International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Focus Challenge Based Presentation Attack Detection in Face Authentication Systems Using Generative Adversarial Network

Rohini B. R.[1], Yogish H. K.[2]

Submitted: 06/05/2023 **Revised**: 15/07/2023 **Accepted**: 08/08/2023

Abstract: Presentation based deceiving of biometric recognition especially face recognition systems have become very common. Detecting such attacks is very important to ensure attack resilience of authentication systems built on biometric recognition. This work proposes a focus challenge based presentation attack detection for Face authentication systems using Generative Adversarial Network (GAN). The difference between the focus varying GAN generated images to the real images is compared in terms of their Deep learning group signature to detect fakes. The focus challenge is very shift giving almost no chance for presentation attack to deceive it. The proposed GAN based presentation attack detection system is very resilient to presentation attacks with noninvasive detection process. Testing under different environmental conditions, the proposed solution is found to have less than 1.74 Attack Presentation Classification Error Rate (APCER) which is atleast 1.3 times less than existing works.

Keywords: *GAN, Face Recognition Systems, Deep Learning*

1. Introduction

Biometric based authentication systems have become common due to various inherent merits like uniqueness, convenience, difficult to steal etc. Compared to token based systems like cards, knowledge based systems like passwords etc, they provide strong security. But with recent presentation attacks and ability to generate more realistic fake biometric samples, this strong security claims are challenged. Compared to iris, fingerprint etc, creating fake biometric samples for face is easier. The fake samples can be created or captured from person (with/without his knowledge) and can be used to break the biometric based authentication systems used in applications like smart homes, protected environment, banking systems etc. In spite of various challenges like illumination, pose etc, face based biometric authentication is preferred due to it simplicity and ability to authenticate in a non-invasive manner. From the face image acquired by cameras, features are extracted after various pre-processing. The features are matched against the features in database to recognize the person. Presented attacks exploit the loop holes in image acquisition process and present printed photographs, previously captured images or videos to deceive face authentication process. In absence of presented attack detection, the authentication gets passed and any attacker can bypass the frontline security provided by the systems. Various solutions have been proposed to detect presentation attacks. These solutions are based on analysis of image features and comparison of features between real and attack samples. Statistical or machine learning based comparison of features is done to detect presentation attacks. But the techniques are static can be deceived with recent deep learning based fake sample generation techniques. The effectiveness of presentation attack detection can be improved by adding dynamism to process of face acquisition. The dynamism must be a random challenge response mechanism, so that faking cannot be done shiftily in pace with challenge and the presentation attack gets detected with higher probability.

Motivated by this observation, this work proposes a GAN based presentation attack detection system for face recognition systems. The challenges are presented in terms of varied focus and face image is acquired. The expected response in presence of focus challenge is generated using GAN and the deep learning group signatures of GAN generated images is compared against the deep learning group signature of captured face images. Higher difference in signature match is an indication of presentation attack. Following are the novel contributions of this paper work.

(i) A novel focus based challenge response mechanism to detect presentation attacks is proposed. The focus challenge is random and the response matching is shift, making it difficult for attackers to predict the focus challenge pattern and provide response. This makes the face based authentication system fully resilient against attacks. (ii) A novel GAN based group signature to match expected challenge response to acquired face images. GAN signature matching is quite fast compared to pixel based matching. The matching error is very less resulting in lower false positive rate in the proposed solution. The rest of the paper is organized as follows, In

[1] *Dept. of AI&DS, Global Academy of Technology, INDIA*
ORCID ID : 0000-0003-2410-6558
rohini.br@gmail.com
[2]*Dept. of IS Ramaiah Institute of Technology, INDIA*
ORCID ID : 0000-0002-1814-3523
** Corresponding Author Email: yogishhk@gmail.com*

section 2, related works on presentation attack detection systems for face recognition are discussed and the research gaps are detailed. In section 3, the proposed focus challenge based presentation attack detection with GAN is discussed. In section 4, the performance results of proposed solution and comparison with state of works are presented. In section 5, the conclusion and future scope of work are presented.

2. Related Work

Mohamed et al. [1] built a model that detects live and non-live faces with Convolutional Neural Networks based on the Celebi-Spoof dataset [2]. According to the results of the tests against CelebA-Spoof, the method achieved 87% accuracy. Real samples were not tested with this method. Using this approach, liveliness is detected by analyzing differences between facial components, but this method can easily be deceived if a realistic image is presented.

Kim et al. [3] detected the presence of liveliness using the effect of defocus. Real and fake faces are compared on the basis of their depth information. Two cameras with different focuses are used to acquire the face images. By analyzing the images, we can determine the focus, power, gradient, and orientation histograms. In order to check for consistency, two face images are compared for differences in features. Differences between real and fake faces are high for real faces, but they are minor for fake faces. Observing the difference between the projection of noise and ear in different views can result in the approach failing if the ears are hidden.

Souza et al. [4] have developed LBPnet, which is a variation of a typically convolutional network that incorporates LBP into the first layer of the convolution, thereby making the convolution work on pixels' LBP values instead of pixels' original values. It is possible to detect artificially spoofed images with the help of the resulting deep learning features. Unless illumination varies greatly, real images cannot be analyzed by this approach.

Parveen et al. [5] used skin texture analysis to detect face liveliness using a Dynamic Local Ternary Pattern (DLTP). It is best to use DLTP features to extract the textural properties of facial skin. Spoofed attacks are detected by comparing real and fake features. By presenting faces with varying illuminations, the method can be easily deceived.

Akhtar et al. [6] identified discriminative patches that correlate well with the detection of spoofing attacks in face images. When local intensity in-homogeneity is observed, discriminative patches can be found. The detection of discriminative patches of face images is based on statistical inference. Detecting spoofing involves comparing discriminative patches between stored and acquired faces. When presenting the same stored images, the method could be easily deceived.

