Turkish Journal of Computer and Mathematics Education Vol.12 No.13 (2021), 4363-4373

Research Article

Face Spoofing Detection Using Deep CNN

Noor Al-Huda Taha [a], Taha Mohammed Hassan [b] , Mohammed Akram Younis [c]

[a,b] University Of Diyala, College of Science, Department of Computer science
[c] University Of Mosul, College of Science, Department of Computer science

Email [1] : scicompms16@uodiyala.edu.iq , Email [2] dr.tahamh@sciences.uodiyala.edu.iq , Email [3]: mohammed.akram@uomosul.edu.iq

Abstract: 3D mask face spoofing attack is an important challenge in recent years and draws further study. Because of the number of deficiencies and small differences in the database, however, a few methods can be proposed to aim at it meanwhile, most developed databases are focused on countering various types of threats and neglect environmental developments in implementations in the real world. The 3D mask against spoofing is used in this paper to simulate the real-world scenario with other options. The database used in the proposed method includes 10 different subject masks (7 subject 3D latex masks and 2 subjects for 2D paper masks and 1 for half mask from below the eye is using to testing the result). Therefore, the total size is 440 videos 400 is fake videos, and 40 is a real video. The directions for future study are shown in the benchmarking experiments. We intend to release the database platform to evaluate various methods this system has been used for a deep Convolution neural network. The result is robust for an eye-blink recognition technique. There are three basic steps in the proposed system: Firstly, video pre-processing, facial recognition, and finally, the output step whether the video is true or falsified. The method used is stronger than most techniques. The suggested approach in this study was used for the MLFP dataset and was very reliable and accurate as a result of the whole experiment and the accuracy obtained is (99.88).

Keywords: face spoofing attack, eye blinking, anti-spoofing, face detection, face recognition.

Introduction

Face recognition (FR) systems have obtained an excellent accuracy but still have no important limitation in terms of their reliability and security in detecting a PA (3d mask and Print attack) and Face anti-spoofing. Presentation attacks can be classified as 2D or 3D, based on the kind of tool used to make an attack. To reliably work the FR method, the identification of all types of PA is essential. The majority of 2D attacks (print, visual display) can be accurately identified by RGB data alone or by the use of a supplementary data acquisition medium such as infrared, thermal, or depth. However, it is a difficult job for RGB (visible spectrum) to recognize 3D masks. In this paper, the issue of detected 3D mask attacks from a real person (mask made of soft silicone) has been discussed. Where the term 3D mask applies to a wide range of quality and material masks. Fig.1 presents examples of various attack types.

Figure1. Some types of attacks.

For 3D face attack detection a series of studies (PAD) has been made easier because of the vulnerability of current face recognition technologies to reliable face appearance attacks [1]. Existing approaches attempted to examine the disparity in the use of deep [3,4] or motion [5,6] of real face skin, multi-spectral imaging, [8,9, 2],and 3D mask fake face materials. In several current 3D face spoofing databases [7,8,9,10] they have achieved promising detection results. However, as facial masks are mostly made of paper, latex, or silicone, these 3D face masking databases have limited database dimensions (mostly of less than 30 subjects), low authenticity (some based on a two-dimensional plane or no custom masking's), and low variety of subject and registration processes, which could hamper the evolution of efficient and realistic PAD schemes [9, 11]. Several studies [1, 12, 13] have shown that a range of 3D PAD approaches are suffered from weakening efficiency in databases where the face of 3D spoofing is more diverse and realistic. Print attacks[26,27], video replay attacks [28], and 3D mask attack [29] are among the most popular spoof-attack. There are some distinctions between real and fake faces, mostly expressed in information about image textures, motion and depth. We may use these differences to distinguish the real and fake faces by designing a range of face anti-Spoofing methods. The study into the detection of face spoofing has been quickly improved in recent years, with several useful study outcomes. This work focuses on two matters: the deep learning method, and the methodology of detection of the blinking eye, and the development of antiracial development As follows, the rest of the paper is structured.

Section II gives a brief historical of related work in different methods. The definition of methodology in Section III and discusses the basic concepts of CNNs that are important for a deep neural network method to be understood and discuss the eye blinking detection. The key CNN-learning and training ideas are outlined in this section. In Section V, explain the dataset used in this work for training and testing and discuss the result obtain in this method, Section VI contains conclusions and section VII contain the references.

