Visible/Infrared face spoofing detection using texture descriptors

Shaimaa Mohamed[1,*], *Amr* Ghoneim[1], and *Aliaa* Youssif [2]

[1]Computer Science Department, Faculty of Computers & Information Helwan University, 11795 Ain Helwan, Cairo, Egypt
[2] College of Computing and Information Technology Arab Academy for Science, Technology and Maritime Transport AASTMT Smart Village, Giza, Egypt

Abstract. With extensive applications of face recognition technologies, face anti-spoofing played an important role and has drawn a great attention in the security systems. This study represents a multi-spectral face anti-spoofing method working with both visible (VIS) and near-infrared (NIR) spectra imaging. Spectral imaging is the capture of images in multiple bands. Since these attacks are carried out at the sensor, operating in the visible range, a sensor operating in another band can give more cues regarding the artifact or disguise used to carry out the attack. Our experimental results of public datasets proved that the proposed algorithms gain promising results for different testing scenarios and that our methods can deal with different illuminations and both photo and screen spoofing.

Keywords: Biometric Security; Face Spoofing Detection; Multispectral Biometric, Spoofing attacks, Texture Analysis.

1 Introduction

Face recognition systems are increasingly being deployed in a diversity of scenarios and applications. [1-5] Due to this widespread use, they have to withstand a high variety of attacks. Biometric signatures are increasingly employed since they have been shown to be relatively more secure when compared to other traditional security approaches. The face is specifically a unique and popular biometric choice on account of its simplicity to be employed; as it doesn't require intricate devices nor invasive procedures / interactions, but simply a camera. Nonetheless, recently, hackers found a way to attack other security tools like password hackers, they don't miss face spoofing. They found their way with accessing a system with fake face photos or videos or most recent with 3D face mask. Therefore, nowadays the question is not any more whether or not biometrics can be copied or forged, but rather to what extent systems are robust to these attacks and if they incorporate the necessary countermeasures to detect them. Spoofing attacks for NIR face recognition systems have not received as much attention, but recent spoofing attempts indicate on their weakness too. [6] Taking into consideration that the driving force of attackers is not how hard systems are to spoof, but how valuable are the resources they protect, it is not pessimistic to expect more and more advanced spoofing attacks in near future. The focus of this paper is on such problem and its scope has been limited to the development of more accurate method to detect disguised or fake face images of printed photos or screens.

2 Related Works

2.1 Spoofing detection in Visual Spectrum

Most of the existing spoofing detection techniques are based on images taken in visual spectrum (VIS spectrum). The existing systems make use of images captured in the visible range (350-740 ηm) of the electromagnetic spectrum [9]. However, these systems suffer from performance degradation due to different illumination conditions. Infrared is another method that can be used by existing systems. The advantage it has over visible is due to its invariance to ambient lighting [10].
VIS spectrum may be roughly divided into three main categories: texture based methods, motion based methods, and methods using other spoofing clues
Texture features are extracted from face images under the assumption that printed faces produce certain texture patterns that do not exist in real ones. Texture is probably the strongest evidence of spoofing most of the works.

2.2 Spoofing detection in Infrared Spectrum

Most researchers suggested Infrared (IR) imagery and Thermal Infrared in the micron wavelength as an alternative way for face identification and recognition. They investigated a lot of modalities, Single and multi-modality. We will discuss some of them next coming sections. But first let's discover the difference between

