Ministry of Higher
Education and Scientific Research
University of Diyala
College of science
Department of Computer Science

Hybrid Techniques for Face Spoofing Recognition using Deep Learning and Eye Blinking

A Thesis Submitted to the Department of Computer Science/ College of Science/ University of Diyala
In Partial Fulfilment of the Requirements for the Degree of Master in Computer Science

By

Noor Al-Huda Taha Jabbar

Supervised By

Prof. Dr .Taha Mohammad Hasan

2021AC 1442AH

بسم اللّٰه الرحمن الرحيم

﴿ نَرْفَعُ دَرَجَاتٍ مَّن نَّشَاءُ وَفَوْقَ كُلِّ ذِي
عِلْمٍ عَلِيمٌ ﴾

صدق اللّٰه العظيم

سورة يوسف
الاية (76)

Acknowledgments

First of all, praise is to **Allah** the lord of the whole creation, on all the blessing was the help in achieving this research to its end.

I wish to express my thanks to my supervisor **Dr. Taha Mohammad Hassan** for supervising this research and for the generosity, patience, and continuous guidance throughout the work. It has been my good fortune to have advice and guidance from him. My thanks to the academic and administrative staff at the Department of the computer sciences\University of Diyala.

I would like to express my gratitude to **my father**, my **mother's soul**, there are no words enough to thank **my sisters, my brothers,** and my **dear husband** for being supportive and believing in me all the time, and for their encouragement during the period of my study and my **dear son**. Not to forget to mention Dear brother and colleague who helped me and stood by me from the beginning of my research journey to this day, all thanks and appreciation to **Mr. Mohammed Al-Saati,** this accomplishment would not have been possible without them.

Dedication

To...

To my mother's soul

My dear father and my brothers and sisters

My dear husband & my son

My Dear friend Mohammed Al-Saati

All our distinguished teachers those who paved the way for our science and knowledge

Noor Al-Huda Taha Jabbar

(Supervisor's Certification)

We certify that this research entitled "Face Spoofing Detection Method" was prepared by **Noor Al-Huda Taha Jabbar** under our supervision at the University of Diyala Faculty of Science Department of Computer Science, as partial fulfillment of the requirement needed to award the degree of Master of Science in Computer Science.

Signature:
Name: Prof. Dr.Taha Mohammad Hasan
Date:

Approved by the University of Diyala Faculty of Science Department of Computer Science.

Signature:
Assist. Prof. Dr. Bashar Talib Hamed
Date:

Head of Computer Science Department

(Linguistic Certification)

I certify that this research entitled "**Face Spoofing Detection Method**" was prepared by **Noor Al-Huda Taha Jabbar** and was reviewed linguistically. Its language was amended to meet the style of the English language.

Signature:

Name:

Date:

(Scientific Amendment)

I certify that the thesis entitled “**Face Spoofing Detection Method**” was prepared by **Noor Al-Huda Taha Jabbar** has been evaluated scientifically; therefore, it is suitable for debate by the examining committee.

Signature:

Name:

Date: /   / 2021

(Scientific Amendment)

I certify that the thesis entitled “**Face Spoofing Detection Method**” was prepared by **Noor Al-Huda Taha Jabbar** has been evaluated scientifically; therefore, it is suitable for debate by the examining committee.

Signature:

Name:

Date: / / 2021

Abstract

Face recognition systems are now being used in many applications such as border crossings, banks, and mobile payments. The wide-scale deployment of facial recognition systems has attracted intensive attention to the reliability of face biometrics against spoof attacks, where photos, videos, or a 3D mask of a genuine user's face can be used to gain illegitimate access to facilities or services.

The widespread deployment of face recognition-based biometric systems has made face Presentation Attack Detection (PAD) an increasingly critical issue. Most existing Face Anti-Spoofing (FAS) methods capture various cues (e.g., texture, depth, and reflection) to distinguish the live faces from the spoofing faces. All these cues are based on the discrepancy among physical materials (e.g., skin, glass, paper, and silicone).

In this thesis, an effective system against face spoofing attacks is proposed. In this system, two algorithms have been used which are (CNN, eye blinking).

The advent of deep learning algorithms has further increased the performance of face recognition systems, which in turn, has led to its increased usage in commercial applications and access control environments. A deep Convolution Neural Network (CNN) has been used in the proposed method. An eye blink detection method is added to robust the result.