Wen et al. [7] developed a distortion-based scheme for the detection of face spoofs. The images are acquired from multiple devices to detect spoofing, and features like reflections, blurs, and chromatic moments are compared between the images. Across different acquisition devices, these features differ. Spoofing can be easily detected when a fake sample is displayed from a different device. Hackers can create fake samples with information to deceive attack detection when the device and acquisition parameters are hacked.

Tirunagari et al. [8] investigated how liveliness could be detected using a classification pipeline that combines dynamic mode decomposition, local binary patterns, and support vector machines. Various dynamics are used to detect liveliness in the face, such as eye blinks, lip movements, and face changes. The solution cannot dedetect replay attacks in videos with long durations that include face movements.

Boulkenafet et al. [9] detected facial spoofing based on textural colour distortions. Different colour textures on luminance and chrominance are compared to distinguish between genuine and fake samples. However, replay attacks cannot be defended using this method.

Zhou et al. [10] devised a mechanism for extracting multiscale features from colour images with a powerful representation capability. It is possible to resist illumination variations and noise by using local directional number patterns with derivative Gaussian masks. Local directional number patterns accommodate spatial-temporal variations. However, the method fails to detect replay attacks when artificially created faces are used.

Li et al. [11] tested face recognition systems for antispoofing using pulse detection. An ROI was selected near the cheeks and chin. As the cardiac pulse changes, the colour values of the pixels in the ROI region change. There is a difference between the power spectral density of real and fake faces in these regions, which can be used to detect liveliness. A print attack works best with this method, but a replay attack does not.

Hasan et al. [12] illustrated how dynamic texture features could work in conjunction with contrast features to detect spoofs. The features are extracted by modifying DoG filtering and using local binary patterns variance. Support vector machines are then used to detect spoofed photos based on the features. It is possible to deceive the proposed work using replay attacks.

Cai et al. [13] proposed two-stream hierarchical fusion networks to detect spoofing. A recent deep learning model is used to extract meta patterns from images. Spoofing is detected by fusing Meta patterns with RGB image features. In spite of the fact that the method can detect photo spoofing attacks, it cannot detect replay attacks.

Zheng et al. [14] used multiscale, and depth features to detect spoofing attacks. Feature extraction from images is proposed using two-stream spatial-temporal networks. Based on these features, a fully connected network layer classifies the samples as spoofs or genuine. Spoof photo attacks work best, but replay attacks fail when the method is used.

Song et al. [15] captured the face image with a binocular camera. We extract new depth and texture features from it. In order to distinguish real faces from spoofed ones, the features are classified using an SVM classifier.

Cai et al. [16] developed a deep learning method for extracting local features that discriminate. Different local regions are represented by convolutional neural networks and recurrent neural networks. Spoofed faces are detected by fusing regional features with global features. Artificial faces created with Deepfakes are the best candidates for the method, but replay attacks cannot be detected.

Yu et al. [17] redefined the problem of face spoof detection as an issue of material recognition since spoofs are created with materials like skin, glass, paper and silicone. Bilateral convolutional neural networks are used to extract intrinsic material-based patterns. Spoofs are classified based on the depth of information presented in the material-based patterns. In order to learn false depth information, real pattern noises can be added to the approaches.

Tu et al. [18] analyzed the movement of the eye, mouth, and head to determine temporal features. Face spoofing based on motion cues can be detected with a CNN-LSTM network. With CNN, high discriminating features are extracted from the image, which is then classified by LSTM into different movements. A spoof video is one that contains no distinguishable movements from the camera. Nevertheless, real-world videos and replay attacks cannot be included in this method.

Wang et al. [19] explored the possibility of multimodal face recognition an attention-based PAD that uses spatial and channel information. Using RGB, depth, and combined input modalities, a face image is acquired. A softmax classifier is used to classify each modality's output as a real or fake face using RESNET features. Photo spoofing attacks can be detected, but video replay attacks cannot.

Liu et al. [20] introduced a light reflection-based face antispoofing technique. The light captcha is generated by generating a random sequence of light cues and intensities, which are then manipulated to cast light. In order to predict the liveliness of the frames, a multitask CNN is used to capture and analyze the frames. To capture the image, the solution used multimodality based on varying light intensities. The accuracy of liveliness detection can be reduced by adding noise through artificial light sources in the data acquisition process.