* Corresponding author: shaimaa_muhammad@fci.helwan.edu.eg

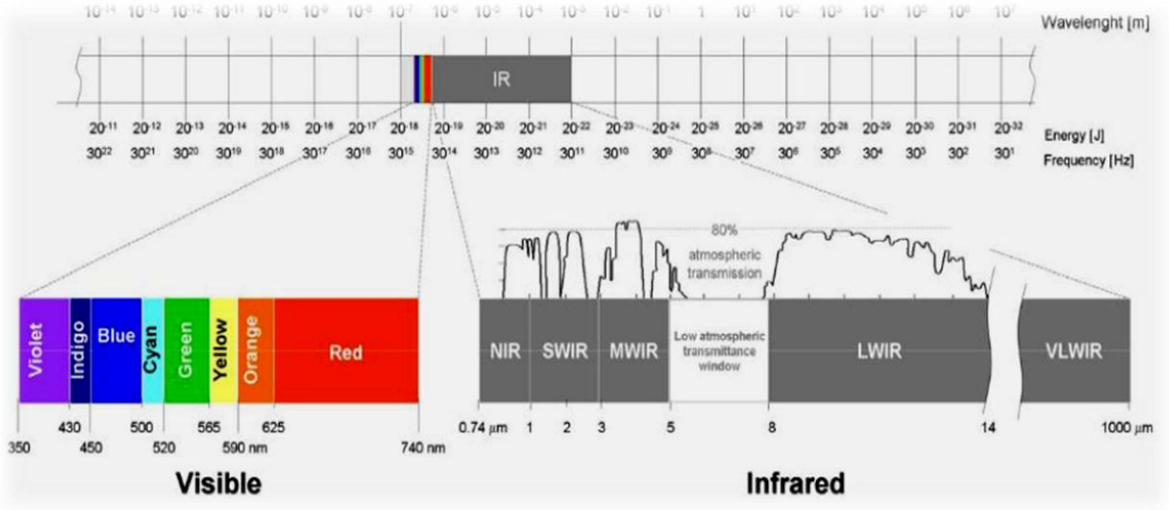

Fig.1, Electromagnetic spectrum: visible and infrared bands and its wavelength

infrared imagery and Thermal Infrared imagery emitted from cameras. Identification of visible light faces gives weak results due to the lack of illumination. For night time surveillance, thermal infrared images is commonly used because of the thermal IR sensors radiations emit from the human face. Thermal face recognition advantages in detecting masked faces or even in low light conditions. Any pattern of a face image consists of eyebrow, eyes, nose, ears, cheeks, mouth, lips, teeth, skin, and chin. The face of any human contains other features like expressions, beard, mustache etc. The biometric identification uses the patterns of specific organs like eyes. These patterns are derived from blood vessels under the skin. As the veins and the tissues of the face structures are unique, therefore the IR images are also unique.

3 Proposed Methods

Two sets of images are required at least, to make a biometric system working. One set consists of images of individuals that are known to the system. A feature extraction process helps in reducing the information in these images to a set of features, which are then saved as templates, in the gallery set (training set of images). Gallery set images are obtained in the registration phase, where the trait of an individual is enrolled in the system, which is captured by a sensor. The other set of images, consists of images of individuals that are presented to the biometric system for personal identication. These images are also converted to respective templates by feature extraction. This set is called the probe set (set of testing images). Classification methods are used to learn from the extracted features in the gallery set and with the help of those, classify the probe image. As stated earlier, spectral imaging of face involves using different spectral bands of the electromagnetic spectrum to collect information in those specific bands, as in Fig1 and Fig.2. The term spectral imaging, includes, both multispectral imaging and hyper spectral imaging. The former has the ability to image a scene in tens of spectral bands whereas the latter has the ability to image a scene in hundreds of spectral bands [7, 8]. The results of a spectral sensor depend especially on the optical properties of the objects in the scene and the illumination geometry. These sensor results have to be converted to reflectance units. The result is a reflectance spectrum which can also be referred to as spectral signature of the objects in the image.

Band	Wavelength
Visible	350-740 ηm
Near InfraRed (NIR)	740-1000 ηm
Short-Wave InfraRed (SWIR)	1-3 μm
Mid-Wave InfraRed (MWIR)	3 - 5 μm
Long-Wave InfraRed (LWIR)	8-14 μm

Fig.2, Different bands and their wavelength

3.1 Illumination variations av nd feature extraction

VIS images are completely sensitive to illumination conditions, so finding a correlation directly between VIS and NIR images could be a tough idea, especially when we need to make our anti-spoofing methods compatible with different lighting conditions. We divide face images horizontally and vertically into l parts equally, obtaining m = l × l patches for each image. Inspired by one recent research, Local Binary Pattern (LBP) [11] based method still plays an important role in spoofing detection and achieves relatively good results.2 It is noted that LBP with (8,1) neighbourhood, denoted as LBP (8,1), satisfies the requirements above. For every image pixel in each patch, LBP (8, 1) descriptor chooses 8 sampling points on a circle of radius 1 clockwise. If sampling point pixel's value is greater than the centre's value, forming a bit '1'; otherwise, forming '0'. This gives an 8-digit binary number corresponding to the centre pixel. Finally, by computing the frequency of every number obtained in the image patch, a histogram could be counted as texture descriptors. This simple local descriptor will achieve promising results in our proposed pipeline as shown in Fig.3. Edge information can also be considered for texture representation. In order to describe edges, Difference Of Gaussians (DoG) [12, 13] are used to remove lighting variations while preserving high frequency components.