Related work

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 [30]. 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 work in this study have been reviewed in the literature:

In 2017 proposed two novel features for face liveness detection systems to protect against printed photo attacks and replayed attacks for biometric authentication systems by S.Yi Wang et al.[31], texture difference between the red and green face channels inspired by the observation of the properties of skin blood flow in the face that differentiates between real and fake face spoofing.The second feature measures the color distribution in face images local regions instead of whole images, because in small areas of face image, image quality may be more discriminating. 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, proposed A new LBP network for face spoofing detection end-to-end learning.

In 2018 proposed Network offers three distinct advantages through the integration by L.L. et al [32]. of fixed sparse binary filters and derivable statistical histograms functions :

(i) Reduce network parameters drastically in convergence and fully connected layers; (ii) Train parameters directly with limited data effectively; (iii) complete the fundamental process of LBP extraction.

The stochastic gradient descent (SGD) algorithm is used during the training stage[33] to learn the network parameters, The method presented with an empirical study both for the identification of vulnerabilities and for the detection of presentation attacks on commercial face reconnaissance systems (FRS) using a custom silicone mask for real people by R. Ramachandra, et al. [34], 2019, Bonafide Presentations were made on three different, Samsung S7 , iPhone X, Smartphones and Samsung S8, for the corresponding subjects as well as for Pas, in 2020 Propose the way of revising the state-of-the-art approach of face spoofing based on the Fully Convolution Network (FCN) by W. Sun, et al. [35], Various supervisory schemes are thoroughly explored, including the global and local label supervision.A general theoretical analysis and related simulation is offered to show that the local label supervision, for local tasks with poor training samples such as facial spoofing, is more suited than global label supervisory.Based on this research, an FCN part and aggregation part are proposed to be the Pixel level Local Classification (SAPLC). The networks are then trained using stochastic gradient descent with a mini-batch size of 10 examples and a momentum of (0.9). The proposed SAPLC is 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 proposed SAPLC achieves an overall AUC of 92.58%, ranks first in the comparison of four methods.

1. Methodology

The pipeline for spoof detection in this paper is first discussed in this section. Then descriptions of the descriptors of the local images and the deep architectures used in this work. The figure shows the methodology for the detection of face spoof.

Deep CNNs are an excellent choice for the extraction of complex characteristics for a variety of applications. The typical CNN consists of a series of convolution layers (Conv) and followed by one or more max-pooling layers and followed by layers fully connected (FC). In the input image, the convolution layers learn local regions' characteristics; in which form similar characteristics are captured using the FC layers [14]. The FC layer is the output of a CNN spatially codes.

More discriminative for the detection of mask attacks would be the best representation of information from the final of CNN, with less spatial details. This paper proposes a new neural network to extract the features from the input image to achieve this representation if the input is real or fake and supports the result of this method by using eye blinking detection. In [15], it is seen that CNN is a good process to extract or practice PAD functionality. For this purpose, we also see CNN as the backbone of the proposed method.

The Stages of the proposed method:

a) Face Detection and Extraction

Face detection is the first and necessary step for the detection of facial spoof and is used to recognize the faces of photographs. In this paper, a deep learning face detector is used based on the Dlib [16]. The Dlib depends on the Histogram of oriented gradients (HOG).

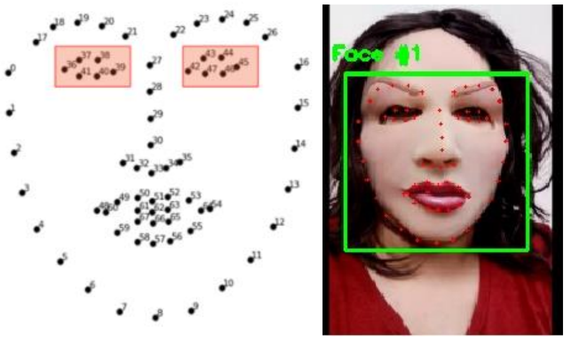

Figure 2. A simple process by Dlib’s shape predictor.

Histogram of oriented gradients is an image processing feature descriptor and computer vision that is for object recognition purposes. The technology includes the occurrence of gradient orientation in image fields [22]. This approach is similar to the method used in histograms of edge orientation, invariant descriptors, and sort contexts, but varies because it is measured on a dense grid of cells with uniform separation and uses a normalization of local contrast to increase accuracy as shown in Fig. 3.

The HOG descriptor is calculated as follows from the source image:

1- Standardization of color and gamma correction is done.

2- Vertically and horizontally, the gradient values are calculated.
3- The image is subdivided into an uniform cell grid.