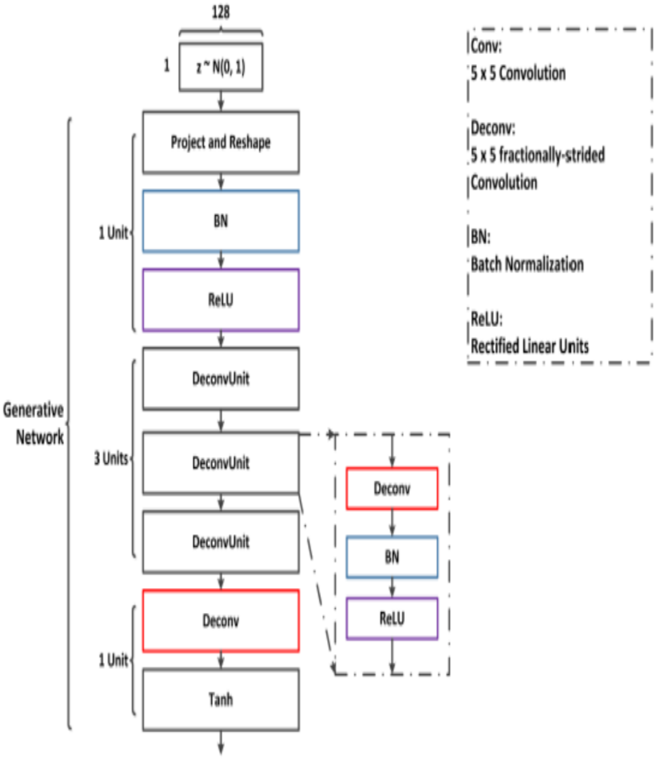

Fig. 1. Generative network

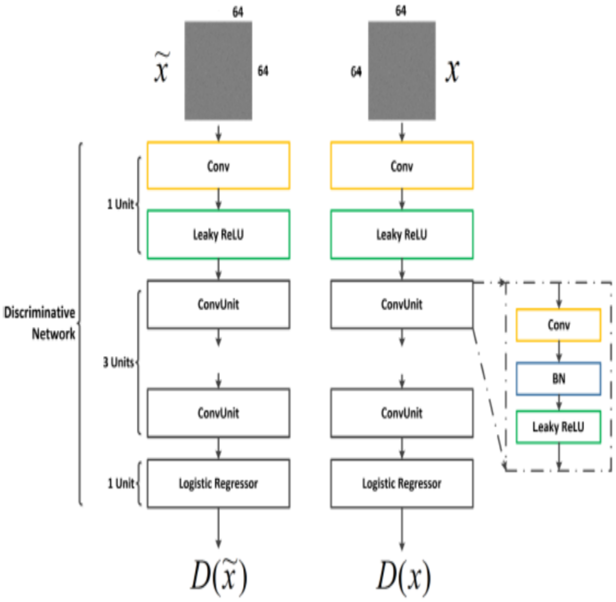

Fig. 2. Discriminative network

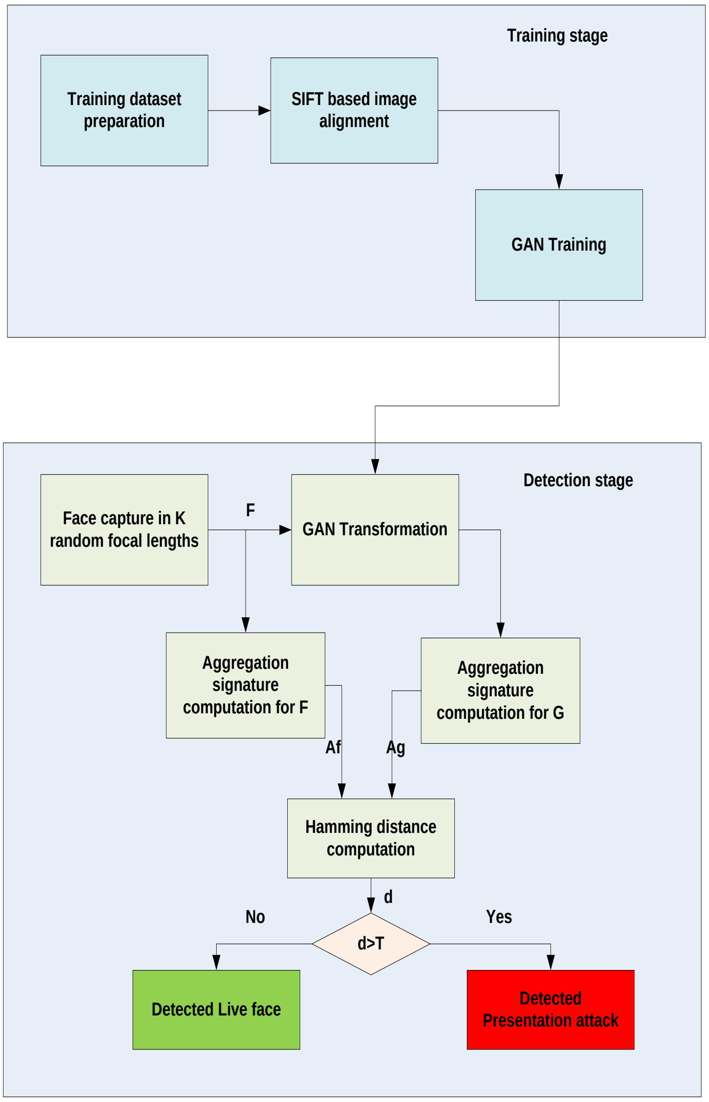
Training stage
Training dataset preparation
SIFT based image alignment
GAN Training
Detection stage
Face capture in K random focal lengths
F
GAN Transformation
Aggregation signature computation for F
Aggregation signature computation for G
Af
Ag
Hamming distance computation
d
d>T
No
Yes
Detected Live face
Detected Presentation attack

Fig. 3. Architecture of focus challenge based PAD

Table 1. Literature summary

Author	Solution	Gap
Mohamed et al. [1]	Convolutional Neural Networks is used for matching between captured faces and live pattern to detect liveliness	Without any challenge, attacker can deceive the CNN matching process by presenting captured videos
Kim et al. [3]	Real and fake faces are compared on the basis of their depth information to detect liveliness	Observing the difference between the projection of noise and ear in different views can result in the approach failing if the ears are hidden
Souza et al. [4]	By detecting illumination variations using deep learning network, liveliness is detected	Unless illumination varies greatly, real images cannot be analyzed by this approach
Parveen et al. [5]	Using skin texture analysis to detect face liveliness using a Dynamic Local Ternary Pattern (DLTP).	By presenting faces with varying illuminations, the method can be easily deceived
Wen et al. [7]	The images are acquired from multiple devices to detect spoofing, and features like reflections, blurs, and chromatic moments are compared between the images.	Hackers can create fake samples with information to deceive attack detection when the device and acquisition parameters are hacked
Tirunagari et al. [8]	Proposed a classification pipeline that combines dynamic mode decomposition, local binary patterns, and support vector machines to detect liveliness	The solution cannot detect replay attacks in videos with long durations that include face movements.
Zheng et al. [14]	Used multiscale, and depth features to detect spoofing attacks	The method fails for replay attack
Song et al. [15]	Two cameras capture images and depth features are extracted from it and compared to detect liveliness.	Face video can be prepared with multi focus to deceive the liveliness process.
Tu et al. [18]	Face spoofing based on motion cues can be detected with a CNN-LSTM network	The method fails in presence of real-world videos and replay attacks
Liu et al. [20]	light reflection-based challenge is introduced and liveliness is detected.	The accuracy of liveliness detection can be reduced by adding noise through artificial light sources in the data acquisition process

The summary of the literature is presented in Table 1. Most the approaches fails in presence of captured videos are presented. They are not resilient against replay attacks. Liu et al [20] proposed a challenge mechanism based on light reflection but it can be made ineffective by introducing artificial light sources. This work solves the problem of replay attack in existing approaches and proposes an effective challenge response mechanism.