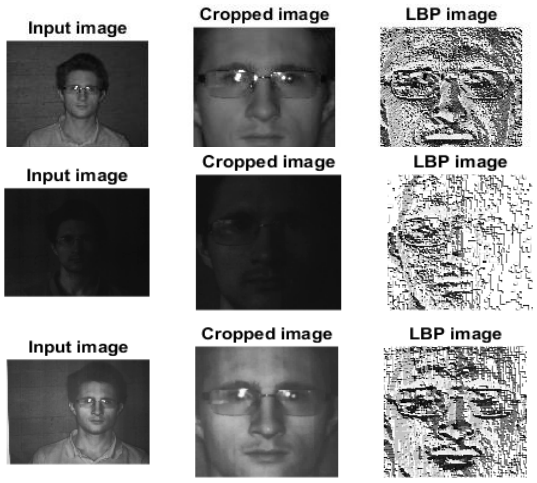

Fig.3, The first column represents different input images: Real and Normalized, Real, and Fake faces. The Second Column represents the face images after pre-processing. The third column represents the result of LBP on each face image from Msspoof database.

3.2 Convolutional Neural Network

Following a recent trend, the anti-spoofing community started experimenting with approaches where the anti-spoofing features are automatically learned directly from the data. This is in contrast to the previously discussed approaches, where the features are inspired by some particular characteristics that can be observed as common either for real accesses or for some types of spoofing attacks. It is argued, however, that the features engineered in this way are not suitable for different kinds of spoofing attacks. Material composition, texture, illumination and shape of 2D surface from presented faces can generate a variety of features. These heterogeneous features can be self-learned using convolutional neural networks (CNN) [18] and were used in our investigation. CNN is a type of neural network, which mainly uses convolution pattern (multiple layers) to connect the neurons. This network can be viewed as a combination of linear and nonlinear image processing operations. CNN network is consists of different kinds of layers, such as convolution layer, pooling layers, ReLU layers, normalization layer, fully connected layers, and loss layers. The network will learn the parameters automatically through forward propagation and backward propagation. A good architecture is a combination of different layers, which can help to extract different unique features when different inputs are provided. Visible and NIR images can be used to enhance the performance presentation attacks detection and was used in this work. Firstly, we use network consists of 11-layer VGG with 2 derived VGG-11 networks have been trained, which can offer the best performance from trained network by individually using NIR and visible images. We input the face images to the network and let the network learn the features from these face images.

4 Experimental Setup and Results

4.1 Database

Msspoof Dataset: Multispectral-Spoof dataset is provided by Chingovska [14], and it contains real access face images and spoofing attack ones for a total of 21 identities. All the images are recorded in VIS and NIR spectra with different controlled ambient illumination, and all the images are gray-level ones. The dataset is made up of several predefined subsets: train, dev and test, containing 9, 6, and 6 non-overlap subjects respectively; according to the dataset regulation, the enrol subset would not be used in our experiment since we are going to employ a binary classifier that distinguishes real accesses from spoofing attacks. The details of this dataset and the number of genuine and spoofing images are shown in Table. 1

Table 1, Msspoof Database

Spectrum	Access	Train	Dev	Test	ALL
NIR	Genuine	265	179	179	623
	Photo Spoof	638	433	432	1503
VIS	Genuine	270	180	180	630
	Photo Spoof	641	434	432	1507

4.2 Results

The experiments on Msspoof Database show promising results shown in Table 2, Fig. 4 and Fig. 5 shows the Area Under Curve for using both proposed algorithms on our selected Msspoof Database. As can be observed our CNN networks gives an overall high classification rate as compared with the chosen local binary pattern method. Since in the previous work [15], a CNN is used for feature extraction and SVM for classification, our CNN architecture directly classify an input face image into either genuine face or fake face.

Table 2, Results for Msspoof database, using DoG-LBP + SVM in terms of percentage of accuracy

Image Type	AUROC	ACC%
VIS	0.974	94.49
NIR	0.973	**95.24**
VIS & NIR	0.726	91.95

Table 3, Results for Msspoof database, using CNN in terms of percentage of accuracy