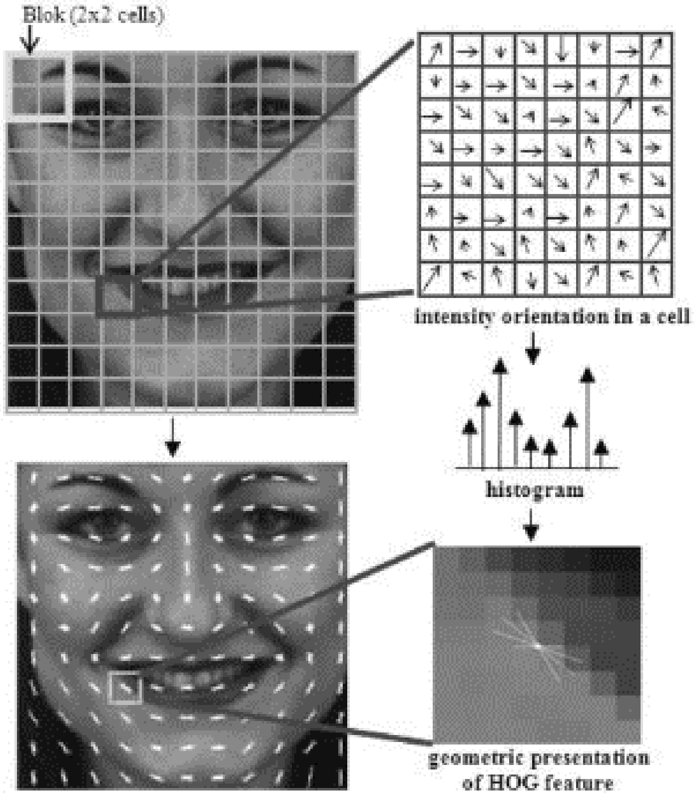

Figure 3**.** Examples of a histogram of oriented gradients

The base unit of the HOG descriptor is a pixel-sized block, rectangular area.A block consists of cells assigned a histogram of gradient directions (inclination in relation to the horizontal).The HOG descriptor is the vector for normalized histogram components of all block areas.As a general rule, blocks overlap, meaning that more than one last descriptor includes each cell.

b) **Deep CNN Features Extraction**

Convolutional Neural Network (CNN) is trained in classification. CNN consists of several hidden layers including Convolution Layer, Activation (Relu), Pooling layer, fully connected layers between the input and the final output layer. The neurons in the secret layer learn the characteristics of the input images and predict the true and flawed groups. The output layer forecasts the input image and provides each class with the percentage of an input image. The highest likelihood class is the end classify of a real or fake image. The image from the video as a real or spoof is classified after the model is trained with the real or spoof data set. The CNN architecture consists of eighteen layers (convolution layer, max-pooling layer, and fully connected layer). Table 1 shows the initialization of the trained CNN used in this model and used the early stopping to avoid excessive training of an iterative learner and to determine the number of epochs shown in Fig (4). These methods update the learner to ensure that each iteration is more suited to the training data.

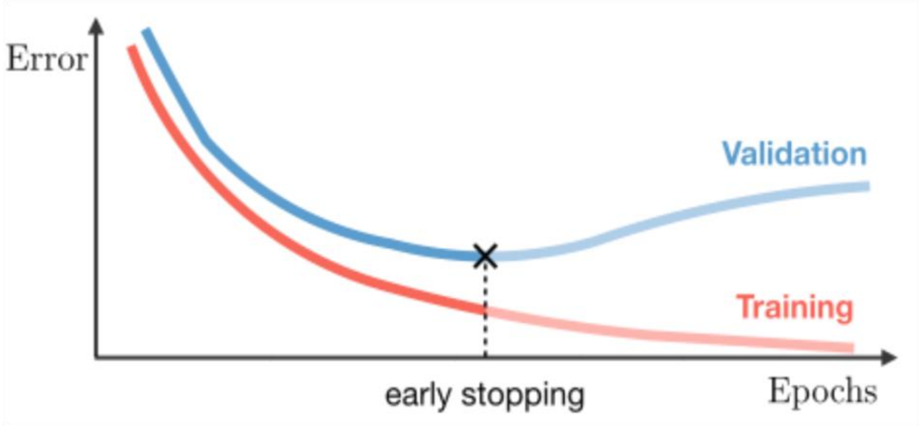

Figure 4. Using early stopping in Tensorflow

Table 1. CNN configuration.