The proposed method in this thesis is conducted using the Multispectral Latex Mask-based Video Face Presentation Attack (MLFP) dataset (contain two types of masks the dataset is divided into 80% for training and 20% for evaluation during the training process. The accuracy obtains in our methods is 100%, Precision is 100%, recall is 100% and F1-score is 100%.

List of Contents

List of Tables

List of Figures

List of Abbreviations

Abbreviations	Meaning
ACC	Accuracy
AI	Artificial Intelligence
ANN	Artificial neural network
AUC	Area Under Curve
BSIF	Binaries Statistical Image Features
CNN	Convolutional Neural Network
CPU	Central Processing Unit
DARPA	Defence Advanced Research Projects Agency
Dlib	Digital Library
DNN	Deep Neural Network
EAR	Eye Aspect Ratio
EER	Equal Error Rate
FAR	False Accept Rate
FAS	Face Anti-Spoofing
FRS	Face Recognition System
FRVT	Face Recognition Vendor Tests
GPU	Graphics Processing Unit
HTER	Half Total Error Rate
IDA	Image Distortion Analysis
LBP	Local Binary Pattern
LPQ	Local Phase Quantization
MLFP	Multispectral Latex Mask based Video Face Presentation Attack
NIST	National Institute of Standards and Technology
PAD	Presentation Attack Detection
ReLU	Rectified Linear Unit
ROI	Region of Interest
SAPLC	Spatial Aggregation of Pixel-level Local Classifiers
SGD	Stochastic Gradient Descent

SVM	Support-Vector Machines
HOG	Histogram of Oriented Gradients

Chapter One

General Introduction

CHAPTER ONE
GENERAL INTRODUCTION

1.1 Introduction

In recent years, face recognition has been widely used in various interactive and payment scenes due to its high accuracy and convenience. However, such a biometric system is vulnerable to presentation attacks (PAs). Typical examples of physical presentation attacks include print, video replay, 3D masks, and makeup. To detect such PAs and secure the face recognition system, face anti-spoofing (FAS) has attracted more attention from both academia and industry point of view [1].

In the past decade, several hand-crafted feature-based [2] and deep learning-based [3, 4] methods have been proposed for presentation attack detection (PAD). On one hand, the classical hand-crafted [5, 6] descriptors leverage local relationships among the neighbors as the discriminative features, which is robust for describing the detailed invariant information (e.g., color texture, moire's pattern, and noise artifacts) between the live and spoofing faces. Furthermore, due to the stacked convolution operations with nonlinear activation, the convolutional neural networks (CNN) hold strong representation abilities to distinguish the bona fide from PA. However, most existing CNN and handcrafted features are designed for universal image recognition tasks, which might not represent fine-grained spoofing patterns in FAS tasks.

In 2021, 4.66 billion people were active internet users, in the world out of the 7.83 billion global population encompassing 59.5 % of the global population [7], of which, 92.6% access the world wide web using mobile devices, with a very large amount of images and videos are uploaded to the

Internet each day. This includes millions of photos and over 400 hours of video content uploaded to Social media and YouTube every minute.

1.2 Related Works

The detection of manipulated videos is much harder than fake image detection due to the solid corruption and degradation of the data of the frame after video compression [8]. A great challenge for methods designed to detect fake videos because of their temporal attributes that are differed among frame-sets. Several related methods to the proposed method in this study have been reviewed in the literature.

H. Phuong Nguyen, et al. [9] in 2016, present a method for studying the problem of spoofing attack detection for facial recognition systems was proposed. There are real and fake faces in front of the sensing device (a camera in the phone) that has variations in the surface of micro-textures used to discriminate against pictures of face-spoofing. First, the wavelet-based de-noising method is used in estimates the noise of the image. The system exploits the distribution statistical behavior of local noise variances, which function differently between real and imaginary representations of faces. Testing the method was on two databases that were developed in their laboratory. Support Vector Machine (SVM) is used for the classification system. The results of the experiment indicate that the proposed method has an encouraging result for the NLV-3 features classification, can arrive at a high sensitivity of detection larger than 80% for just a very small rate of false-positive and less than 2.5%.