Image Type	AUROC	ACC%
VIS	0.975	94.89
NIR	0.988	95.11
VIS & NIR	0.995	**96.78**

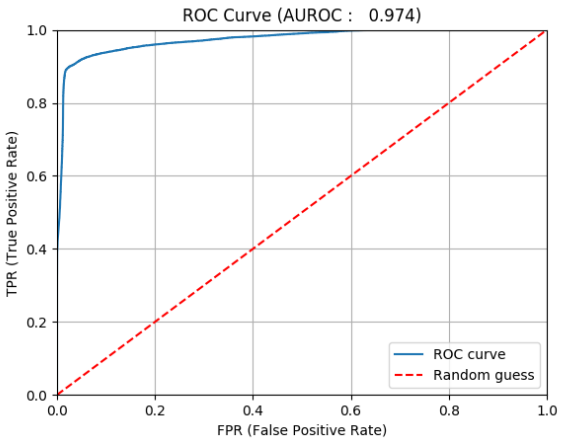

Fig. 4, ROC Curve for using LBP on Msspoof Database Visible Images

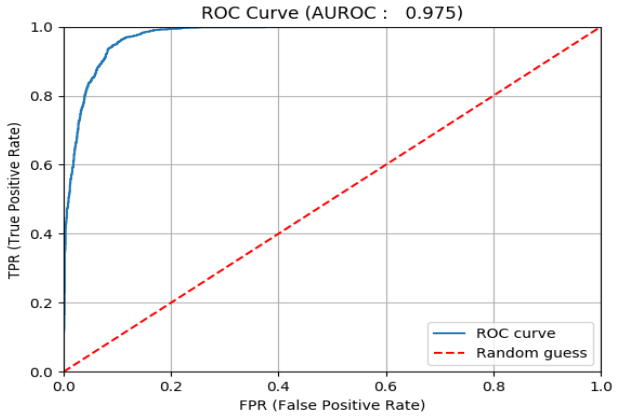

Fig. 7, ROC Curve for using CNN on Msspoof Database Visible Images

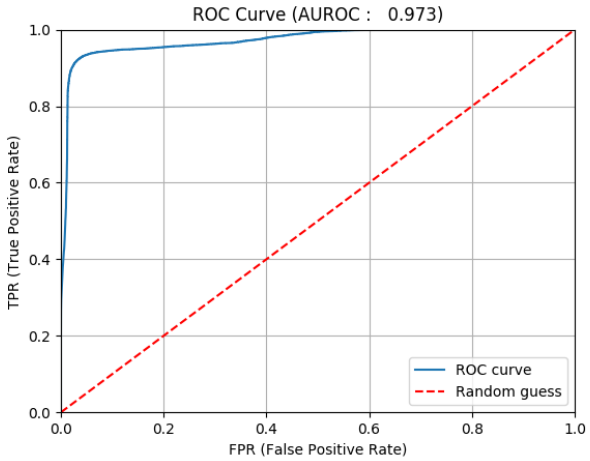

Fig. 5, ROC Curve for using LBP on Msspoof Database Near Infrared Images

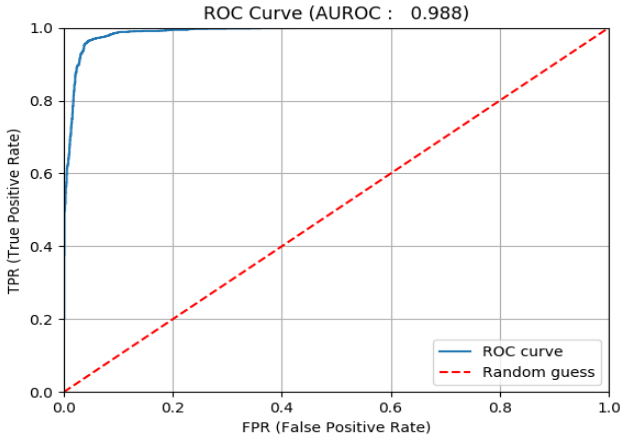

Fig. 8, ROC Curve for using CNN on Msspoof Database Near Infrared Images

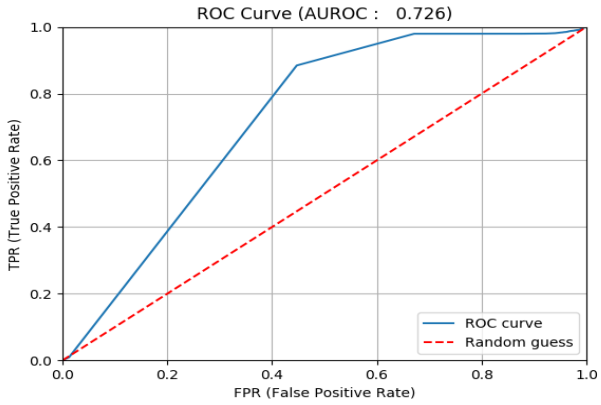

Fig. 6, ROC Curve for using LBP on Msspoof Database Both VIS & NIR face images

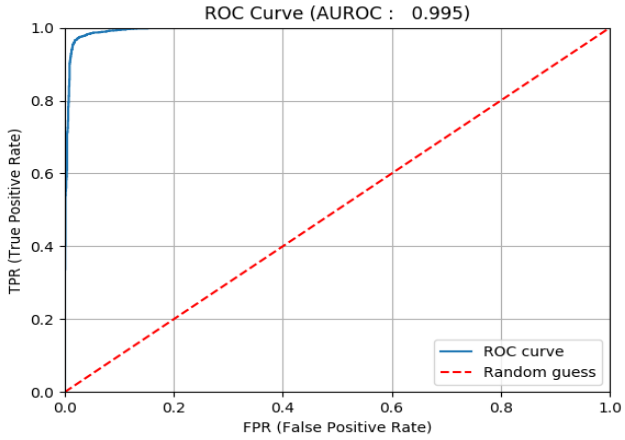

Fig. 9, ROC Curve for using CNN on Msspoof Database Both VIS & NIR face images

5 Conclusion and Future work

In our paper, we proposed different approaches on VIS & NIR face images using Msspoof database. The work provided a comparison and a combination of 2D face spoofing detection from simultaneously acquired near-infrared and visible illumination images.
Additionally, an accurate training strategy has been proposed for training CNN architecture when training data is limited.
Future work will investigate large set of training and testing datasets.

References

[1] A. Pentland and T. Choudhury, "Face Recognition for Smart Environments," IEEE Computer, 33(2):50-55, (2000).

[2] P. Phillips et al., "The Feret Database and Evaluation Procedure for Face Recognition Algorithms," Image and Vision Computing, May (1998), pp. 295-306.

[3] L. Wiskott et al., "Face Recognition by Elastic Bunch Graph Matching," Trans. IEEE Pattern Analysis and Machine Intelligence, 19(7):775-779, (1997).

[4] B. Moghaddam and A. Pentland, "Probabilistic Visual Recognition for Object Recognition," Trans. IEEE Pattern Analysis and Machine Intelligence, 19(7):696-710, (1997).

