo Attack Dataset

For the printed photo attack, one facial frame is first manually extracted from the frontal facial video clips captured from each participant at the fixed distance of 50cm between the camera and face. The choice of 50cm is because majority of people will not release their facial photos/videos taken at such a close range as 20cm due to the obvious distortion that will be seen on the face image. Secondly, these extracted facial frames are then printed on a A3 paper (so that the size of the face will close to that of a real face) using a quality industrial Direct Impression printer.

Finally, each of the printed photos are then presented to the controlled capture system as describe in section 5.2.1. using the mobile phone and the simple interface program. The system records 2 facial video clips for each image at controlled distance D1=20 and D2=50cm, each for 3 seconds at 24fps. In total, the printed-photo attack dataset consists of 80 pairwise videos (that is 80 for D1=20cm and 80 for D2=50 for each photo).

Screen-Photo Attack Dataset

For the screen-photo attack dataset, the manually extracted facial frames from frontal facial video clips captured at distance D= 50cm between the camera and face, are display to the controlled capture system as describe in section 5.2.1 using a mobile device. the system records 2 videos for each image at controlled distance D1=20 and D2=50cm, each for 3 seconds at 24fps. In total, the screen-photo attack dataset consists of 80 pairwise videos (that is 80 for D1=20cm and 80 for D2=50 for each photo).

Replay Attack Dataset

For the replay Attack dataset, the frontal facial video clips captured from each participant at the fixed distance of 50cm between the camera and face we used. Each video clip is replayed on a looped mode using a mobile device and presented to the controlled capture system consisting of the HUAWEI mobile phone and the simple interface program. The replay device is adjusted closer or farther from the camera until the face region fits to the ellipse, then the system records the video. The system records 2 facial video clips for each replayed video at controlled distance D1=20 and D2=50cm, each for 3 seconds at 24fps. Hence, a total of 80 pairwise videos (that is 80 for D1=20cm and 80 for D2=50 for each photo) are obtained for the replay attack dataset.

5.4. Classification Models

This section discusses the various classification models that were employed for training of the dataset. Four different classification algorithms were exploited consisting of both machine and classifiers and deep learning models, which include: Support Vector Machines (SVM), Linear Discriminant Analysis (LDA), K-Nearest Neighbor KNN) and Convolutional Neural Network (CNN).

5.4.1. Support Vector Machine (SVM)

Support vector machine (SVM) is a kind of classifier with generalized model for performing binary classifications based on supervised learning. By utilizing the kernel method, nonlinear classification tasks can also be accomplished. Because it has an outstanding property in terms of sparsity and robustness, SVM is often employed in resolving face recognition related tasks [77]. The decision boundary of Support Vector Machine is the maximum margin hyperplane of the learning samples' solution. Also, in order to optimize structural risks SVM makes use of hinge loss functions for the calculation of empirical risks and adding regularization terms to the solution system. Face spoof detection are mostly considered as binary classification task, of which SVMs are classifiers with the great potential to cope with such task. It is worth of note that, the feature size obtained in this work through the hand-crafted feature-based approach is relatively large, and Support Vector Machines performs greatly when learning datasets of high-dimensional feature vectors.

5.4.2. Linear Discriminant Analysis (LDA)

Face recognition as a field has seen a broad application of Linear discriminant analysis (LDA) which as a supervised approach [78], so also in face presentation attack detection [79]. LDA can be used for both dimensionality reduction and classification. The objective of LDA is to find a

proper projection that maximizes the between-class scatter matrix and minimizes the within-class scatter matrix in the projective feature space. In the past, image data are often directly used as input for classification, but in dealing with the high-dimensional face data, linear discriminant analysis often suffers from the problem of small sample size. Since the feature size obtained in this work through the hand-crafted feature-based approach is relatively large and of high dimension, it is assumed that LDA is a good fit.

5.4.3. K-Nearest Neighbor (KNN)

The KNN classifier is viewed to have the simplest procedure amongst other machine learning classifiers. More to that In [80], authors approach spoof detection task using Local Binary Pattern descriptors, analysis of their result shows that KNN performed better both in terms of accuracy and execution time. In the same vein authors in [81] extract features using the Eigen vector technique. Their result also shows a better performance by KNN as compared to SVM both in terms of accuracy and execution time. KNN is also employed in this work for comparison.

5.4.4. Convolutional Neural Network (CNN)

Deep learning can be harness achieve promising results in computer vision as a field, and has proven effective in tackling face spoof detection task [72]. The CNN architecture used in this worked is described below:

Firstly, is a 1-dimentional convolutional layer (Conv1D) with 32 units as the input layer, with input dimension of (2280 x1) to extract the various features from the input images. The kernel of the input layer is set to 3 with a Relu activation function to learn and approximate the relationship between variables of the network. Then it is followed by a Maxpooling stride 2 to decrease the size of the feature map obtain from the Convolutional layer in order to reduce the computational costs. Thirdly is a fully connected layer with 1024 units. And lastly is a fully connected layer with 2 units and a sigmoid function (since it is binary classification) to serve as the output layer. Two Dropout layers set at 0.4 were introduce to tackle overfitting. One after the pooling layer and the other after the first fully connected layer. Also a flatten layer is introduced just before the fully connected layer. Adam optimizer is utilized and the model is set at a learning rate of 0.001 and decay of 1e-6. The summary of the model is displayed as shown in Table 5.2.

Table 5.2. Summary of the CNN architecture employed

#Layer	Layer type	Activation Function	Value
1	Convolution (Conv1D)	Relu	Unit=32 Kernel size=3
2	MaxPooling1D	None	Stride= 2
3	Dropout	None	0.4
4	Flatten	None	None
5	Dense	Relu	Unit=1024
6	Dropout	None	0.4
7	Dense	Sigmoid	Unit=2

5.5. Evaluation Metrics

Kanika and Jaspreet [82] outlined five parameters commonly used for the evaluation in face spoof detection. They include: False rejection rate (FRR), false acceptance rate (FAR), equal error rate (EER), half total error rate (HTER) and accuracy. These parameters help in determining real threats that affects any face recognition system by a spoofing database. These metrics will be explain using the confusion metrics parameters described in Figure 5.7 below.

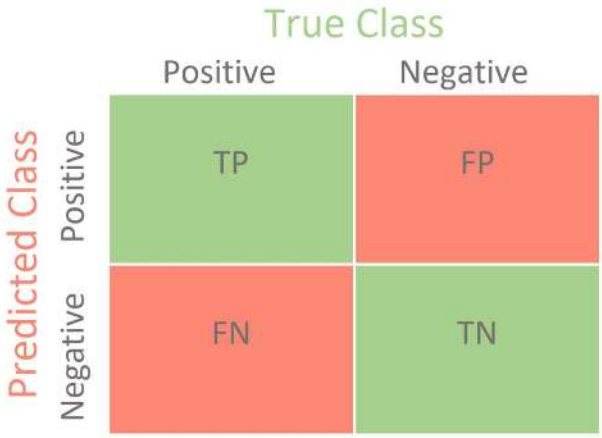

Figure 5.7. Confusion matrix chart for binary classification

- **True Positive(TP):** The number of attacks recognized as the attacks.
- **True Negative (TN):** The number of genuine samples recognized as genuine ones.
- **False Positive (FP):** The number of genuine samples recognized as attacks.
- **False Negative (FN):** The number of attacks recognized as genuine samples.