3. Focus Challenge Based Pad

The proposed focus challenge PAD is built on foundations of GAN [30]. GAN has two networks: generative network and discriminative. Generative network produces samples with aim to deceive the discriminative network and the discriminative network attempts to check if the sample is real or created by the generative network. With competition of both these networks, a close to real sample is produced by the generative network. GAN networks are being adopted to generate synthetic data due to ability to adapt to complex distributions.

The objective function GAN is given as

$$L_{GAN} = E_{\bar{x}\sim P_g}[D(\bar{x})]\text{-}\ E_{\bar{x}\sim P_r}[D(x)] + \lambda E_{\bar{x}\sim P_x}[(||\nabla_{\bar{x}} D(\bar{x})||_2 - 1)^2]$$

The distribution over V is given as P_r. The distribution over which generator produces data is given as P_g. The uniform samples over P_r and P_g is given as P_x

Range of focal length for Camera is fixed between a interval of x-n…x…..x+n where x is the default focal length, x+n is the maximum focal length and x-n is the minimum focal length.

The architecture of the proposed solution for presentation attack detection is given in Figure 3. The solution has two stages : (i) training stage and (ii) detection stage.

In the training stage, the GAN is trained with training set with training dataset created as follows. A training dataset of face image at focal length x and the focal length f as input and the face image at focal length x+f as output is created. GAN is trained with the dataset. The trained GNN, taking face image y and focal length g as input produces the face image taken at y+g as output. The architecture of the generator and discriminative network used in this work is given in Figure 1 and Figure 2.

In the detection stage, face image of any person P is captured in K different focal length which are selected randomly apart from the first which is the default focal length x. The corresponding faces F={F_x, F_1,.. F_(K-1)}. The faces are first aligned based on SIFT features with SIFT alignment procedure given in [31]. Once the faces are aligned, the face F_x and the K focal length in order are passed to GAN to get the focal length adjusted faces G={F_x, G_1,.. G_(K-1)}.

Aggregation signature is computed for the set F and G. The aggregation signature computation procedure for each image set is as follows.

Quaternion Discrete Cosine Transform (QDCT) is applied over each of the images in the set. QDCT for an image f(x,y) is calculated as

f(x,y)=A_n^q f(x,y)+ ∑_(s=1)^n▒〖[D_(s,1)^q f(x,y)+D_(s,2)^q 〗 f(x,y)+D_(s,3)^q f(x,y)] (2)

Where A_n^q f(x,y) is the low frequency band and D_(s,1)^q f(x,y) is the high frequency band of the image. After QDCT is applied on the image a low frequency part, n groups of high frequency parts are obtained.

The frequency of the coefficients is given below.

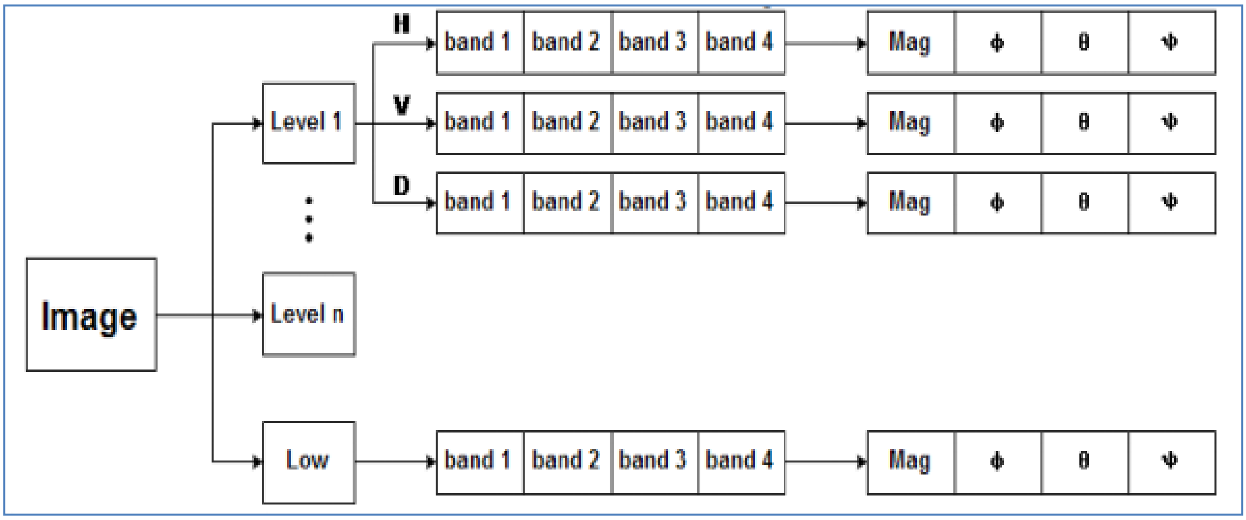

Fig. 4. QDCT Coefficients of Image

To reduce the dimension of the coefficients, average fusion is done for low frequency sub bands. High frequency sub bands are fused using a fusion rule based on maximum value of energy of coefficients. The average fusion rule for fusing the low frequency bands is given as average of the coefficients pair wise between the Low frequency coefficients of two patch images. The fusion rule for fusing the high frequency sub bands is given as selecting the maximum value of coefficient between the pair wise high frequency sub bands.

The QDCT coefficients are given as input to a frequency domain convolutional neural network as given in Figure 5.