S.Yi Wang et al.[10] in 2017, proposed two novel features for face liveness detection systems to protect against printed photo attacks and replayed attacks for biometric authentication systems. The first feature obtains the texture difference between red and green channels of face images inspired by the observation that skin blood flow in the face has properties that enable distinction between real and spoofing face images. The second feature estimates the color distribution in the local regions of face images, instead of whole images, because image quality might be more discriminative in small areas of face images. These two features are concatenated together, along with a multi-scale local binary pattern feature, and a support vector machine classifier is trained to discriminate between living and spoofing face images. The experimental results showed that four public domain databases (the NUAA, CASIA, Idiap, and MSU databases) showed encouraging success for face spoof detection on images. While the proposed method did not reliably produce the best performance for each database, it performed excellently in terms of accuracy rate (96.69%), EER (7.01%).

A.Agarwal et.al. [11] In 2017, presents a unique multispectral video face database for face presentation attacks using latex and paper masks. The proposed Multispectral Latex Mask-based Video Face Presentation Attack (MLFP) database contains 1350 videos in visible, near-infrared, and thermal spectrums. Since the database consists of videos of subjects without any mask as well as wearing ten different masks, the effect of identity concealment is analyzed in each spectrum using face recognition algorithms. They also present the performance of existing presentation attack detection algorithms on the proposed MLFP database. It is observed that the thermal imaging spectrum is most effective in detecting face presentation attacks. Similarly, using RDWT+Haralick features in the thermal spectrum, indoor videos are detected with 89.9% accuracy as compared to 88.8% accuracy for outdoor videos

X.Sun, et al. [12] in 2018, presents a novel multimodal face anti-spoofing method, which makes full use of available information on RGB-D images and no manually chosen regions are needed. For every pair of RGB-D images, first of all, they calculate the correlation between color and depth images to detect multimodal properties; then, by analyzing the consistency of sub-regions extracted from the depth image, they were able to distinguish flat spoofing faces from genuine human beings. Both anti-spoofing features were fused to make final antispoofing decisions. Experiments on both self-collected and pubic 3DMAD datasets show that our proposed method is effective for intra-dataset and cross-dataset testing scenarios and that their method could deal with different presentation attacks carried by photos, tablet screens, and face masks. , the mean accuracy is 99.40% for proposed frame-level results during several cross-validation procedures.

Y. Liu, et al. [13] in 2018, introduce noise modeling and de-noising algorithms, they identify a new problem of face de-spoofing, for anti-spoofing: inversely decomposing a spoof face into a spoof noise and a real face, and then utilizing the spoof noise for classification. A CNN architecture with proper constraints and supervision is proposed to overcome the problem of having no ground truth for the decomposition. Evaluated their work on three face anti-spoofing databases, with print and replay attacks: Oulu-NPU, CASIA-MFSD, and Replay-Attack. Oulu NPU is a high-resolution database, considering many real-world variations. Oulu NPU also includes 4 testings. The results show promising improvements due to the spoof noise modeling. Moreover, the estimated spoof noise provides a visualization that helps to understand the added spoof noise by each spoof medium with achieving an accuracy of 82%.

H. Li, et al. [14] in 2018, proposed a method that how to train a network with limited face samples in a particular environment. In face anti-spoofing, it is practical to collect sufficient fully labeled training data samples captured under one invariant environment (e.g., limited types and modes of attacks). First, trained deep neural network based on reasonably sufficient labeled data in an attempt to “teach” a neural network for the application-specific domain for which training samples are scarce. Subsequently, trained sample pairs from both domains and formulate a novel optimization function by considering the cross-entropy loss, as well as a maximum mean discrepancy of features and paired sample similarity embedding for network distillation. Thus expect to capture spoofing-specific information and train a discriminative deep neural network on the application-specific domain**.** In the testing stage, the remaining samples in the corresponding database which was adopted as the application-specific domain are used for performance evaluation. This protocol also simulates the scenario for which it is impractical to collect samples from many different identities in the training stage for the application-specific domain. Since they focus on the problem of face anti-spoofing with limited face samples, the Equal Error Rate (EER) is adopted to measure the detection performance**.**