5.5.1. False Acceptance Rate (FAR)

The false acceptance rate, or FAR, is the measure of the likelihood that the biometric security system will incorrectly accept an access attempt by an unauthorized user. It is simply the percentage of identification instances in which unauthorized persons are incorrectly accepted. A system's FAR typically is stated as the ratio of the number of false acceptances divided by the number of identification attempts. The formula for FAR is as show below:

$$FAR = \frac{FP}{FP+TN} \quad 5$$

5.5.2. False Rejection Rate (FRR)

The false rejection rate account for the tendency of a biometric system to incorrectly reject or deny access to a genuine user. It is simply the percentage of identification instances in which authorized persons are incorrectly rejected. A system's FRR typically is stated as the ratio of the number of false rejections divided by the number of identification attempts. The formula is stated below:

$$FRR = \frac{FN}{TP+FN} \quad 6$$

5.5.3. Equal Error Rate (EER)

A decrease in false acceptance rate (FAR), will result in an increase in false rejection rate (FRR) and vice versa. The point at which the lines meet is referred to as Equal Error Rate (EER). As seen in Figure 5.8, the EER is the point of equality in FAR and FRR. Equal Error Rate, inherently references to an algorithmic approach of Error margin, where you equalize false rejections and false acceptances. This rate, is the common value, or common ground, of said result. The lower this error margin is; the greater accuracy the biometric system has. Thus, basically, it’s a Mathematical way of scoring out errors and error margins in terms of false data.

EER is important to determine accuracy of data, and make comparative analysis between two systems and results.

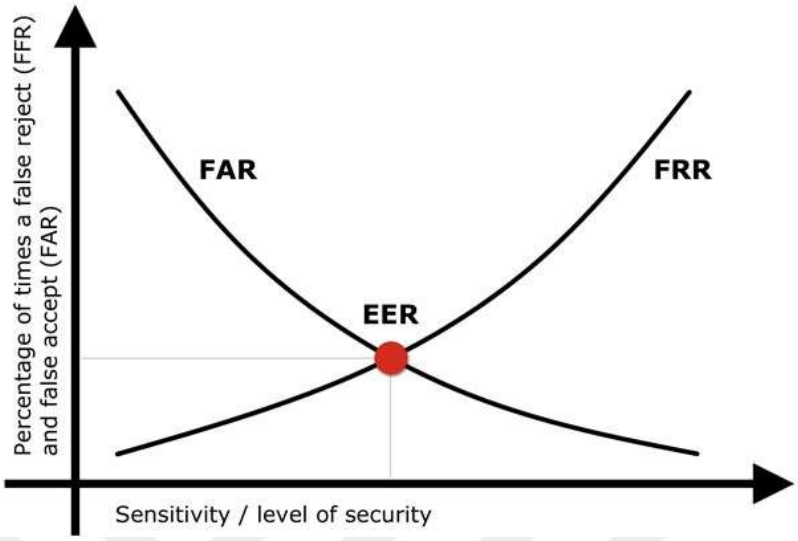

Figure 5.8. Illustration of EER derived from the plot of FAR vs FRR

Reducing the false acceptance rate to barest minimum, is likely to cause a sharp rise in the false rejection rate. In other words, the more secure your access control, the less convenient it will be, as users are falsely rejected by the system. Therefore, it can be stated that the choice of FAR or FRR is a matter of preference that will results in either a system that is more secure (but less user friendly) or less secure (but more user friendly).

Few systems exist in the market which allows a high level of security to be achieved and at same time having user-friendly access control. If an organization does not opt for such a system, it will often priorities user convenience over security [83].

5.5.4. Half Total Error Rate (HTER)

The half total rate is an error rate that corresponds to the average of the FAR and the FRR error rates. The formula HTER is as shown below.

$$HTER = \frac{FAR+FRR}{2} \quad 7$$

Most biometric systems make use of HTERs or its other variations for measurement and comparison. Bear in mind that in most benchmark databases used in literatures relating to biometric systems, there is a notable imbalance between the number of genuine accesses and that of impostors accesses. In all probability, it is owing to the relatively greater cost of obtaining the former with respect to the latter. This imbalance account for the reason HTER is used for comparison of models and not the usual classification error used in the machine learning literatures [84].

5.5.5. Accuracy

The accuracy of a biometric system is simply the measurement of ratio of correctly predicted samples as compared to the total predictions. It is calculated thus:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+F} \quad 8$$

It should be noted that the smaller EER or HTER represents the better detection result.

6. RESULTS AND DISCUSSION

The results of this work will be presented and discussed in terms of five (5) evaluation metrics described in section 5.4. which include: Accuracy, false acceptance rate, False rejection rate, equal error rate, and half total error rate. Also three different analytical experiment were carried to verify the effectiveness of the approach.

6.1. Experimental Setup

To determine the liveness of a face, the frame selection module firstly selects K-frames from each pairwise facial videos taken at distance D1=20cm and D2=50cm as discussed in Section 5.1.1. For each pairwise frames D_1f_i *and* D_2f_i, for i= (1,2, …, K), the feature extraction module extracts the features forming a vector of 1 x 2280 as described in section 5.1.2. For this study, K is to 30 for attack samples, while K is set to 60 for genuine samples. This is in order to avoid the case of imbalance in the classes. Therefore, for each pairwise videos of attack cases, 30 samples are extracted making 30 x 80 =2400 for each spoof case with a total of 2400 x 3(printed-photo, screen-photo and replay attack) = 7200 samples are generated for the attack dataset. Also for the genuine videos, 60 x 80=4800 samples are generated for the legitimate dataset. In total a dataset of 12000 samples are obtained which forms a 12000 x 2280 matrix for the legitimate and spoof data samples.

Three experiment are carried out to validate the effectiveness of the approach. First, four different models (SVM, LDA, K-NN and a CNN model) were utilized in training of the entire dataset of 12000 samples in ratio of 3:1 for training and testing (9000 and 3000 respectively). Results are presented in section 6.2.

Secondly, the datasets composed of each spoof case separately are been fed to the trained models for classification in order to have an insight into their performances based on the various spoof cases. Results are shown in the figures in section 6.2.

And finally, each model is saved. And a simple program is designed that incorporate the saved models and the video capture controlled system described in section 5.2 to test all the model in real time (see Figure 6.5, Figure 6.6 and Figure 6.7). The systems were presented with live face and all the attack types.

6.2. Presentation of Results and Discussion

Four algorithms are being utilize for the training of the datasets. Which include: SVM, LDA, K-NN and CNN. The result of these algorithms are presented and compared and discussed in the next five sections.

6.2.1. Results from SVM Classifier

The SVM classifier is first train on the main dataset consisting of 12000 samples, of which 4800 are legitimate, while 7200 are spoof cases. As seen in the Table 6.1, SVM classifier attained an accuracy of 96.07%. This indicates the general ability of the model to correctly identify both real faces and spoof faces. It also attains an FAR and FRR of 0.053 and 0.031 respectively. A false acceptance rate of 0.053 implies that in a sample of one thousand (1000), the model is likely to incorrectly accept or grant access to 53 unauthorized users. Likewise, a false rejection rate of 0.031 implies that in a sample of one thousand (1000) genuine faces, the model can correctly accept 969 while failing in only 31cases. This also shows that the SVM model's performance in identifying genuine faces is slightly better than its ability to correctly identify spoof cases. The model also attained EER and HTER of 0.034 and 0.042 respectively, which are indications of good performance as smaller value of EER and HTER, the better the performance.

