DEEP LEARNING BASED CHALLENGE RESPONSE LIVELINESS MATCHING FOR PRESENTATION ATTACK DETECTION IN FACE RECOGNITION BIOMETRIC AUTHENTICATION SYSTEMS

Rohini B R
Assistant Professor, Department of CSE, Don Bosco Institute of Technology, Bangalore , Affiliated to VTU, Karnataka, India – 560074,
rohini.br@gmail.com

Dr.Yogish H K
Professor, Department of ISE, Ramaiah Institute of Technology, Bangalore, Autonomous under VTU, Karnataka, India – 560054,
yogishhk@gmail.com

Dr.Deepa Y
Professor, Department of ISE. Don Bosco Institute of Technology, Bangalore, Affiliated to VTU, Karnataka, India - 560074
deepayogish@gmail.com

Abstract

Face recognition based biometric authentication systems are being widely adopted but they are vulnerable to presentation attacks. Detecting presentation attack is important to enhance the security level of face recognition biometric systems. Many presentation attack detection systems (PAD) have been proposed based on comparison of real and presentation image features. But these solutions can be deceived easily by creating the exact replica of real face. To solve this problem, this work proposes a liveliness approach which solves PAD as a challenge response problem. The response of face to a challenge is measured and analyzed to detect PAD. The challenge response matching is realized using a novel Face action unit biased convolutional neural network which selectively skips feature learning in non action unit areas. This novel deep learning model speeds up the challenge response face matching, increases the accuracy of liveliness matching and robust against environmental distortionss.

Keywords**: Face recognition, Spoofing, PADS, liveliness detection, Deep learning, CNN, Emotions mapping.**

1. Introduction

Biometric features have become the most widely adopted secure and reliable authentication systems due to various advantages like difficult to steal, uniqueness, high recognition accuracy and convenience. Biometric authentication systems were found to provide stronger security compared to token based methods (cards, keys etc) and knowledge based methods (username/password), but this is getting challenged recently. Various technological advancements make it easy to generate fake biometric samples with close resemblance to real samples. Though the fake samples can be created for any biometric features like face, iris, fingerprint etc, this work address the problem of faking in face based biometrics. Face recognition based biometric authentication is used in various applications like Smartphone/computer login, passport control, premises access control. In spite of various challenges in illumination and pose variations, it is still being used in biometric authentication systems. Users face presented in front of cameras are captured. Features extracted from the face are matched to features stored in database to recognize the person. Various attacks exploit the loop holes in face acquisition process like printed photographs, masks, or video displays and reduce the security level of the system. These attacks are called as presentation or spoofing attacks. These attacks must be detected and recognition process must be barred to enhance the security level.

Various presentation attack detection (PAD) algorithms are designed for face recognition systems based on color, texture, depth, light reflection analysis. Most the works extract various handcrafted features for real and fake images. The differences in handcrafted features are analyzed either statistically or through machine learning algorithm to detect presentation attacks. However with latest technological advancements like Deep fake, it becomes easy to create fake samples as close to the real image. Thus using features in image alone to detect presentation attack becomes ineffective. Multi modal approaches involving collection from multiple devices looks promising to detect fake, but they too fail in case of careful video replay attacks.

This work proposes a challenge response mechanism to detect presentation attacks. Challenges are presented in terms of sequence of random images and the facial responses are collected in the same channel as face acquisition. The power spectral density variations in the captured face video is compared against the expected results from challenges using a novel Facial action biased convolutional neural network feature. The difference of deep learning feature to the expected feature is then analyzed to detect liveliness. Following are the novel contributions of this paper work.

(i) A novel Facial action unit biased convolutional neural network for liveliness detection with higher accuracy, reduced time for matching and better resilience to environmental distortions

(ii) A novel Challenge response mechanism based on emotion arousal in regions of Facial action units in the 2D scaleogram which is difficult to deceive.

The rest of the paper is organized as follows, In 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 liveliness based PAD scheme 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 works