S. Priya, et al. [15] in 2019, Histogram of Directed Gradients (HOG), Local Binary Pattern (LBP), SIFT, VGG16, Shallow CNN, and Inception-ResnetV2, for face spoofing detection, are a comparative study of various local description and off-shelf deep networks for feature extraction. Besides, the function extracted via local descriptors and deep networks evaluates three classifiers trees for decision, (SVM), and Artificial Neural Network (ANN). A conveniently accessible YALE face dataset embodying actual and false facial photographs was used to perform the assessment. The dataset consists of 5121 real entries and 7508 images that are fake. The study findings showed that when graded with ANN, the optimum forecast accuracy of spoof and real were achieved with the outset of ResnetV2 features, and approximately 96.23 percent accuracy is achieved.

R. Ramachandra, et al. [16] in 2019, Proposed the method that presents an empirical study on both vulnerability analysis and presentation attack detection for commercial face recognition systems (FRS) using custom 3D silicone face masks corresponding to real subjects. Bonafide presentations for the corresponding subjects, as well as PAs, have been collected using three different smartphones, iPhone X, Samsung S7, and Samsung S8. This is the largest custom silicone mask dataset (compared to earlier works) collected so far. The experiments indicate that the two commercial FRS are vulnerable to PAs based on custom 3D silicone face masks, especially when the operating threshold corresponds to higher values of False Accept Rate (FAR). When the threshold is set at the lower values of FAR (e.g., FAR = 0.01%), both commercial FRS are not vulnerable to the custom silicone mask PAs. The state-of-the-art feature extraction algorithms evaluated in their study, all characterize image-texture information using: Local Binary Patterns (LBP) [28], Binaries Statistical Image Features (BSIF), Local Phase Quantization (LPQ), Image Distortion Analysis (IDA), and Color texture.

W. Sun, et al. [17] in 2020, Proposed the method that a state-of-the-art face spoofing detection method based on a depth-based Fully Convolution Network (FCN) was revisited. Proposed the Spatial Aggregation of Pixel-level Local Classifiers (SAPLC), which is composed of an FCN part and an aggregation part. The networks were then trained using stochastic gradient descent with a mini-batch size of 10 examples and a momentum of (0.9). The training is stopped after 100,000 iterations. Since all datasets contain segmented videos, each video has a single label. In the testing stage, the frame-level testing probabilities are temporally averaged to obtain the video-level decisions. The performance reported in their experiments is based on video-level decisions. The proposed SAPLC was compared with representative deep networks and some state-of-the-art methods in experiments on the CASIA-FASD, Replay-Attack, OULU-NPU, and SiW datasets. The performance of the SAPLC on the cross-dataset evaluation is generally lower than that on the intra-dataset evaluation. In the future, domain adaptation methods can be introduced into the SAPLC framework such that the cross-domain performance can be improved. Besides, the partial attack in the SiW-M dataset is another novel and tough problem that needs to be handled. The proposed SAPLC achieves an overall AUC of 92.58%, ranks first in the comparison of four methods.

K. Kotwal and S. Marcel. [18] In 2020, proposed a CNN-based face PAD method to detect 3D mask attacks in the NIR channel. This method employs a patch-pooling mechanism to learn textural cues from the final Conv layer of CNN. They have also demonstrated that a CNN, pertained for FR using visual spectrum data, can be directly used to compute the patch-pooled feature descriptor. The proposed PAD method has been tested on two publicly available datasets that consist of masks made of paper, latex, and silicone. Excellent results, on both datasets, indicate that the patch pooling mechanism is well-suited for discriminating mask-based PAs in the NIR channel. The accuracy of the worst trial is above 97%.

Shan Jia et al. [19], 2020, From the perspective of fine-grained classification, address the problem of detecting these realistic 3D face presentation attacks and suggest a novel Anti-spoofing process. This system, which uses factorized bilinear coding of multiple color channels (MC FBC) to learn subtle fine-grained differences between real and fake images, aims to learn subtle fine-grained differences between real and fake images. They developed a principled approach to 3D face spoofing detection by extracting discriminative and fusing complementary information from RGB and YCbCr spaces. To aid the study of 3D face presentation attack detection, a large-scale wax figure face database (WFFD) of both images and videos has been collected as super-realistic attacks. Extensive research results show that the proposed method outperforms the competition on both WFFD and other face spoofing databases in a variety of intra-database and inter-database testing scenarios. The accuracy obtained from this method is 94.74.all these related works illustrate in Table (1.1).