Table 6.1. Performance Results of the SVM Model on the dataset

SVM Performances				
Accuracy	**FAR**	**FRR**	**EER**	**HTER**
96.07%	0.053	0.031	0.034	0.042

For the second experiment where the SVM model is tested for performance on individual spoof cases. As seen in Figure 6.1 below, result shows that it is more robust in detecting screen-photo attacks than the printed-photo and replay attacks.

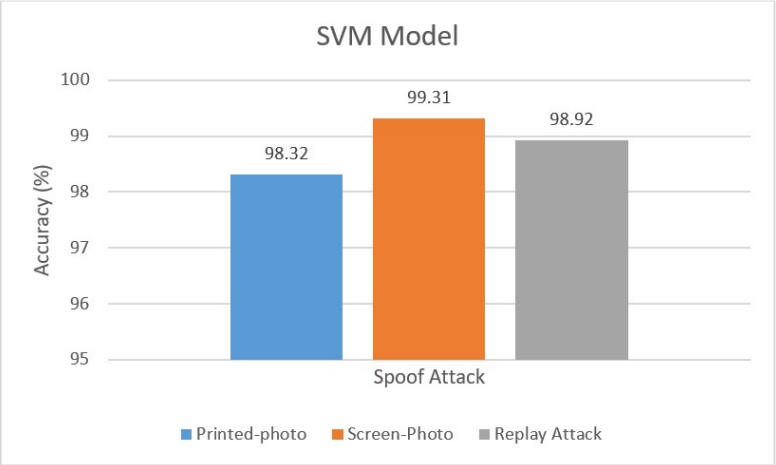

Figure 6.1. Graphical Accuracy of the SVM classifier on each attack type

6.2.2. Results from LDA Classifier

Result from linear discriminant analysis reveals a very impressive performance across all metrics with accuracy of 98.03%, FAR and FRR of 0.022 and 0.018 respectively. This shows that in a sample of one thousand (1000), the LDA is likely to falsely accept or grant access to only 22 unauthorized users. Likewise, a false rejection rate of 0.018 implies that in a sample of one thousand (1000) genuine faces, the model is likely to wrongly reject or deny access to as low as 18 authorized users. An EER and HTER 0.014 and 0.020 shows the true robustness of the model (See Table 6.2).

Table 6.2. Performance of the LDA Classifier on the dataset

LDA Performances				
Accuracy	**FAR**	**FRR**	**EER**	**HTER**
98.03%	0.022	0.018	0.014	0.020

Figure 6.2 shows that the performance of Linear Discriminant Analysis on the unit spoof attacks testing performs slight better in detecting screen photo attacks, followed by the printed-photo attack, with the least performance in replay attack with accuracy of 99.44%, 99.48% and 99.26% respectively.

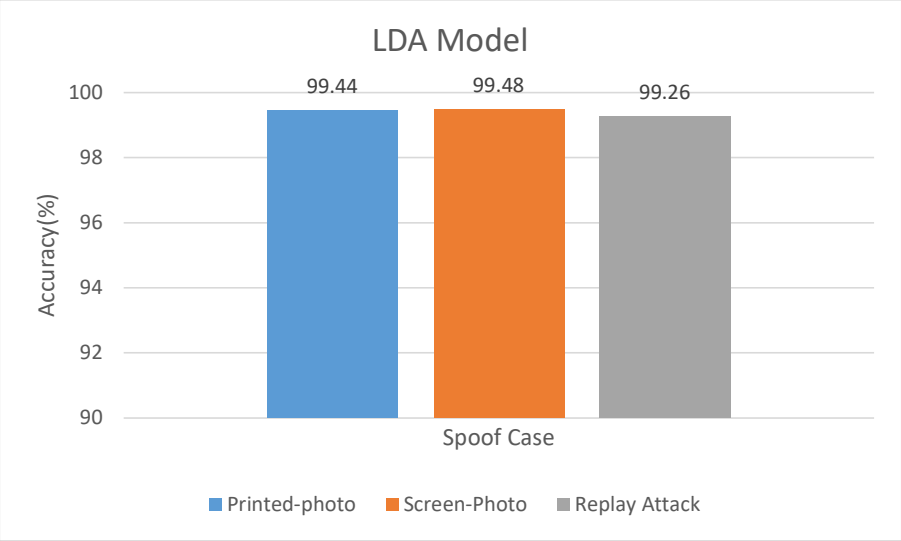

Figure 6.2. Accuracy of the LDA classifier on each attack type

6.2.3. Results from K-NN Classifier

The K-NN classifier has the least performance with an accuracy of 92.90%, FAR and FRR of 0.082 and 0.062 respectively (See Table 6.3). This implies when presented with a sample of 1000 genuine and spoof faces, each a wrong prediction of 82 and 63 are expect. The low performance could be as result of high dimensionality of the dataset. This correlate fact that K nearest neighbors (KNN) are known as one of the simplest nonparametric classifiers but in high dimensional setting its accuracy is affected by nuisance features [85].

Table 6.3. Performance of the K-NN Classifier on the dataset

K-NN Performances				
Accuracy	**FAR**	**FRR**	**EER**	**HTER**
92.90%	0.082	0.063	0.053	0.073

In the unit testing phase, as shown in Figure 6.3, contrary to other classifiers which were more robust in detecting screen-photo attacks, the K-NN classifier had its best performance in Printed-photo attacks with 97.07% accuracy which is slightly better than the performance in screen-photo and replay attacks having 96.64% and 94.52% respectively as the accuracies.

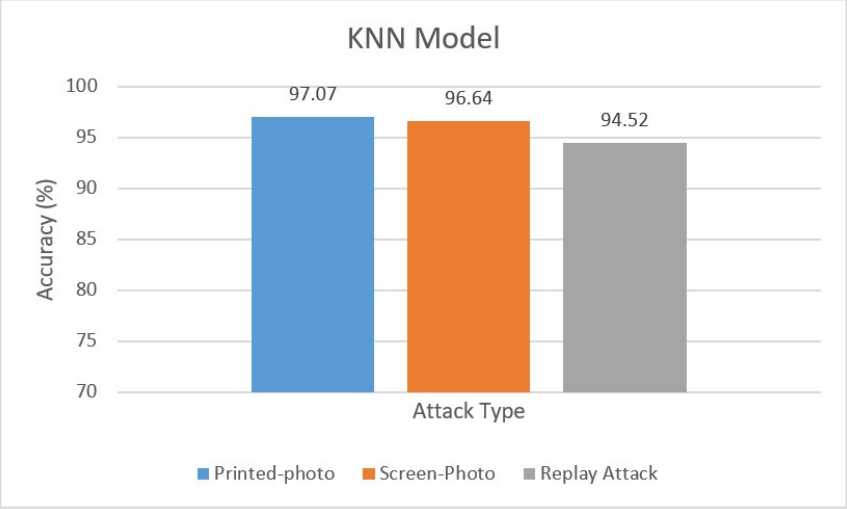

Figure 6.3. Accuracy of the K-NN classifier on each attack type

6.2.4. Results from CNN Model

The CNN model is trained using the architecture described in section 5.4.3 with learning rate set to 0.001, weight decay is 1e-6 and the max iteration of 100 accordingly. The result as seen in Table 6.4 below shows that the CNN model performed very impressively with an accuracy as high as 98.18%, with FAR and FRR of 0.036 and 0.007 respectively. An FRR as low as 0.007 shows the model has a very high recall for genuine samples. Also CNN model attained an EER and HTER as lows 0.23 and 0.21, which is an indication of good performance.