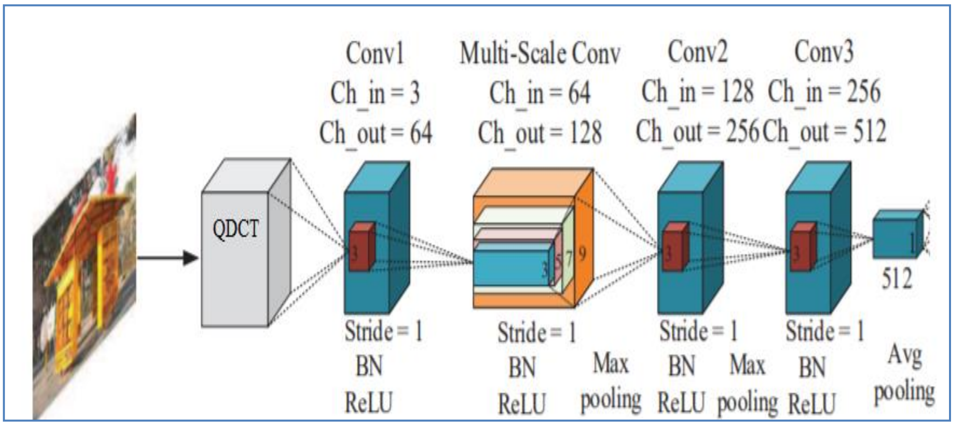

Fig. 5. QDCT CNN

The coefficients pass through a sequence of ReLU and max pooling layer and a final average pooling layer to provide an output of 1× 512 dimension feature vector. The CNN configuration used for feature extraction is given in Table 1. An aggregation signature is constructed from the feature vectors belonging to same image patch as below:

A unit random vector of dimension d (d<512) is generated {r_0,r_1,…r_d}. Each element is sampled from a Gaussian function with mean 0 and variance 1. The d vector is put together into a matrix D of dimension 512 × d. This is generated on time at time of collecting the video as input for tracking.

A inner product between the feature vectors v and the matrix D is done to get vector u=D^T v

For every vector u, following transformation function tf is applied produce the transformed feature vector u¯.

$$tf(u) = \begin{cases} 1\ r.u\ \geq 0 \\ 0, r.u < 0 \end{cases}$$

$$\bar{u} = \{\mathrm{tf}_{\mathrm{r1}}(\mathrm{u}), \mathrm{tf}_{\mathrm{r2}}(\mathrm{u}), \ldots . \mathrm{tf}_{\mathrm{rd}}(\mathrm{u})$$

The feature vectors belonging to same image patch is now represented as bit stream of length d called as aggregation signature of the target image patch.

The benefits of converting the features of same patch to binary bit stream of aggregation signature have two benefits of: compressed form and reduced time complexity for matching the aggregation signature.

The aggregation signature is computed for both set F(A_f) and G(A_g). The hamming distance is measured between aggregation signatures of the two sets.

d=Hamming_distance(A_F,A_G) (3)

When d is less than threshold (T), the face is detected as lively and when d is greater than T, presentation attack is detected.

The pseudo code for detection process is given below

Algorithm: Detection

```
Input: captured face F={F_x, F_1,.. F_(K-1)}.
Output: attack or live
G={F_x}
For m=1:K-1
    G = { G U GAN(F_m) }
end
A_f=Aggregation_singature(F)
A_g=Aggregation_singature(G)
d=Hamming_distance(F,G)
if d>T
    return attack;
else
    return normal
```

4. Results

The performance of the proposed solution is evaluated against 30 different person faces. The faces were acquired using Microsoft LifeCam studio in different focal length. Since there were no standard datasets with faces acquired in different focal lengths, this method is adopted for data collection. A dataset with 40 different person faces with valid faces in different focal length for 30 person and presentation attack for 10 person is created and used for testing.

The performance of the proposed solution is compared against defocus method proposed by Kim et al (2015) [3] attention based solution proposed by Zheng et al (2021)[14] and spatial gradient solution proposed by Wang et al (2020) [29]. The performance is compared in terms of Attack Presentation Classification Error Rate (APCER), Bona Fide Presentation Classification Error Rate (BPCER), and Average Classification Error Rate (ACER). The lower the values of these error rates, the performance is better.

The performance test is conducted in four environments as given in table 2

Table 2. Environment for testing

Env1	under random lighting and background.
Env2	random attack media.
Env3	transformation of the attack camera equipment.
Env4	All above three factors combined

The performance in Env1 are measured and given in Table3

Table 3. Env1 results

Env1			
Solution	**APCER**	**BPCER**	**ACER**
Zhen et al (2021)	1.4	1.8	1.0
Wang et al (2020)	1.0	0.0	1.0
Kim et al (2015)	1.2	1.6	1.0
Proposed	0.60	0.49	0.30

The performance in Env2 are measured and given in Table 4.

Table 4. Env2 results

Env2			
Solution	**APCER**	**BPCER**	**ACER**
Zhen et al (2021)	2.6	0.8	1.7
Wang et al (2020)	2.5	1.3	1.9
Kim et al (2015)	2.3	0.7	1.3
Proposed	0.58	0.51	0.39

The performance in Env3 are measured and given in Table 5.

Table 5. Env3 results

Env3			
Solution	**APCER**	**BPCER**	**ACER**
Zhen et al (2021)	2.0	3.9	2.8
Wang et al (2020)	3.2	2.2	2.7
Kim et al (2015)	2.1	2.5	2.3
Proposed	0.69	0.59	0.57

The performance in Env4 are measured and given in Table6

Table 6. Env4 results

Env4			
Solution	**APCER**	**BPCER**	**ACER**
Zhen et al (2021)	4.2	4.6	4.4
Wang et al (2020)	6.7	3.3	5.0
Kim et al (2015)	4.1	3.1	4.4
Proposed	1.74	1.50	2.31

From the results, the proposed solution is found to have lower values of error compared to existing works. The proposed solution is more robust to changes in attack pattern and lighting. Use of random challenge has lowered the error in classification between real and fake samples.

Kim et al used focus based liveliness detection similar to the proposed solution, but it used handcrafted features and feature fusion. Compared to it, the deep learning based signature used in proposed solution performed better in term of errors.

The focal length was varied in four values of 8cm, 16cm, 32 cm and 48 cm. The Half total error rate (HTER %) was measured for various focal length and the result is given in Figure 6. HTER % of proposed solution is compared against defocus solution by Kim et al .