Table 1.1: Related works of some researchers for face spoofing detection

Study	Year	Dataset	Type of attack	Classifier	Feature extraction	Accuracy.
H. Phuong Nguyen, et al. [9]	2016	1317 images 673 real face 644 fake face	images	SVM	exploits the statistic behavior of the distribution of noise's local variances	80%
S.Yi Wang et al. [10]	2017	NUAA, CASIA, Idiap, and MSU	printed photo attacks and replayed attacks	SVM	skin blood and multi-scale local binary pattern	96.69%
Akshay Agarwal et.al [11]	2017	MLFP	3D Mask 2D Mask	binary	RDWT+Haralic k featur	88.8% in outdoor 89.9% in indoor

Xudong Sun [12]	2018	3DMAD	Photo attack	SVM	Four different features (color diversity, blurriness, specular reflection, chromatic moment)	99.40%
Y. Liu, et al. [13]	2018	Oulu-NPU, CASIA-MFSD, and Replay-Attack	Print attack Replay attack	Noise modeling	CNN	82%.
H. Li, et al. [14]	2018	Idiap REPLAY-ATTACK CASIA	Image video	Binary classifier	Hand-crafted Features Deep Learning-based Feature	Not mention
Sandan Priya, et al [15]	2019	YALE face database	Photo attack (Printed photo)	-Decision Tree. - ANN. - SVM	Local Binary Pattern (LBP), Shallow CNN, SIFT, Histogram of Directed Gradients (HOG), VGG16, and Inception-ResnetV2.	96.23%
R. Ramachandra† et al. [16]	2019	six custom silicone masks	3D mask attack (high quality)	convolutional neural network (CNN)	Eye-brows and facial make-up.	93.57 %
Wenyun Sun et al. [17]	2020	OULU-NPU and SiW	Replay attack	Fully Convolutional Network (FCN)	Depth-based Fully Convolutional Network (FCN)	92.58%

Yunxiao Qin et al. [85]	2020	CASIA	warped Photo	Meta-Learning	Hand-crafted feature	99.1%
Ketan Kotwal and Sebastien Marcel [18]	2020	WMCA and MLFP datasets	masks made of paper, latex, and silicone	Binary classification	patch-pooled feature descriptor	97%
Shan Jia et al. [19]	2020	wax figure face database (WFFD)	Video attack And photo attack	fine-grained	(RGB vs. YCbCr) extracted via CNN model	94.74%

1.3 Statement of the Problem

Face recognition systems are now being used in many applications such as border crossings, banks, and mobile payments. The wide-scale deployment of facial recognition systems has attracted intensive attention to the reliability of face biometrics against spoof attacks, where a photo, a video, or a 3D mask of a genuine user's face can be used to gain illegitimate access to facilities or services.

The difficulty is that they are vulnerable to identity theft assaults. To trick the hackers of recognition systems, employ numerous approaches, such as using face pictures or videos of people in the database of the system. A new kind of attack has arisen using 3D face masks. This form of attack is quite effective. Another problematic feature of counter spoofing is the difficulties of collecting enough samples for a specific application and representative attacks, as might a large percentage of hackers who utilize 3D masses.

1.4 The Aim of Thesis

We aim to develop a strong face spoofing system to discern real and fraudulent videos through the usage of the CNN and Eyeblink methods. CNN consists of 18 layers and uses a Video Presentation Attack database with a multispectral Latex mask (MLFP). The use of 2D paper masks and 3D latex masks in the visible spectrum replicates facial attacks in this database. Serious actions must be taken and methods to detect face spoofing videos must be developed to restrict the production of this phenomenon.

1.5 Thesis organization

The thesis is segmented into five chapters; a brief description of their contents is given below:

Chapter One: This chapter introduces an overview of the work and related works.

Chapter Two: This chapter introduced methods and descriptions for the theoretical background and techniques that are used in this thesis.

Chapter Three: This chapter describes the proposed method with its design and implementation and the execution of the stages of the proposed method.

Chapter Four: This chapter presents the tests and the results of the proposed method.

Chapter Five: This chapter offers conclusions and systems for future work.