BASE_NETWORK :	**19 layer**
PRETRAINED_MODELS :	**False**
IMG_SIZE :	**[224, 224, 3]**
TRAIN :	**{'LEARNING_RATE': 1e-3, 'NUM_EPOCH': 20}**

Below it is the fundamental layers of a CNN:
The layer of input: The input typically consists of a multidimensional array of data, which transmit data to the network [6].
Convolutional layers or stages: It's CNN's principal building block. The main aim of the convolution is to extract different features from the input. Krig [17] outlines the existence of such layers as a filter or learning kernel, which is designed to extract local inputs and which is used by each kernel to compute a function map or kernel map.
In the first convolution layer, meaningful characteristics such as borders, angles, textures, and lines are extracted. The next convolution layer extracts higher characteristics, but at the final convolution layer, the highest features are extracted [18]. The kernel size corresponds to the size of the filter along with the function map when the sliding process of the filter is stride. It governs how the filter gathers around the characteristic map. The filter then slides one unit across the various input layers of the function map [19]. Another important function of CNNs is a padding that allows data input to be extended [21]. the output size and kernel width W need to be controlled independently, zero padding input is used [20].
Pooling layer
The pooling layer is a nonlinear operation that usually reflects the downsampling process to minimize the feature dimensions and complexity of the whole network.
Among the several options for bundling layers, such as average pooling (AVE), max pooling (MAX), L2 standard pooling, etc.
In this work, maximum pooling was chosen since the non-linearity of the whole Network is not only increased, but the computer complexity is also reduced.
Max pooling may be illustrate in Fig (5).

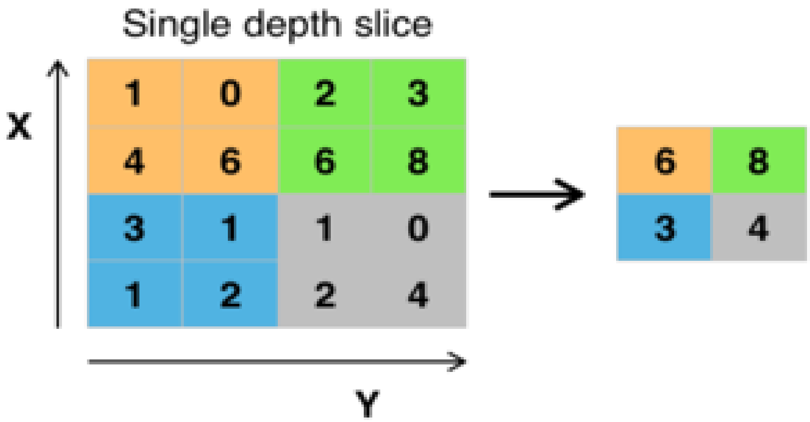

Figure 5. example of max pooling layer

Fully connected layer

The final classification is achieved using fully connected layers after several convolutionary and max pooling layers.Neurons in a fully connected layer have links to all previous layer activations as shown in regular artificial (non-convolutionary) neural networks.

Their activation may thus be calculated using a matrix multiplication and then bias compensation as an affine transformation (vector addition of a learned or fixed bias term).

D) Eye blinks Detection Stage

Specifically for the area of interactions between impaired individuals and computers[36], somnolence [37] and cognition load [38], blink detection technology has been recently employed.Thus, blink periods, blinking counts and frequency analysis of the eye are a significant source of information about a subject's condition and helps to assess how external variables affect change in emotional conditions.The blink of the eye is characterized as a swifter eyelid closure and re-open, usually from 100 to 400 ms [39].

Previous eye-blink detection systems evaluate ocular status as open or closed [40] or eye-closure pathways [41].

For each image, other approaches employ template matching, which lists open and/or closed eye templates and computes a standard cross-reference coefficient for each eye region[42].In order to begin with a blink detection, the eye must first be detected.

This is done by employing a pretrained model that offers us sixty-eight face points.Then map the landmarks on the face found.The process is in three sections implemented.Load the pretrained model's contents into an object first.Next, use this object to map the signs on the face seen in the frame and then, as shown in the figure(6), to extract the Cartesian coordinates of these landmarks .