Table 6.4. Performance of the CNN Model on the dataset

CNN Model				
Accuracy	**FAR**	**FRR**	**EER**	**HTER**
98.18%	0.036	0.007	0.023	0.021

In the unit testing phase, as shown in Figure 6.4 below, like SVM and LDA classifiers, the CNN model performed better in the screen-photo attack than in the other two attacks with accuracy of 99.74%, 99.66% and 99.53% for screen-photo attack, replay attack and printed-photo attacks respectively.

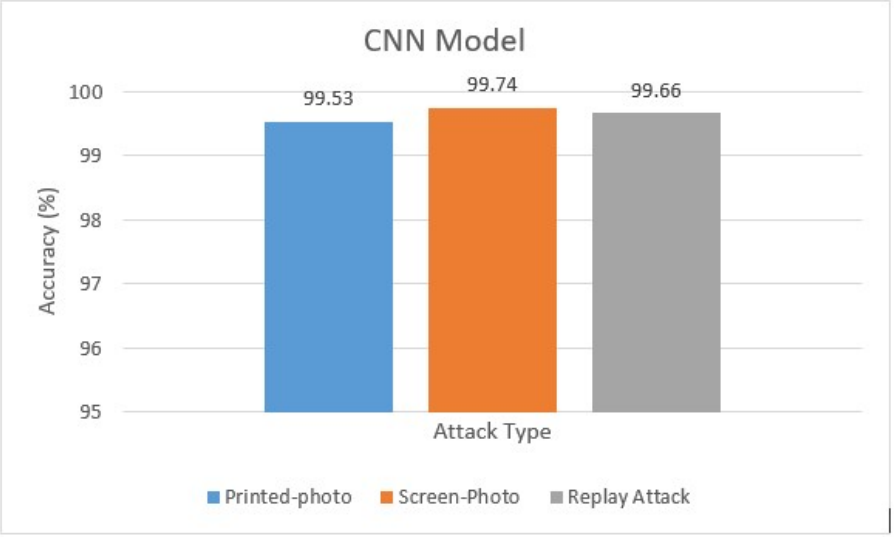

Figure 6.4. Accuracy of the CNN Model on each attack type

6.3. Validation on Real-Life Test Scenario

Here the results of the real live test cases using the saved trained models and the capture control system will be presented based on each test case. Figures 6.5 and Figure 6.6 shows the capturing of live face at distances 50cm and 20cm respectively from the camera. While Figure 6.7 is the output of the models performances

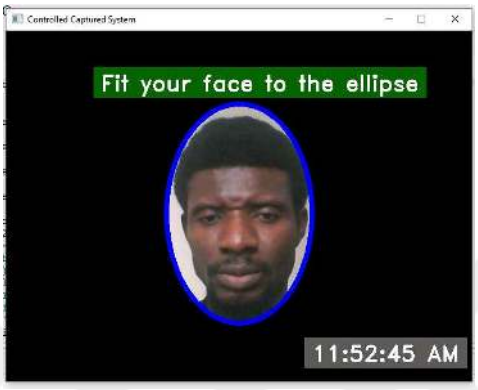

Figure 6.5. The capture control system capturing live-face at distance=50cm

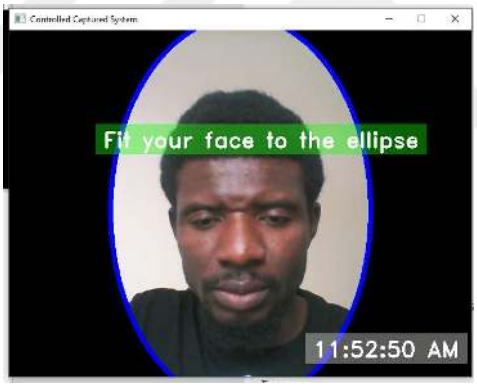

Figure 6.6. The capture control system capturing live-face at distance=50cm

Figure 6.7. Image depicting the result of the models on live face

Test result on the various attacks

Here live tests are performed using the different attack types. Figure 6.8 shows the output of the printed-photo attack; Figure 6.9 is the result of the screen-photo attack scenario; And finally Figure 6.10 is the result of the replay attack live test.

Figure 6.8. Result of the models on printed-photo attack

Figure 6.9. Result of the models on screen-photo attack

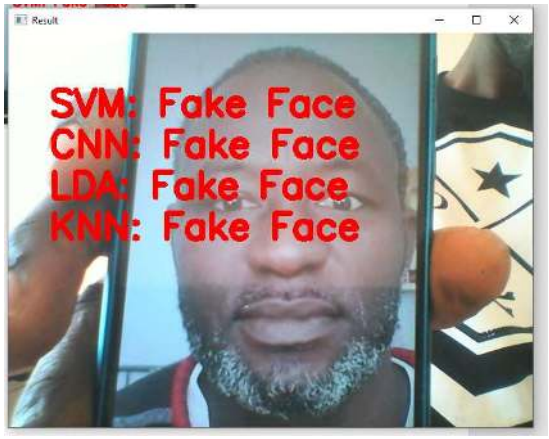

Figure 6.10. Result of the models on replay attack

6.4. Comparison of Models Performances

Table 6.5 below is a comparative table of the performances of all the model across all the five metrics for evaluation. Values in bold shows best performance on that metrics

Table 6.5. Comparison of Models' Performances across the five metrics

Model	Accuracy	FAR	FRR	EER	HTER
SVM	96.07%	0.053	0.031	0.034	0.042
LDA	98.03%	**0.022**	0.018	**0.014**	**0.020**
KNN	92.90%	0.082	0.063	0.053	0.073
CNN	**98.18%**	0.036	**0.007**	0.023	0.021
Best Permanence	**98.18%**	**0.022**	**0.007**	**0.014**	**0.020**

Also the saved models were tested against time. Each model is fed with an input of 2400 data samples for prediction. Figure 6.11 reveals their performances. The LDA model is extremely fast with the best performance with respect to time taken in predicting the entre 2400 samples within in approximately 3miliseconds. Whereas the CNN model with the highest execution of 10.901seconds. The SVM and KNN performance with respect to time of execution are almost the same with 3.712seconds and 3.869 seconds for SVM and LDA respectively.

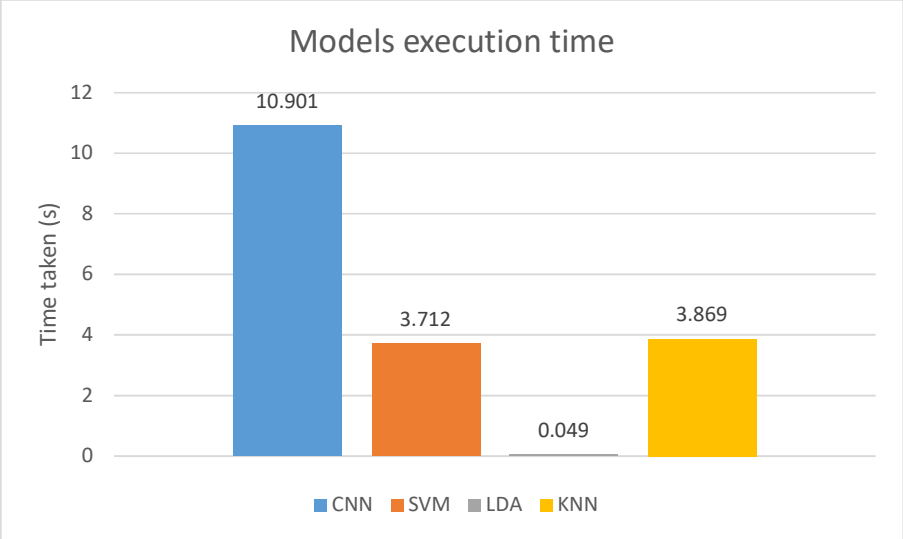

Figure 6.11. Results of the models performances against time taken for prediction

6.5. Comparison of other active-based approaches with the proposed approach

Table 6.6 below is a comparison of the proposed approach to other approaches while highlighting the their strength and limitations.