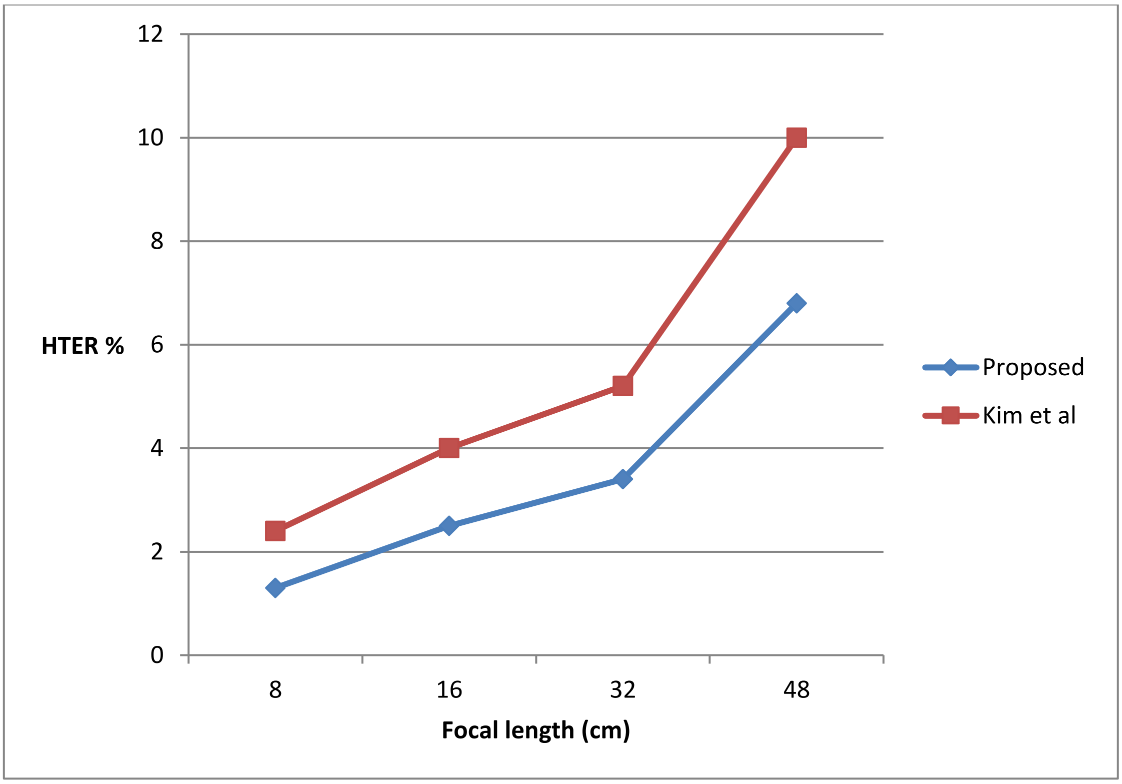

Fig. 6. HTER vs focal length

The HTER % has reduced by 54.2% in proposed solution compared to Kim et al. The error has reduced by large value in proposed solution due to use of deep learning based features compared to handcrafted histogram features in the Kim et al.

Receiver operating characteristics plot is given in Figure 7.

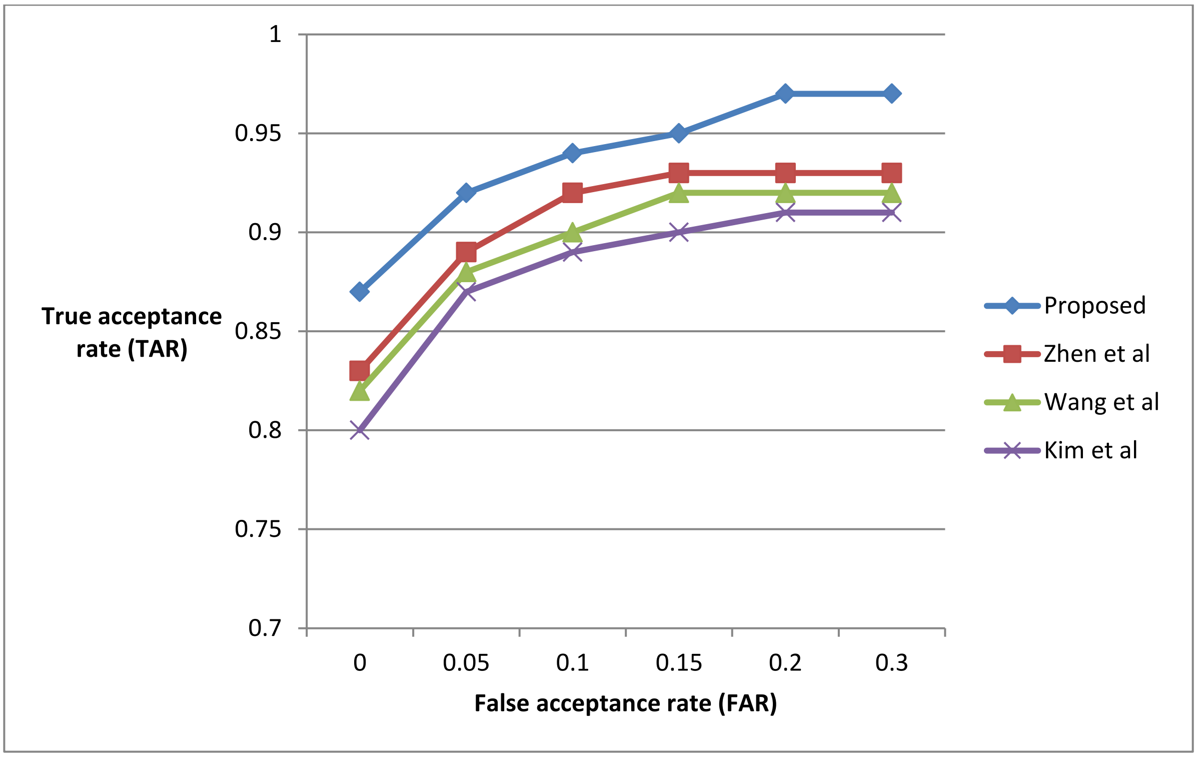

Fig. 7. ROC plot

The ROC area is higher in proposed solution compared to existing works indicating a better sensitivity of proposed solution.