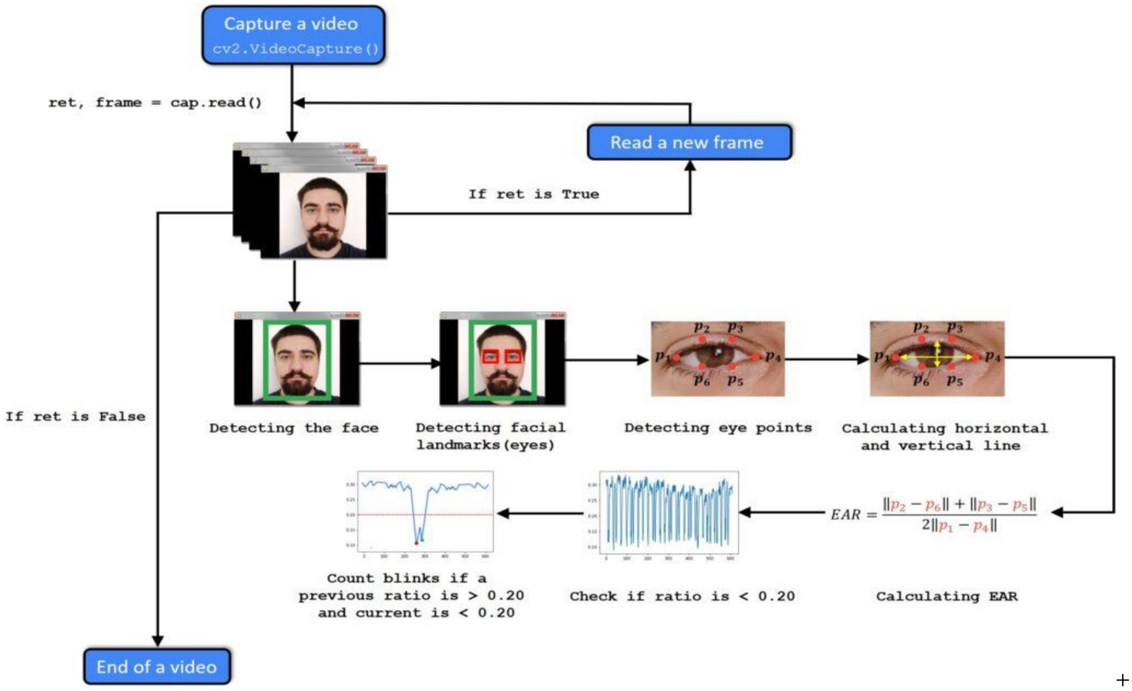

Figure 6. Block diagram of eye blinking detection method.

Experimental

a) Dataset

The MLFP dataset is made up of (440) videos for (10) subjects with and without a face mask. Videos were recorded in unconstrained environments at various places (indoors and outdoors). Of the 440 database videos, 400 are masked and the remaining 40 have no mask. The dataset comprises 10 individuals between the age of 23 - 38 years (4 females and 6 males). The minimum video length is 10 seconds in the database. There are over 46,000 frames in the database. The characteristics of the MLFP database proposed are summarized in Table (2). Two kinds of masks were used in the MLFP database (3D mask attack, 2D print attack):

• **Masks 3D Latex**: Seven latex masks are being used in the series of databases. The masks are soft and therefore adapt to the face of the subject. These masks enable the mouth and the face to move in a life-like manner as shown in fig. (7).

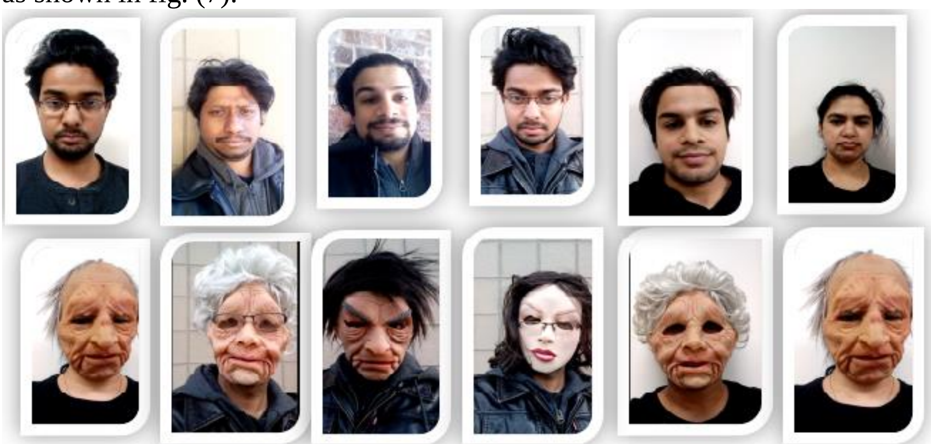

Figure 7. Different Environments of Some samples of 3D mask and real in MLFP dataset.