Mohamed et al [1] trained a sequential convolutional neural network model with CelebA-Spoof dataset [2] to detect live and non live faces. The method was able to achieve 87% accuracy when tested against CelebA-Spoof dataset. The method was not tested for real samples. The approach detects liveliness based on the differences in facial components, The approach works only for artificial images created programmatically and it can be deceived easily by presenting the realistic face image. Kim et al [3] used the effect of defocus to detect face liveliness. The work is based on the depth information differences between real and fake faces. The face images is acquired are two cameras with different focus. The features of focus, power and gradient location and orientation histogram are acquired from the images. The differences in features from two face images are compared to look for consistency. In case of real faces, the difference is high, but fake face present minor difference. The approach can be failed by hiding the ears as it works based on observing the difference between projection of noise and ear in different views.Souza et al [4] proposed LBPnet a variation of typical convolutional neural network accommodating LBP in first layer of convolution by making the convolution to work on LBP values of pixels instead of the original pixels. The resulting high level deep learning features can detect artificially spoofed images. But when real images are presented without much variation is illumination, the approach fails. Parveen et al [5] proposed a Dynamic Local Ternary Pattern (DLTP) for face liveliness detection using skin texture analysis. Textural properties of facial skin are best extracted from DLTP features. The features are then compared between real and fake faces to detect spoofed attacks. The method can be deceived easily by presenting faces with illumination variations.Akhtar et al [6] identified discriminative patches in the face images which has higher correlation to spoofing attack detection. The discriminative patches are found by observing the local intensity in-homogeneity. Statistical inference based observation is made to detect discriminative patches of face image. The discriminative patches are compared between stored and acquired face to detect spoofing. But the method could be easily deceived by presenting same stored images.
Wen et al [7] proposed a face spoof detection scheme based on distortion analysis. Image is acquired from multiple devices and features like reflection, blur and chromatic moment are compared between images acquired from multiple devices to detect spoof. These features have characteristic differences across different acquisition device. By this way, when a fake sample acquired from different device is shown, spoofing can be detected easily. When the device and the acquisition parameters are hacked, then fake samples can be created with information to deceive attack detection.Tirunagari et al [8] proposed a classification pipeline consisting of dynamic mode decomposition, local binary pattern and support vector machine to detect liveliness. Various dynamics like eye blinks, lip movements and face changes are used to detect the liveliness of the face. The solution cannot detect replay attacks with long duration videos which has movements in face regions.Boulkenafet et al [9] used color textural distortions to detect face spoofing. The color textural differences on the luminance and chrominance channels are compared to differentiate between real and fake

samples. But the method is not resilient against replay attacks.Zhou et al [10] proposed a feature extraction method to extract multi scale features from color images with powerful representation capacity. Local directional number pattern with derivative Gaussian mask is used which can resist illumination variations and noises. Spatial temporal variations are accommodated in the local directional number pattern. The method can detect artificially created faces but fails for replay attacks. Li et al [11] used pulse detection for detecting anti spoofing in face recognition systems. The regions near cheeks, chin selected as ROI. The pixels in the ROI region have color value changes with the cardiac pulse change. Power spectral density in these regions between the real and fake faces shows a difference and this can be used to detect liveliness. The method works best for print attacks but fails for replay attacks. Hasan et al [12] proposed a spoof detection technique based on the contrast and dynamic texture features. A modified version of DoG filtering along with local binary patterns variance is used for extraction of features. The features are then classified using support vector machine for detecting spoofed photos. The proposed work can be deceived with replay attack. Cai et al [13] proposed a two-stream Hierarchical Fusion Network to detect spoofing attacks. Meta patterns are extracted from the images using Resnet deep learning model. The Meta patterns are fused with features extracted from RGB images to detect spoofing. Though the method can detect photo spoofing attacks, it fails for replay attack. Zheng et al [14] used depth and multi scale features to detect spoofing attacks. Two stream spatial temporal networks is proposed to extract both depth and scale information from images in form of features. A fully connected network layer classifies these features to spoof or genuine samples. The method works best of spoof photo attacks but fails in presence of replay attack.

Song et al [15] used binocular camera to capture the face image. From it, a novel depth and texture feature is extracted. The features are then classified with SVM classifier to differentiate between real and spoofed faces.

Cai et al [16] proposed a deep learning method to extract discriminating features from local regions. Convolutional neural network along with recurrent neural network is used to learn the representation features of various local regions. The local regions features are fused with global image level features and used for detecting spoofed face. The method works best for artificial faces created using Deepfakes but it cannot detect replay attacks.Yu et al [17] rephrased the problem of face spoof detection as material recognition problem as spoof are done through materials like skin, glass, paper and silicone. Intrinsic material based patterns are extracted using bilateral convolutional neural networks The material based patterns present information about the depth and from this depth information spoofs are classified. The approaches can be deceived by adding real pattern noises to learn false depth information.Tu et al [18] extracted temporal features based on eye, mouth and head movements. A joint CNN-LSTM network is proposed for detecting face spoofing based on the motion cues. CNN extracts the high discriminative features from image and this is classified by LSTM to various movements. The input videos from cameras without any discriminative movements were classified as spoofs. But the method does not work for replay attacks and realistic videos as inputs.Wang et al [19] proposed a multi modalPAD using spatial and channel attention. Four input modalities are used to acquire face image – RGB, Depth, combined. Resnet features are extracted from each modalities output and classified using softmax classifier to real or fake face. The method is able to detect photo spoof attack but fails for video replay attacks.Liu et al [20] proposed a face anti-spoof method using light reflection properties. A random sequence of light cues and intensities called as light captcha is generated and screen is manipulated to cast light as mentioned in light captcha. The frames are captured and analyzed by a multi task CNN to predict the liveliness. The solution used multi modality in terms of varying light intensity to capture the image. But noises can be added through artificial light sources in data acquisition process to reduce the accuracy of liveliness detection.