Table 6.6. Comparison of other approaches with the proposed approach

Technique	Category	Strength	Limitation	Performance
Fusion of face and Voice [19]	Active	• Very robust • Can detect 3D attacks	• Highly Intrusive • Extra Hardware requirement • Low application range • Expensive • Computational complexity	FAR=0.017 FRR= 0.011 FTC= 0.031 EER= 0.01
Pupil tracking using led sensor[5]	Active	• Robust against print attacks	• Intrusive • Helpless in cut-photo and replay attacks • Extra Hardware requirement • Cannot easily fit into existing system • Time consuming	-
DoG[33]	Passive	• Fast response time • Non-Intrusive	• Poor generalization (Vulnerable to variation in acquisition) •	EER=0.17
Ensemble of Texture descriptors (LBP+GDP+GLTP+LDiP+LGBPHS+ LPQ)	Passive	• Non-Intrusive •	• Computational complexity • Poor generalization • Limited to 2D attacks	RR= 0.98.39 FRR= 0.49
3D sensor [7]	Passive	• Robust • Non-intrusive	• Time complexity • Require extra hardware • Helpless in 3D attacks • Cannot fit into existing systems easily	Accuracy=100%
Eye-blinking and Mouth movement[1]	Acvtive	• Robust against • No extra hardware required • Less expensive	• Instrusive • Weak in detecting cut-photo and replay attacks • Slow response time	-
Face and Finger print[10]	Active	• Very robust • Captures 3D attacks	• Highly Intrusive • Extra Hardware requirement • Low generalization ability • Expensive	EER=0.012
Facial distortion changes	Active	• Robust • No extra hardware requirement • Less expensive	• Intrusive • Computational complexity	Accuracy= 99.48%

Temporal and depth information	Passive	• Robust • Non-intrusive	• Computational Complexity • Require extra hardware • Helpless in 3D attacks • Cannot fit into existing systems easily	Intra DB: ACER= 0.73 on SiW; ACER= 1.3 on OULU-NPU Cross DB: HTER = 17.5 on Replay Attack DB; HTER = 24.0 on CASIA-MFSD
Proposed approach	Active	• Can easily fit into existing system • Less intrusive • No extra hardware requirement • Less expensive • Robust • Less complexity	• Intrusive • Limited to 2D attacks	Accuracy=98.18% FAR=0.036 FRR=0.007 EER=0.023 HTER=0.021

6.6. Discussion

The results presented in this work proved that the method proposed in this study is robust and thus viable with all models having an accuracy beyond 92%. Top on the list is the CNN model and the LDA classifier. Performance evaluation using the four models based on five metrics shows that the CNN model performed better than the rest in two of the metrics, with accuracy and false rejection rate (FRR) of 98.18% and 0.007 respectively. Which means that given a set of data samples, the CNN model will produce higher accuracy rate and the tendency of denying access or incorrectly rejecting an authorized user is very low as indicated in an FRR of 0.007. Showing that in a sample of 1000, only 7 incorrect rejections are expected. On the other hand, the linear discriminant analysis outperformed the others in terms of three other metrics which are FAR, EER and HTER with values of 0.022, 0.014 and 0.020 respectively. Although the CNN has a very low FRR, it has a much higher FAR, which is the reason why it is beaten by the LDA in the values of EER and the HTER. A lower EER indicates a good margin between FAR and FRR.

Contrary to the results shown in [80] and[81] where KNN performed better than other classifiers, in this study, KNN has the least performance across all metrics. This correlate with opinion of authors in [85] that in the face of high dimensional feature space KNN performance is affected by nuisance features.

In the test for execution time, CNN had the least performance with 10.901 seconds execution time over 2400 samples prediction. LDA had the best performance with execution time of

0.049seconds (49milliseconds) for the 2400 samples. The fast the good performance of LDA is due to the dimensionality reduction which is not the case with the other models.

A good question to ask is "which model should be chosen between the CNN and LDA in this case?". The fact is, it depends on the goal and priority of the system to be build or deployed to. Some system goal may prioritize FAR where the other may prioritize FRR. Therefore, it can be stated that the choice of FAR or FRR is a matter of preference that will results in either a system that is more secure (but less user friendly) or less secure (but more user friendly)[83]. Consider an instance where you want to avoid people queuing up at the entrance because of the system not working properly (false rejection). This could have caused people to spend all the morning stocked in traffic. In cases, such as these, convenient being the priority may be acceptable to users. Nonetheless, the reverse is the case where high level of security is the expectation of the users. However, generally, when it comes to biometric systems, low equal error rate is worth given due consideration.

In comparison to the efforts in [20] which is closer approach to the one proposed in this work, although they attained a slightly higher accuracy, in their approach, a single sample is represented by a 7 by 2147 matrix as feature set extracted from 8 frames at different distances. Whereas in approach proposed in this study, only two frames captured at 20cm and 50cm from the camera are needed to form the feature vector for a single sample. This makes their approach more computationally complex and time consuming.

Also, table 6.5 clearly highlight the advantages of the proposed approach over other active-based approaches. In [19] and [10] where face recognition is combined with either voice or finger print, a user is also requested to either speak of provide his/her fingerprint which is captured using an audio device or fingerprint reader. Although they are very robust against all forms of face spoof attacks, they are very intrusive and requires extra hardware such as fingerprint reader which makes it also expensive. Whereas in the proposed approach, there is no need of an extra hardware, and it is less intrusive as a user will just need to bring his face closer or farther, which is normal for users before any face recognition system to bring their face closer or farther even without instruction. The effort proposed in [1] which does not require extra hardware and tends to fit into existing systems, a user is requested to blink his/her eye and open/close mouth. Unfortunately, it is highly intrusive and very weak or helpless when faced with cut-photo or replay attacks. Unlike the approach proposed here which has achieved accuracy as high as 99.66% against replay attacks in the unit testing scenario. Beside every user facing such a system knows that the liveness test is based on blinking of the eye and movement of the mouth, as such it suggests to the impostor the next point of action, unlike the method proposed in this study where the liveness cue is not clear to the user.

7. CONCLUSIONS

This study drew inspiration from the common phenomenon in photography on distortion of faces in video where the camera and the face are in close proximity, which is caused by the uneven 3D surface of the human face, and proposed a face spoof detection approach that utilizes the distortion changes of video frames of users face captured at different distance from the camera to deliver a robust active-based spoof detection, yet less intrusive, less expensive and a more generic applicability than other active-based approaches using both traditional machine learning classifiers and a deep learning model. Firstly, video frames of 80 volunteers were captured using front facing camera of a mobile phone at distances 20cm and 50cm between the face and the camera.