Table 2**.** The Multispectral Latex Mask-based Face Presentation Attack (MLFP) database's characteristics.

Spectrum	**Visible**
Masks Type	Latex (7) and paper (3)
Videos Count	440
Videos Type	Real (40) and fake (400)
Subjects Count	(4 Female, 6 Males)
Environmental Variations	Outdoor/ Indoor with fixed/random background
Video Duration	Min. 10s and Max.15s

• **2D Masks of paper**: Three masks with eye cut-outs are used. This is done on high-quality card paper with high-resolution photographs as shown in fig. (8). As opposed to 3D latex masks, 2D paper masks are smoother and show lighter. On the other hand, 3D latex masks, like wrinkles and some facial hair, have textured features. In both latex masks and 2D paper masks, gender and age differences appear.

Figure 8. Different Environments of Some samples of 2D paper mask in MLFP dataset

b) **Result Evaluation**

The test and result of the whole experiment done on the MLFP dataset, each test was carried on the real video and the fake (3D Silicon mask, 2D print paper mask) video. The experiment result gave good and accurate results with a few nuances.

The experimental results are classified into two main groups depending on some criteria as described in each group:

Group (1): This group of results holds the real-video and fake-video data that have a different possibility of a fake-video and real video test. The videos which were less than 0.5 possibilities the error of detecting because the less the number of frames and low resolution of videos maybe lead to false detect, and equal to or larger than 0.5 possibilities for correct detect. In each testing case, the count of frames (real and fake frames), the eye blink classification, and CNN classification are shown.

Group 2: The experimental results of some 3d mask videos were selected from the YouTube website. These videos consist of 20 different masks (5 females and 15 males) between the ages of 23- 38 years. The minimum video duration in the database is 6 seconds. By examining all of these videos, excellent results were obtained almost 100%.

c) **Proposed Algorithm vs. Related Works**

The comparison was made for the proposed system of face spoofing detection. Table **3** shows a comparison with the related work that using the same dataset [24,25].

Table 3. Comparison of Classification Accuracy with Earlier Studies

Researchers	Methodology	Accuracy
Akshay Agarwal [24]	RDWT+Haralick features	Indoor videos 89.9% outdoor videos 88.8%
Ketan Kotwal and Sebastien Marcel [25]	Deep CNN	97%
The proposed algorithm	Deep CNN and Eye blinking	99.88%

d) **Results**

This phase is used to evaluate the performance of the face spoofing detection classifying an approach. In this system, 20% remaining of the MLFP dataset is used for the testing phase. At then, testing the 80 fake videos and 8 real videos of test data and calculated the percentage of accuracy, Precession, Recall as shown in table 4 below.

Table 4**.** The results achieved in the proposed method

Metrics	CNN of MLFP dataset	Some video from YouTube
Accuracy	98.86%	100%
Recall	98.75%	100%
Precision	100%	100%

Conclusion

The system for detecting 3D mask attacks in the visible spectrum was proposed by CNN-based presentation attack detection (PAD) and eye blinking detection. This method employs a max-pooling process to learn CNN's final Conv layer textural proof. Two public tests of the proposed PAD approach were carried out Data sets available consisting of paper masks, and silicone. And some videos from YouTube consisting of 3D latex masks. Excellent results show that the maximum pooling phase can be discriminated against with Visible spectrum mask-dependent PA and achieve excellent accuracy in both the dataset and YouTube videos.

References

[1] S. Jia, G. Guo, and Z. Xu, “A survey on 3d mask presentation attack detection and countermeasures,” Pattern Recognition, vol. 98, p. 107032, 2020.

[2] Y. Wang, S. Chen, W. Li, D. Huang, and Y. Wang, “Face anti-spoofing to 3d masks by combining texture and geometry features,” in Chinese Conference on Biometric Recognition. Springer, 2018, pp. 399–408.

[3] R. Shao, X. Lan, and P. C. Yuen, “Joint discriminative learning of deep dynamic textures for 3d mask face anti-spoofing,” IEEE Transactions on Information Forensics and Security, vol. 14, no. 4, pp. 923–938, 2018.

[4] Y. Qin, C. Zhao, X. Zhu, Z. Wang, Z. Yu, T. Fu, F. Zhou, J. Shi, and Z. Lei, “Learning meta-model for zero- and few-shot face anti-spoofing,” in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, no. 07, 2020, pp. 11 916–11 923.