The computation complexity is measured in the proposed solution for various challenge size and the result is given in Figure 8.

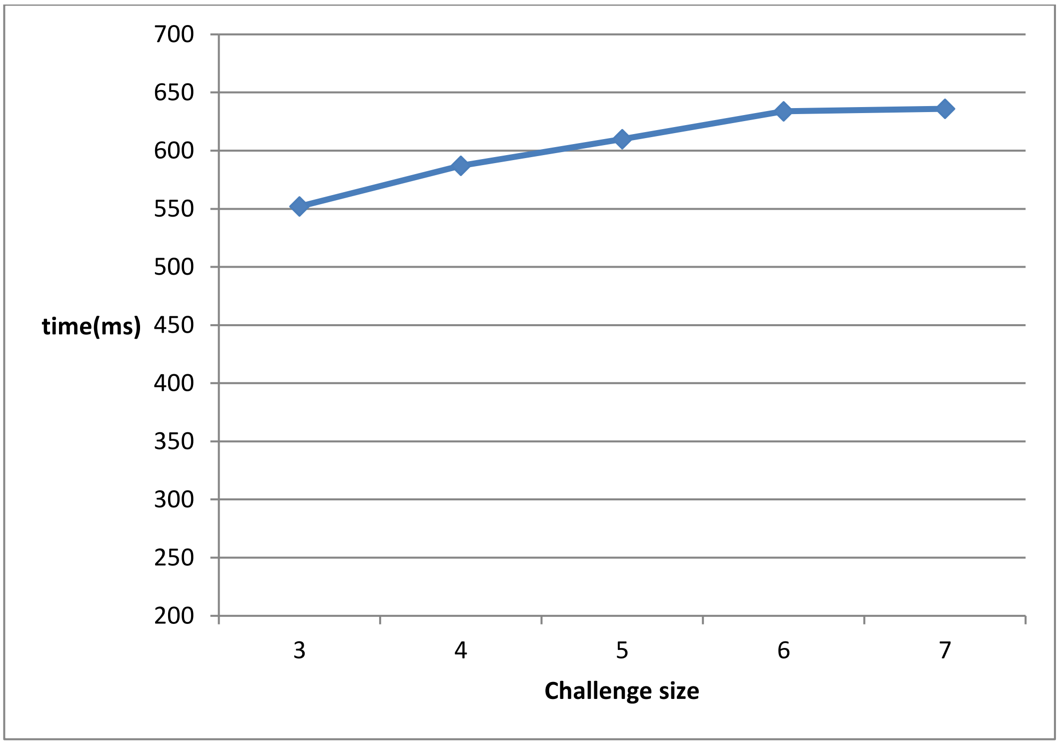

Fig. 8. Computational complexity

For challenge size of 3, the time taken for matching in proposed solution is 552 ms and when the challenge size is doubled to 6, the computation time increased only by 14%. Thus with increase in challenge size, the computation time increases only by smaller margin in the proposed solution.

5. Conclusion

Focus challenge based presentation attack detection was proposed in this work. The proposed solution used GAN to generate expected facial images for different random focal length. Aggregation signature computed for these facial images are compared to acquired face images from camera to detect presentation attacks. Use of GAN along with deep learning based aggregation signature provided higher accuracy of presentation attacks detection. The proposed solution was resilient to various disturbances during image acquisition. The Attack Presentation Classification Error Rate (APCER) in proposed solution is atleast 1.3 times less compared to existing works.

Acknowledgements

This research was supported by Ramaiah Institute of Technology and Global Academy of Technology. We thank our colleagues from both the institutions who provided insight and assisted the research.

Author contributions

Rohini B R: Has formulated the concept and methodology along with preparing and writing the original draft of all the sections with analysis. **Yogish H K:** Validation were carried out by along with writing, review and editing were guided.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] A. Mohamed, M. M. Nagah, M. G. Abdelmonem, M. Y. Ahmed, M. El-Sahhar and F. H. Ismail, "Face Liveness Detection Using a sequential CNN technique," 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC), 2021, pp. 1483-1488,.

[2] Y. Zhang, Z. Yin, Y. Li, G. Yin, J. Yan, J. Shao, and Z. Liu, “Celebaspoof: Large-scale face anti-spoofing dataset with rich annotations,” arXiv preprint arXiv:2007.12342, 2020.

[3] Kim, S.; Ban, Y.; Lee, S. Face liveness detection using defocus. Sensors 2015, 15, 1537–1563.

[4] De Souza, G.B.; Da Silva Santos, D.F.; Pires, R.G.; Marana, A.N.; Papa, J.P. Deep texture features for robust face spoofing detection. IEEE Trans. Circuits Syst. II Express Briefs 2017, 64, 1–5.

[5] Parveen, S.; Ahmad, S.M.S.; Abbas, N.H.; Adnan,W.A.W.; Hanafi, M.; Naeem, N. Face liveness detection using dynamic local ternary pattern (DLTP). Computers 2016, 5, 10.

[6] Akhtar, Z.; Foresti, G.L. Face spoof attack recognition using discriminative image patches. J. Electr. Comput. Eng. 2016, 2016, 1–14.

[7] D. Wen, H. Han, and A. K. Jain, “Face spoof detection with image distortion analysis,” IEEE Transactions on Information Forensics and Security, vol. 10, no. 4, pp. 746–761, 2015.

[8] S. Tirunagari, N. Poh, D. Windridge, A. Iorliam, N. Suki, and A. T. S. Ho, “Detection of face spoofing using visual dynamics,” IEEE Transactions on Information Forensics and Security, vol. 10, no. 4, pp. 762–777, 2015.