Analytical experiment to verify the effectiveness of the approach was carried out and evaluated across five metrics: accuracy, FAR, FRR, EER and HTER. Result shows very impressive and competitive performance with the best performances in CNN and LDA models with accuracy = 98.18%, EER=0.023, HTER=0.021 and accuracy = 98.03%, EER=0.022, HTER=0.020 respectively. The least performance is seen in the K-NN model with accuracy = 92.90%, EER=0.053, HTER=0.073, which shows the weakness of K-NN classifier in the face of high dimension feature space. In the light of the above, it can be concluded that the approach proposed in this work is viable and the handcrafted features are discriminative enough against 2D attacks and can be explored further against other forms of attack such as the 3D attacks. It is also worthy of note that the proposed approach holds several advantages over other active-based approaches including: less intrusiveness, no hardware requirement, fast response time, less computational complexity, and can easily fit into existing system.

RECOMMENDATIONS

Currently, face recognition has become an identity authentication system that has being broadly used, it is continually increasing in terms of acceptance, deployment and application. However, the rising number of attacks and diversity of face spoof attacks is a major concern. Hence, building a robust spoof detection measure against such fraudulent activities has become necessary and unavoidable. Spoof detection tasks are usually viewed as a binary classification problem despite diverse forms of attacks. Hence, it is recommended that future works should address spoof detection in terms of multi-class classification considering the different forms of attacks. By so doing, feedbacks of the common forms of attacks presented to the face recognition system can be collected, analyzed and used for further study and improvement. Also cases of unknown or emerging attacks can also be traced.

REFERENCES

[1] A. Singh, P. Joshi, G.C. Nandi, Face recognition with liveness detection using eye and mouth movement, 2014 Int. Conf. Signal Propag. Comput. Technol. (ICSPCT 2014). (2014) 592–597.

[2] D. Wen, H. Han, A.K. Jain, Face Spoof Detection With Image Distortion Analysis, IEEE Trans. Inf. Forensics Secur. 10 (2015) 746–761. https://doi.org/10.1109/TIFS.2015.2400395.

[3] L. Sun, G. Pan, Z. Wu, S. Lao, Blinking-Based Live Face Detection Using Conditional Random Fields, in: S.-W. Lee, S.Z. Li (Eds.), Adv. Biometrics, Springer Berlin Heidelberg, Berlin, Heidelberg, 2007: pp. 252–260.

[4] C.N. Karson, Spontaneous eye-blink rates and dopaminergic systems., Brain. 106 (Pt 3) (1983) 643–653. https://wwww.unboundmedicine.com/medline/citation/6640274/Spontaneous_eye_blink_rates_and_dopaminergic_systems_.

[5] M. Killioğlu, M. Taşkiran, N. Kahraman, Anti-spoofing in face recognition with liveness detection using pupil tracking, in: 2017 IEEE 15th Int. Symp. Appl. Mach. Intell. Informatics, 2017: pp. 87–92. https://doi.org/10.1109/SAMI.2017.7880281.

[6] T. Dhawanpatil, B. Joglekar, A Review Spoof Face Recognition Using LBP Descriptor, in: A.K. Somani, S. Srivastava, A. Mundra, S. Rawat (Eds.), Proc. First Int. Conf. Smart Syst. Innov. Comput., Springer Singapore, Singapore, 2018: pp. 661–668. https://doi.org/https://doi.org/10.1007/978-981-10-5828-8_63.

[7] G. Albakri, S. Alghowinem, The Effectiveness of Depth Data in Liveness Face Authentication Using 3D Sensor Cameras@, Sensors (Basel). 19 (2019).

[8] Y. Ma, L. Wu, Z. Li, F. liu, A novel face presentation attack detection scheme based on multi-regional convolutional neural networks, Pattern Recognit. Lett. 131 (2020) 261–267.

[9] L. Feng, L.-M. Po, Y. Li, X. Xu, F. Yuan, T.C.-H. Cheung, K.-W. Cheung, Integration of image quality and motion cues for face anti-spoofing: A neural network approach, J. Vis. Commun. Image Represent. 38 (2016) 451–460. https://doi.org/https://doi.org/10.1016/j.jvcir.2016.03.019.

[10] P. Wild, P. Radu, L. Chen, J. Ferryman, Robust multimodal face and fingerprint fusion in the presence of spoofing attacks, Pattern Recognit. 50 (2016) 17–25. https://doi.org/https://doi.org/10.1016/j.patcog.2015.08.007.

[11] B. Geng, C. Lang, J. Xing, S. Feng, W. Jun, MFAD: A Multi-modality Face Anti-spoofing Dataset, in: 2019: pp. 214–225. https://doi.org/10.1007/978-3-030-29911-8_17.

[12] T. Chugh, A.K. Jain, Fingerprint Spoof Detection: Temporal Analysis of Image Sequence, CoRR. abs/1912.0 (2019). http://arxiv.org/abs/1912.08240.

[13] V. Holub, J. Fridrich, Digital Image Steganography Using Universal Distortion, in: IH MMSec 2013 - Proc. 2013 ACM Inf. Hiding Multimed. Secur. Work., 2013. https://doi.org/10.1145/2482513.2482514.

[14] J. Cheng, A.C. Kot, S. Rahardja, Steganalysis of Binary Cartoon Image using Distortion Measure, in: 2007 IEEE Int. Conf. Acoust. Speech Signal Process. - ICASSP '07, 2007: pp. II-261-II–264. https://doi.org/10.1109/ICASSP.2007.366222.

[15] P.P.D. Raval, R.R. Sedamkar, S. Kulkarni, Face Spoofing Detection Using Image Distortion Features, in: 2017.

[16] M. Bryson, M. Johnson-Roberson, O. Pizarro, S. Williams, Colour-Consistent Structure-from-Motion Models using Underwater Imagery, in: 2012. https://doi.org/10.15607/RSS.2012.VIII.005.

[17] X. Sun, L. Huang, C. Liu, Context based face spoofing detection using active near-infrared images, in: 2016: pp. 4262–4267. https://doi.org/10.1109/ICPR.2016.7900303.

[18] Mohamed, Shaimaa, Ghoneim, Amr, Youssif, Aliaa, Visible/Infrared face spoofing detection using

texture descriptors, MATEC Web Conf. 292 (2019) 4006. https://doi.org/10.1051/matecconf/201929204006.

[19] B.R. Naidu, P.V.G.D. Reddy, Fusion of face and voice for a multimodal biometric recognition system, Int. J. Eng. Adv. Technol. 8 (2019) 506–515.

[20] Y. Li, Z. Wang, Y. Li, R. Deng, B. Chen, W. Meng, H. Li, A Closer Look Tells More: A Facial Distortion Based Liveness Detection for Face Authentication, in: Proc. 2019 ACM Asia Conf. Comput. Commun. Secur., Association for Computing Machinery, New York, NY, USA, 2019: pp. 241–246. https://doi.org/10.1145/3321705.3329850.

[21] S. Kumar, S. Singh, J. Kumar, A comparative study on face spoofing attacks, in: 2017 Int. Conf. Comput. Commun. Autom., 2017: pp. 1104–1108. https://doi.org/10.1109/CCAA.2017.8229961.

[22] J. Galbally, S. Marcel, J. Fierrez, Biometric Antispoofing Methods: A Survey in Face Recognition, IEEE Access. 2 (2014) 1530–1552. https://doi.org/10.1109/ACCESS.2014.2381273.

[23] D. Menotti, G. Chiachia, A. Pinto, W.R. Schwartz, H. Pedrini, A.X. Falcão, A. Rocha, Deep Representations for Iris, Face, and Fingerprint Spoofing Detection, IEEE Trans. Inf. Forensics Secur. 10 (2015) 864–879. https://doi.org/10.1109/TIFS.2015.2398817.