[5] S. Liu, P. C. Yuen, S. Zhang, and G. Zhao, “3d mask face anti-spoofing with remote photo plethysmography,” in European Conference on Computer Vision. Springer, 2016, pp. 85–100.

[6] B. Lin, X. Li, Z. Yu, and G. Zhao, “Face liveness detection by rppg features and contextual patch-based CNN,” in Proceedings of the 2019 3rd International Conference on Biometric Engineering and Applications, 2019, pp. 61–68.

[7] N. Erdogmus and S. Marcel, “Spoofing in 2d face recognition with 3d masks and anti-spoofing with Kinect,” in Biometrics: Theory, Applications, and Systems (BTAS), 2013 IEEE Sixth International Conference on. IEEE, 2013, pp. 1–6.

[8] S. Liu, B. Yang, P. C. Yuen, and G. Zhao, “A 3d mask face anti-spoofing database with real-world variations,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2016, pp. 100–106.

[9] I. Manjani, S. Tariyal, M. Vatsa, R. Singh, and A. Majumdar, "Detecting silicone mask-based presentation attack via deep dictionary learning," IEEE Transactions on Information Forensics and Security, vol. 12, no. 7, pp. 1713–1723, 2017.

[10] A. George, Z. Mostaani, D. Geissenbuhler, O. Nikisins, A. Anjos, and S. Marcel, "Biometric face presentation attack detection with multichannel convolutional neural network," IEEE Transactions on Information Forensics and Security, 2019

[11] A. Agarwal, D. Yadav, N. Kohli, R. Singh, M. Vatsa, and A. Noore, "Face presentation attack with latex masks in multispectral videos," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2017, pp. 81–89.

[12] R. Shao, X. Lan, and P. C. Yuen, "Deep convolutional dynamic texture learning with adaptive channel-discriminability for 3d mask face antispoofing," in Biometrics (IJCB), 2017 IEEE International Joint Conference on. IEEE, 2017, pp. 748–755.

[13] I. Manjani, S. Tariyal, M. Vatsa, R. Singh, and A. Majumdar, "Detecting silicone mask-based presentation attack via deep dictionary learning," IEEE Transactions on Information Forensics and Security, vol. 12, no. 7, pp. 1713–1723, 2017.

[14] M. Campos, S. Maji, and A. Vedaldi, "Deep filter banks for texture recognition and segmentation," in Proceedings of Conference on Computer Vision and Pattern Recognition, June 2015, pp. 3828–3836.

[15] K. Kotwal, S. Bhattacharjee, and S. Marcel, "Multispectral deep embeddings as a countermeasure to custom silicone mask presentation attacks," IEEE Transactions on Biometrics, Behavior, and Identity Science, vol. 1, no. 4, pp. 238–251, Oct 2019.

[16] https://github.com/davisking/dlib-models.

[17] S. Krig, Computer Vision Metrics. Survey, Taxonomy, and Analysis of Computer Vision, Visual Neuroscience, and Deep Learning. Springer, p. 637, 2016. https://doi.org/10.1007/978-3-319-33762-3.

[18] J. S. Ren, W. Wang, J. Wang, and S. Liao, "An Unsupervised Feature Learning Approach to Improve Automatic Incident Detection," 2012 15th International IEEE Conference on Intelligent Transportation Systems, Sep. 2012. https://doi.org/10.1109/itsc.2012.6338621 [35] C. Affonso, A. D. Rossi, F. H. A. Vieira, and A. C

[19] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning (Adaptive Competition and Machine Learning). The MIT Press, p. 779, 2016.

[20] M. A. Younus and T. M. Hasan, "Abbreviated View of Deepfake Videos Detection Techniques," in 2020 6th International Engineering Conference "Sustainable Technology and Development"(IEC), 2020, pp. 115–120.

[21] I. Namatēvs, "Deep Convolutional Neural Networks: Structure, Feature Extraction, and Training.," Inf. Technol. Manag. Sci., vol. 20, no. 1, 2017.

[22] M. A. Younus and T. M. Hasan, "Effective and fast deepfake detection method based on haar wavelet transform," in 2020 International Conference on Computer Science and Software Engineering (CSASE), 2020, pp. 186–190.

[23] S. Jia, X. Li, C. Hu, G. Guo, and Z. Xu, "3D Face Anti-Spoofing with Factorized Bilinear Coding," IEEE Trans. Circuits Syst. Video Technol., 2020.

[24] A. Agarwal, D. Yadav, N. Kohli, R. Singh, M. Vatsa, and A. Noore, "Face presentation attack with latex masks in multispectral videos," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2017, pp. 81–89.