[9] Z. Boulkenafet, J. Komulainen, and A. Hadid, “Face antispoofing based on color texture analysis,” in Proceedings of the IEEE International Conference on Image Processing (ICIP ’15), pp. 2636–2640, Quebec City, Canada, September 2015.

[10] J. Zhou, K. Shu, P. Liu, J. Xiang and S. Xiong, "Face Anti-Spoofing Based on Dynamic Color Texture Analysis Using Local Directional Number Pattern," in 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 2021 pp. 4221-4228.

[11] Li, X.; Komulainen, J.; Zhao, G.; Yuen, P.C.; Pietikäinen, M. Generalized face anti-spoofing by detecting pulse from face videos. In Proceedings of the 2016 23rd International Conference on Pattern Recognition (ICPR), Cancun, Mexico, 4–8 December 2016; pp. 4244–4249

[12] Hasan, Md Rezwan & Mahmud, S & Li, Xiang. (2019). Face Anti-Spoofing Using Texture-Based Techniques and Filtering Methods. Journal of Physics: Conference Series. 1229. 012044. 10.1088/1742-6596/1229/1/012044.

[13] Cai, Rizhao & Li, Zhi & Wan, Renjie & Li, Haoliang & Hu, Yongjian & Kot, Alex. (2021). Learning Meta Pattern for Face Anti-Spoofing.

[14] W. Zheng, M. Yue, S. Zhao and S. Liu, "Attention-Based Spatial-Temporal Multi-Scale Network for Face Anti-Spoofing," in IEEE Transactions on Biometrics, Behavior, and Identity Science, vol. 3, no. 3, pp. 296-307, July 2021

[15] Song, Xiao & Zhao, Xu & Lin, Tianwei. (2017). Face Spoofing Detection by Fusing Binocular Depth and Spatial Pyramid Coding Micro-Texture Features. 96-100. 10.1109/ICIP.2017.8296250.

[16] R. Cai, H. Li, S. Wang, C. Chen, and A. C. Kot, "DRL-FAS: A Novel Framework Based on Deep Reinforcement Learning for Face Anti-Spoofing," IEEE Transactions on Information Forensics and Security, vol. 16, pp. 937-951, 2020

[17] Z. Yu, X. Li, X. Niu, J. Shi, and G. Zhao, "Face anti-

spoofing with human material perception," 2020, arXiv:2007.02157

[18] Tu, Xiaoguang & Zhang, Hengsheng & Xie, Mei & Luo, Yao & Zhang, Yuefei & Ma, Zheng. (2019). Enhance the Motion Cues for Face Anti-Spoofing using CNN-LSTM Architecture.

[19] G. Wang, C. Lan, H. Han, S. Shan and X. Chen, "Multi-modal face presentation attack detection via spatial and channel attentions," IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2019, pp. 1584-1590

[20] Y. Liu, Y. Tai, J. Li, S. Ding, C. Wang, F. Huang, and R. Ji, "Aurora guard: Real-time face anti-spoofing via light reflection," 2019, arXiv:1902.10311

[21] Chou, Chao-Lung. (2021). Presentation attack detection based on score level fusion and challenge-response technique. The Journal of Supercomputing. 77. 10.1007/s11227-020-03461-1.

[22] Melnikov A, Akhunzyanov R, Oleg K, Luckyanets E (2015) Audiovisual liveness detection. In: International Conference on Image Analysis and Processing (ICIAP 2015), vol 9280, pp 643–652

[23] Boutellaa E, Boulkenafet Z, Komulainen J, Hadid A (2016) Audiovisual synchrony assessment for replay attack detection in talking face biometrics. Multimed Tools Appl 75(9):5329–5343

[24] Benitez-Quiroz C.F., Srinivasan R., Martinez A.M. EmotioNet: An accurate, real-time algorithm for the automatic annotation of a million facial expressions in the wild; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; Las Vegas, NV, USA. 26 June–1 July 2016; pp. 5562–5570.

[25] Tao S.Y., Martinez A.M. Compound facial expressions of emotion. Natl. Acad. Sci. 2014;111:E1454–E1462

[26] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” in CVPR, 2001, pp. 511–518

[27] Asthana, S. Zafeiriou, S. Cheng, and M. Pantic, “Robust discriminative response map fitting with constrained local models,” in CVPR, 2013

[28] Z. Boulkenafet, J. Komulainen, L. Li, X. Feng, and A. Hadid, “Oulu-npu: A mobile face presentation attack database with real-world variations,” IEEE International Conference on Automatic Face & Gesture Recognition (FG), 2017, pp. 612–618

[29] Z. Wang, Z. Yu, C. Zhao, X. Zhu, Y. Qin, Q. Zhou, and Z. Lei, "Deep spatial gradient and temporal depth learning for face anti-spoofing," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 5042-5051.

[30] J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D.Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial networks. Advances in Neural Informa-tion Processing Systems, 3:2672–2680, 2014

[31] Z. -H. Li, Y. Hou and H. -B. Liu, “Alignment of face images based on SIFT feature,” 2014 International Conference on Machine Learning and Cybernetics, 2014, pp. 597-600

[32] Khare, S. ., & Badholia, A. . (2023). BLA2C2: Design of a Novel Blockchain-based Light-Weight Authentication &amp; Access Control Layer for Cloud Deployments. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3), 283–294. https://doi.org/10.17762/ijritcc.v11i3.6359

[33] Verma, D. N. . (2022). Access Control-Based Cloud Storage Using Role-Fully Homomorphic Encryption Scheme. Research Journal of Computer Systems and Engineering, 3(1), 78–83. Retrieved from https://technicaljournals.org/RJCSE/index.php/journal/article/view/46

[34] Yadav, N., Saini, D.K.J.B., Uniyal, A., Yadav, N., Bembde, M.S., Dhabliya, D. Prediction of Omicron cases in India using LSTM: An advanced approach of artificial intelligence (2023) Journal of Interdisciplinary Mathematics, 26 (3), pp. 361-370.