[5] P. Penev and J. Atick, "Local Feature Analysis: A General Statistical Theory for Object Representation, Network: Computation in Neural Systems, Mar. (1996), pp. 477-500.

[6] Yi, D., Lei, Z., Zhang, Z., Li, S.: Face anti-spoofing: Multi-spectral approach. In: Marcel, S.,Nixon, M.S., Li, S.Z. (eds.) Handbook of Biometric Anti-Spoofing, Advances in Computer Vision and Pattern Recognition, pp. 83–102. Springer, London (2014)

[7] D. W. Allen, An overview of spectral imaging of human skin toward face recognition, in: Face Recognition Across the Imaging Spectrum, Springer, (2016), pp. 1-19.

[8] M. Nischan, R. Joseph, J. Libby, J. Kerekes, Active spectral imaging, Lincoln Laboratory Journal **14** (2003) 131-144.

[9] B. Thirimachos, R. Arun, C. Cunjian, H. Lawrence, A study on using mid-wave infrared images for face recognition (2012). doi:10.1117/12.918899.
URL https://doi.org/10.1117/12.918899

[10] M. S. Sarfraz, R. Stiefelhagen, Deep perceptual mapping for thermal to visible face recognition, (2015).

[11] T. Ojala, M. Pietikinen, D. Harwood. A comparative study of texture measures with classification based on feature distributions, Pattern Recognition, 29 (1) (1996) 51-59, Elsevier.

[12] B. Peixoto, C.Michelassi, A. Rocha, "Face liveness detection under bad illumination conditions", in Proc. ICIP, (2011)

[13] Z. Zhang, J. Yan, S. Liu, Z. Lei, D. Yi, S. Li, A face antispoofing data set with diverse attacks, in Proc. ICB, (2012), pp. 26–31.

[14] I. Chingovska, N. Erdogmus, A. Anjos and S. Marcel, Face Recognition Systems Under Spoofing Attacks, Face Recognition Across the Imaging Spectrum (Springer, 2015), pp.165-194.

[15] Y. Abbas, L. Man Po, M. Liu, Deep Learning for Face Anti-Spoofing: An End-to-End Approach, IEEE, (2017)

[16] L. Huang, C. Lu, "MULTISPECTRAL FACE SPOOFING DETECTION USING VIS-NIR IMAGING CORRELATION" International Journal of Wavelets, Multi resolution and Information Processing, (2018)