[25] K. Kotwal and S. Marcel, "CNN Patch Pooling for Detecting 3D Mask Presentation Attacks in NIR," in 2020 IEEE International Conference on Image Processing (ICIP), 2020, pp. 1336–1340.

[26] Z. Boulkenafet, J. Komulainen and A. Hadid, "Face antispoofing based on color texture analysis," in Proc. of IEEE Int. Conf. on Image Processing, Piscataway, NJ: IEEE Press, pp. 2636–2640, 2015.

[27] Z. Boulkenafet, J. Komulainen and A. Hadid, "Face spoofing detection using colour texture analysis," IEEE Trans on Information Forensics and Security, vol. 11, no. 8, pp. 1818–1830, 2016.

[28] Xiaobai Li, J. Komulainen and Guoying Zhao, "Generalized face anti-spoofing by detecting pulse from face videos," in Proc. of IEEE23rd Int. Conf. on Pattern Recognition, Piscataway, NJ: IEEE Press, pp. 4239–4244, 2016.

[29] S. Q. Liu, X. Y. Lan and P. C. Yuen, "Remote photoplethysmography correspondence feature for 3D mask face presentation attack detection," in Proc. of the European Conf. on Computer Vision, Cham: Springer, pp. 558– 573, 2018

[30] D. Afshar, V. Nozick, J. Yamagishi, and I. Echizen, "Mesonet: a compact facial video forgery detection network," in 2018 IEEE International Workshop on Information Forensics and Security (WIFS), 2018, pp. 1-7: IEEE

[31] S.-Y. Wang, S.-H. Yang, Y.-P. Chen, and J.-W. Huang, "Face liveness detection based on skin blood flow analysis," Symmetry (Basel)., vol. 9, no. 12, p. 305, 2017.

[32]L. Li, X. Feng, Z. Xia, X. Jiang, and A. Hadid, "Face spoofing detection with local binary pattern network," J. Vis. Commun. Image Represent., vol. 54, pp. 182–192, 2018.

[33] Z. Yan, L. Hu, H. Chen, and F. Lu. Computer vision syndrome: A widely spreading but largely unknown epidemic among computer users. Computers in Human Behaviour, (24):2026–2042, 2008. 1

[34] R. Ramachandra et al., "Custom silicone face masks: Vulnerability of commercial face recognition systems & presentation attack detection," in 2019 7th International Workshop on Biometrics and Forensics (IWBF), 2019, pp. 1–6.

[35] W. Sun, Y. Song, C. Chen, J. Huang, and A. C. Kot, "Face spoofing detection based on local ternary label supervision in fully convolutional networks," IEEE Trans. Inf. Forensics Secur., vol. 15, pp. 3181–3196, 2020.

[36] Królak, A.; Strumiłło, P. Eye-blink detection system for human–computer interaction. Univers. Access Inf. Soc. 2012, 11, 409–419. [CrossRef]

[37] Rahman, A.; Sirshar, M.; Khan, A. Real time drowsiness detection using eye blink monitoring . In Proceedings of the 2015 National Software Engineering Conference (NSEC), Rawalpindi, Pakistan, 17 December 2015; pp. 1–7.

[38] Wilson, G.F. An analysis of mental workload in pilots during flight using multiple psychophysiological measures. Int. J. Aviat. Psychol. 2002, 12, 3–18. [CrossRef]

[39] Stern, J.A.; Walrath, L.C.; Goldstein, R. The Endogenous Eyeblink. Psychophysiology 1984, 21, 22–33. [CrossRef]

[40] Lee, W.O.; Lee, E.C.; Park, K.R. Blink detection robust to various facial poses. J. Neurosci. Methods 2010, 193, 356–372. [CrossRef]

[41] Garcia, I.; Bronte, S.; Bergasa, L.M.; Almazán, J.; Yebes, J. Vision-based drowsiness detector for real driving conditions. In Proceedings of the 2012 IEEE Intelligent Vehicles Symposium (IV), Madrid, Spain, 3–7 June 2012; pp. 618–623.

[42] Udayashankar, A.; Kowshik, A.R.; Chandramouli, S.; Prashanth, H. Assistance for the paralyzed using eye blink detection. In Proceedings of the 2012 Fourth International Conference on Digital Home (ICDH), Guangzhou, China, 23–25 November 2012; pp. 104–108.

[43] Yogesh Hole et al 2019 J. Phys.: Conf. Ser. 1362 012121