3. Challenge response based liveliness PAD

Many existing works for PAD were discussed in previous section. These approaches used various cues like depth, texture, light reflection, material properties, face movements, multi mode acquisition, dual camera etc in detection of fake samples. But these approaches can be easily deceived by careful construction of videos compensating the various cues and launching replay attacks. The risk of deceiving can be avoided by throwing random challenges and measured the feedback during acquisition. The feedback can be analyzed and compared with expected response to detect liveliness in the video. Random challenge and response is proposed as means to thwart deceiving problem in previous works. Challenge response approaches have been earlier proposed [21-23]. But these approaches are bi-channel with use of speech channel based challenge response. Different from it, this work proposes a challenge response based liveliness PAD using the same visual channel for acquisition and challenge response.

The proposed solution is based on changes in 68 facial landmarks in response to challenge of sequence of images of different emotions. An in-depth study made in [24] correlating the emotions to changes in different facial landmarks (Fig 1) is given in Table 1.

e-ISSN : 0976-5166
p-ISSN : 2231-3850

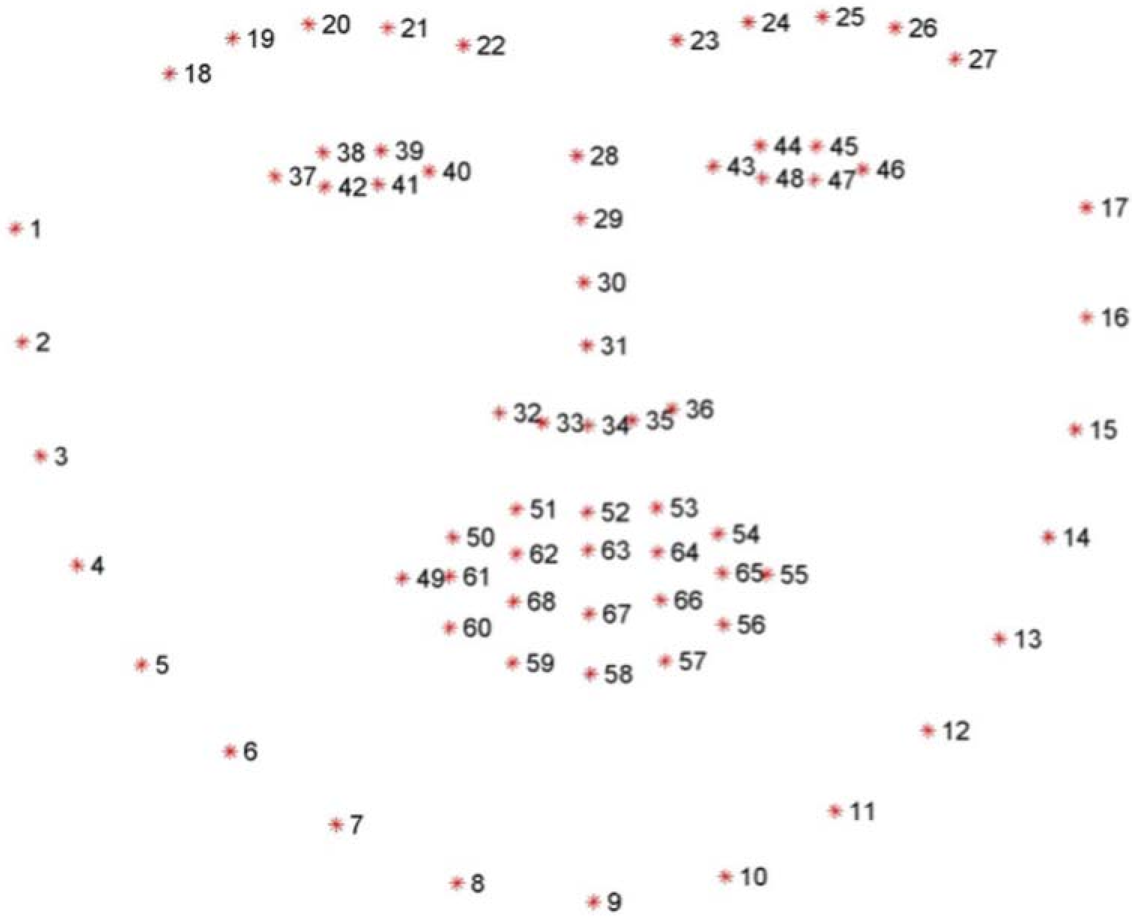

Fig. 1. Facial landmarks

Emotion	Facial landmarks activated
Happiness	12,25
Sad	4,15
Fear	1,4,20,25
Angry	4,7,24
Surprised	1,2,25 ,26
Disgusted	9,10,17
Happily sad	4,6,12,25
Happily surprised	1,2,12,25
Happily disgusted	10.12,25
Sadly fearful	1,4,15,25
Sadly angry	4,7,15
Sadly surprised	1,4,25,26
Awed	1,2,5,25
Appalled	4,9,10
Hatred	4,7,10