[24] A. Pinto, W. Schwartz, H. Pedrini, A. Rocha, Using Visual Rhythms for Detecting Video-Based Facial Spoof Attacks, Inf. Forensics Secur. IEEE Trans. 10 (2015) 1025–1038. https://doi.org/10.1109/TIFS.2015.2395139.

[25] J. Yang, Z. Lei, D. Yi, S.Z. Li, Person-Specific Face Antispoofing With Subject Domain Adaptation, IEEE Trans. Inf. Forensics Secur. 10 (2015) 797–809. https://doi.org/10.1109/TIFS.2015.2403306.

[26] Y. Li, K. Xu, Q. Yan, Y. Li, R.H. Deng, Understanding OSN-based Facial Disclosure against Face Authentication Systems. (2014), in: A.C.C. S. (Ed.), 14 Proc. 9th ACM Symp. Information, Comput. Commun. Secur. June Kyoto. 413-424. Res. Collect. Sch. Inf. Syst. Available, 2014: pp. 4–6.

[27] I. Chingovska, A. Anjos, S. Marcel, On the effectiveness of local binary patterns in face anti-spoofing, 2012 BIOSIG - Proc. Int. Conf. Biometrics Spec. Interes. Gr. (2012) 1–7.

[28] N. Erdogmus, S. Marcel, Spoofing in 2D Face Recognition with 3D Masks and Anti-spoofing with Kinect, in: Biometrics Theory, Appl. Syst., 2013.

[29] Jinjin Zhang, D. Huang, Y. Wang, J. Sun, Lock3DFace: A large-scale database of low-cost Kinect 3D faces, in: 2016 Int. Conf. Biometrics, 2016: pp. 1–8. https://doi.org/10.1109/ICB.2016.7550062.

[30] S. Liu, B. Yang, P.C. Yuen, G. Zhao, A 3D Mask Face Anti-Spoofing Database with Real World Variations, in: 2016: pp. 1551–1557. https://doi.org/10.1109/CVPRW.2016.193.

[31] A. Anjos, S. Marcel, Counter-Measures to Photo Attacks in Face Recognition: a public database and a baseline, in: Int. Jt. Conf. Biometrics 2011, 2011. http://pypi.python.org/pypi/antispoofing.motion.

[32] X. Tan, Y. Li, J. Liu, L. Jiang, Face Liveness Detection from a Single Image with Sparse Low Rank Bilinear Discriminative Model, in: K. Daniilidis, P. Maragos, N. Paragios (Eds.), Comput. Vis. -- ECCV 2010, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010: pp. 504–517.

[33] Z. Zhang, J. Yan, S. Liu, Z. Lei, D. Yi, S. Li, A face antispoofing database with diverse attacks, 2012 5th IAPR Int. Conf. Biometrics. (2012) 26–31.

[34] S.Z.L. Shifeng Zhang, Xiaobo Wang, Ajian Liu, Chenxu Zhao, Jun Wan, Sergio Escalera, Hailin Shi, Zezheng Wang, A Dataset and Benchmark for Large-Scale Multi-Modal Face Anti-Spoofing, in: Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2019: pp. 919–928.

[35] A. Liu, Z. Tan, X. Li, J. Wan, S. Escalera, G. Guo, S.Z. Li, CASIA-SURF CeFA: A Benchmark for Multi-modal Cross-ethnicity Face Anti-spoofing, ArXiv. abs/2003.0 (2020).

[36] Y. Zhang, Z. Yin, Y. Li, G. Yin, J. Yan, J. Shao, Z. Liu, CelebA-Spoof: Large-Scale Face Anti-Spoofing Dataset with Rich Annotations, (2020).

[37] I. Pavlidis, P. Symosek, The imaging issue in an automatic face/disguise detection system, in: Proc. IEEE Work. Comput. Vis. Beyond Visible Spectr. Methods Appl. (Cat. No.PR00640), 2000: pp. 15–

24. https://doi.org/10.1109/CVBVS.2000.855246.

[38] R. TABULA, “Trusted biometrics under spoofing attacks,” Http://Www.Tabularasa-Euproject.Org/. (n.d.).

[39] S. Parveen, S.M.S. Ahmad, N.H. Abbas, W.A.W. Adnan, M. Hanafi, N. Naeem, Face Liveness Detection Using Dynamic Local Ternary Pattern (DLTP), Computers. 5 (2016). https://doi.org/10.3390/computers5020010.

[40] R. Raghavendra, R. Kunte, A Novel Feature Descriptor for Face Anti-Spoofing Using Texture Based Method, Cybern. Inf. Technol. 20 (2020) 159–176.

[41] J. Määttä, A. Hadid, M. Pietikäinen, Face spoofing detection from single images using micro-texture analysis, in: 2011 Int. Jt. Conf. Biometrics, 2011: pp. 1–7. https://doi.org/10.1109/IJCB.2011.6117510.

[42] Z. Boulkenafet, J. Komulainen, A. Hadid, Face Spoofing Detection Using Colour Texture Analysis, IEEE Trans. Inf. Forensics Secur. 11 (2016) 1818–1830. https://doi.org/10.1109/TIFS.2016.2555286.

[43] T. Pereira, J. Komulainen, A. Anjos, J. De Martino, A. Hadid, M. Pietikäinen, S. Marcel, Face liveness detection using dynamic texture, EURASIP J. Image Video Process. 2014 (2014) 2. https://doi.org/10.1186/1687-5281-2014-2.

[44] V. Mohanraj, S. Chakkaravarthy, S.. Gogul, I.. S. Kumar, V.. Kumar, V. Ranajit; Vaidehi, Hybrid feature descriptors to detect face spoof attacks, J. Intell. Fuzzy Syst. (2018) 1–9. https://doi.org/10.3233/JIFS-169436.

[45] D. Karmakar, P. Mukherjee, M. Datta, Spoofed Facial Presentation Attack Detection by Multivariate Gradient Descriptor in Micro-Expression Region, Pattern Recognit. Image Anal. 31 (2021) 285–294. https://doi.org/10.1134/S1054661821020097.

[46] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, in: 2005 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., 2005: pp. 886–893 vol. 1. https://doi.org/10.1109/CVPR.2005.177.

[47] F.A. Pujol, M.J. Pujol, C. Rizo-Maestre, M. Pujol, Entropy-Based Face Recognition and Spoof Detection for Security Applications, Sustainability. 12 (2020). https://doi.org/10.3390/su12010085.

[48] Z. Boulkenafet, J. Komulainen, A. Hadid, Face Anti-Spoofing using Speeded-Up Robust Features and Fisher Vector Encoding, IEEE Signal Process. Lett. PP (2016). https://doi.org/10.1109/LSP.2016.2630740.

[49] A. Benlamoudi, D. Samai, A. Ouafi, S.E. Bekhouche, A. Taleb-Ahmed, A. Hadid, Face spoofing detection using multi-level local phase quantization (ML-LPQ), in: 2015.

[50] K. Mohan, P. Chandrasekhar, S.A.K. Jilani, A combined HOG-LPQ with Fuz-SVM classifier for Object face Liveness Detection, in: 2017 Int. Conf. I-SMAC (IoT Soc. Mobile, Anal. Cloud), 2017: pp. 531–537. https://doi.org/10.1109/I-SMAC.2017.8058406.

[51] G. Thippeswamy, H. Vinutha, R. Dhanapal, A New Ensemble of Texture Descriptors Based On Local Appearance-Based Methods For Face Anti-Spoofing System, J. Crit. Rev. 7 (2020) 644–649. https://doi.org/10.31838/jcr.07.11.118.

[52] J. Li, Y. Wang, T. Tan, A.K. Jain, Live face detection based on the analysis of Fourier spectra, in: SPIE Def. + Commer. Sens., 2004.

[53] Z. Boulkenafet, J. Komulainen, A. Hadid, Face anti-spoofing based on color texture analysis, in: 2015 IEEE Int. Conf. Image Process., 2015: pp. 2636–2640. https://doi.org/10.1109/ICIP.2015.7351280.

[54] J. Yang, Z. Lei, S. Liao, S.Z. Li, Face liveness detection with component dependent descriptor, in: 2013 Int. Conf. Biometrics, 2013: pp. 1–6. https://doi.org/10.1109/ICB.2013.6612955.

[55] J. Yang, Z. Lei, S. Li, Learn Convolutional Neural Network for Face Anti-Spoofing, ArXiv. abs/1408.5 (2014).

[56] J. Galbally, S. Marcel, Face Anti-spoofing Based on General Image Quality Assessment, in: 2014 22nd Int. Conf. Pattern Recognit., IEEE Computer Society, Los Alamitos, CA, USA, 2014: pp. 1173–1178. https://doi.org/10.1109/ICPR.2014.211.

[57] J. Galbally, S. Marcel, J. Fierrez, Image Quality Assessment for Fake Biometric Detection: Application to Iris, Fingerprint, and Face Recognition, IEEE Trans. Image Process. 23 (2014) 710–724. https://doi.org/10.1109/TIP.2013.2292332.

[58] D. Anand, K. Gupta, Face Spoof Detection System Based on Genetic Algorithm and Artificial Intelligence Technique, Int. J. Res. Appl. Sci. Eng. Technol. 6 (2018). https://www.ijraset.com/fileserve.php?FID=18000.

[59] Y. Atoum, Y. Liu, A. Jourabloo, X. Liu, Face anti-spoofing using patch and depth-based CNNs, in: 2017 IEEE Int. Jt. Conf. Biometrics, 2017: pp. 319–328. https://doi.org/10.1109/BTAS.2017.8272713.

[60] Z. Boulkenafet, J. Komulainen, L. Li, X. Feng, A. Hadid, OULU-NPU: A Mobile Face Presentation Attack Database with Real-World Variations, 2017 12th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG 2017). (2017) 612–618.

[61] Y. Liu, A. Jourabloo, X. Liu, Learning Deep Models for Face Anti-Spoofing: Binary or Auxiliary Supervision, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2018.

[62] G. Fumera, G.L. Marcialis, B. Biggio, F. Roli, S.C. Schuckers, Multimodal anti-spoofing in biometric recognition systems, in: Handb. Biometric Anti-Spoofing, Springer, 2014: pp. 165–184.

[63] R.N. Rodrigues, L.L. Ling, V. Govindaraju, Robustness of multimodal biometric fusion methods against spoof attacks, J. Vis. Lang. Comput. 20 (2009) 169–179. https://doi.org/https://doi.org/10.1016/j.jvlc.2009.01.010.

[64] M.A. Turk, A. Pentland, Face recognition using eigenfaces, Proceedings. 1991 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (1991) 586–591.

[65] Nist, fingerprint software, National Institute of Standards and Technology (NIST), (n.d.).

[66] M.K.V. Sudhamani M J, Feature Level Fusion of Iris and Finger Vein Biometrics for Multimodal Biometric Authentication System, Int. J. Recent Technol. Eng. Volume-7 (n.d.).

[67] G. Pan, L. Sun, Z. Wu, Y. Wang, Monocular camera-based face liveness detection by combining eyeblink and scene context, Telecommun. Syst. 47 (2011) 215–225.

[68] G. Pan, L. Sun, Z. Wu, S. Lao, Eyeblink-based Anti-Spoofing in Face Recognition from a Generic Webcamera, 2007 IEEE 11th Int. Conf. Comput. Vis. (2007) 1–8.

[69] P. Zhang, F. Zou, Z. Wu, N. Dai, S. Mark, M. Fu, J. Zhao, K. Li, FeatherNets: Convolutional neural networks as light as feather for face anti-spoofing, in: IEEE/CVF Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work., IEEE/CVF, 2019: pp. 1574–1583. https://doi.org/10.1109/CVPRW.2019.00199.

[70] K. Patel, H. Han, A.K. Jain, Secure Face Unlock: Spoof Detection on Smartphones, IEEE Trans. Inf. Forensic Secur. XX (2016) XXX.

[71] and A.K.J. Keyurkumar Patel, Hu Han, Cross-database face antispoofing with robust feature representation, in: N Chinese Conf. Biometric Recognit., Springer, 2016: pp. 611–619.

[72] Y. Du, T. Qiao, M. Xu, N. Zheng, Towards Face Presentation Attack Detection Based on Residual Color Texture Representation, Secur. Commun. Networks. 2021 (2021) 6652727. https://doi.org/10.1155/2021/6652727.

[73] L. Li, X. Feng, Z. Boulkenafet, Z. Xia, M. Li, A. Hadid, An original face anti-spoofing approach using partial convolutional neural network, in: 2016 Sixth Int. Conf. Image Process. Theory, Tools Appl., 2016: pp. 1–6.

[74] Z. Wang, C. Zhao, Y. Qin, Q. Zhou, Z. Lei, Exploiting temporal and depth information for multi-frame face anti-spoofing, ArXiv. abs/1811.0 (2018).

[75] Dlib Python API Tutorials [Electronic resource] – Access mode: http://dlib.net/python/index.html, (n.d.). http://dlib.net/python/index.html.

[76] D.E. King, Dlib-ml: A Machine Learning Toolkit, J. Mach. Learn. Res. 10 (2009) 1755–1758.

[77] E. Osuna, R. Freund, F. Girosit, Training support vector machines: an application to face detection, in: Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., 1997: pp. 130–136.

[78] H. Yu, J. Yang, A direct LDA algorithm for high-dimensional data—with application to face recognition, Pattern Recognit. 34 (2001) 2067–2070.

[79] S. Bharadwaj, T.I. Dhamecha, M. Vatsa, R. Singh, Computationally efficient face spoofing detection with motion magnification, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Work., 2013: pp. 105–110.

[80] C. Priyanka, Sharma; Neha, Spoofing Face Detection using LBP Descriptor and KNN Classifier in Image Processing, Int. J. Recent Technol. Eng. Volume-8 (n.d.).

[81] K. Samrity, Saini; Kiranpreet, KNN Classification for the Face Spoof Detection, Int. J. Sci. Eng. Res. Volume 10 (n.d.) 1101–1106.

[82] Kanika kalihal ; Jaspreet Kaur, A Review on Different Face Spoof Detection Techniques in Biometric Systems, Int. J. Sci. Res. Eng. Trends. Volume 5 (n.d.).

[83] RecogTech, FAR and FRR: security level versus user convenience, Https://Www.Recogtech.Com/En/Knowledge-Base/Security-Level-versus-User-Convenience, (Retrieved on 31st/01/2022). (n.d.) (Retrieved on 31st/01/2022).

[84] S. Bengio, J. Mariéthoz, A Statistical Significance Test for Person Authentication, Speak. Lang. Recognit. Work. (2004).

[85] H. Raeisi Shahraki, S. Pourahmad, N. Zare, K Important Neighbors: A Novel Approach to Binary Classification in High Dimensional Data, Biomed Res. Int. 2017 (2017) 7560807. https://doi.org/10.1155/2017/7560807.