Table 1. Emotion to Facial landmark mapping

The architecture of the solution is given in Figure 2. The proposed solution has two stages: training and liveliness detection.

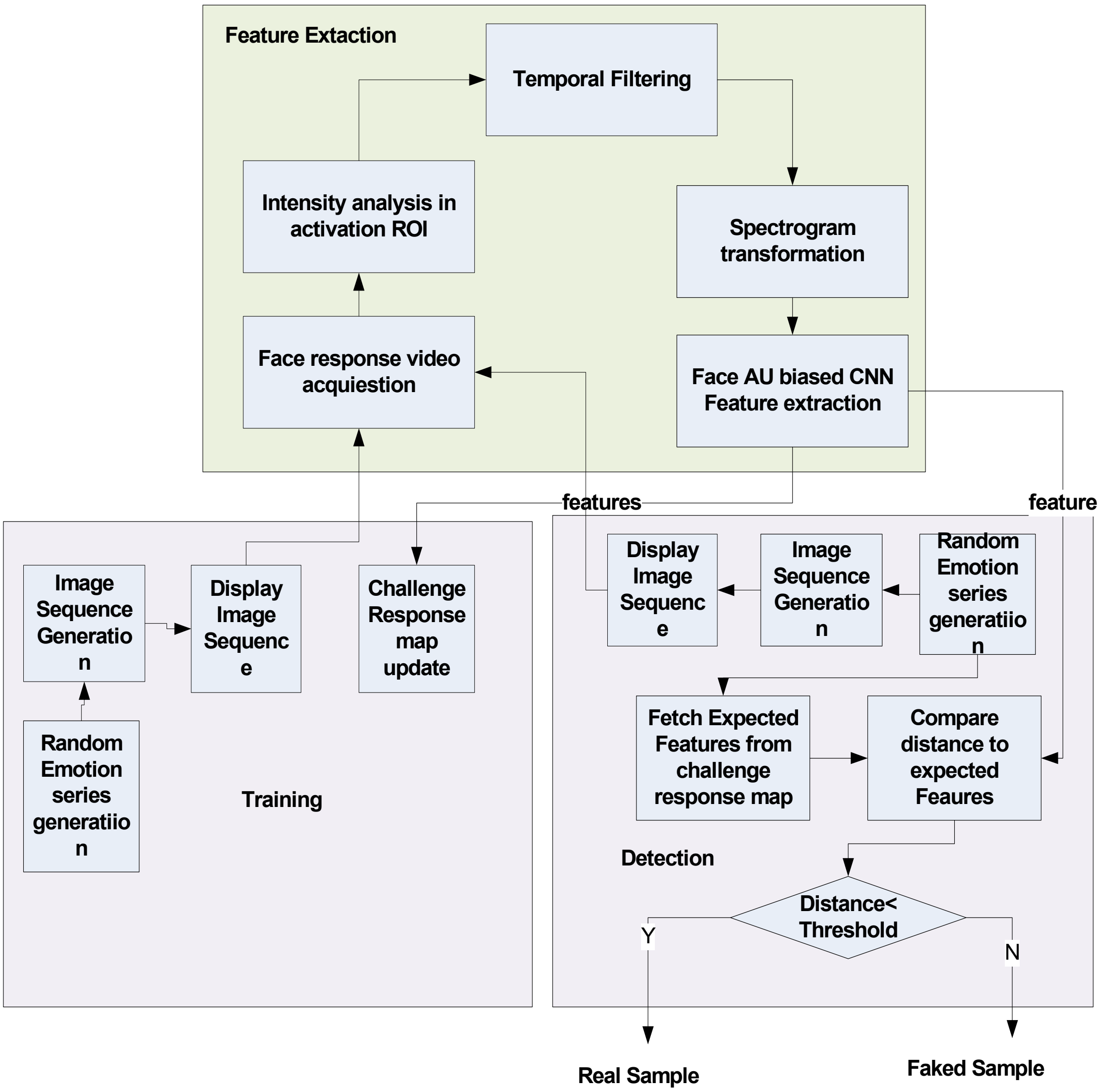

Fig. 2. Architecture of proposed solution

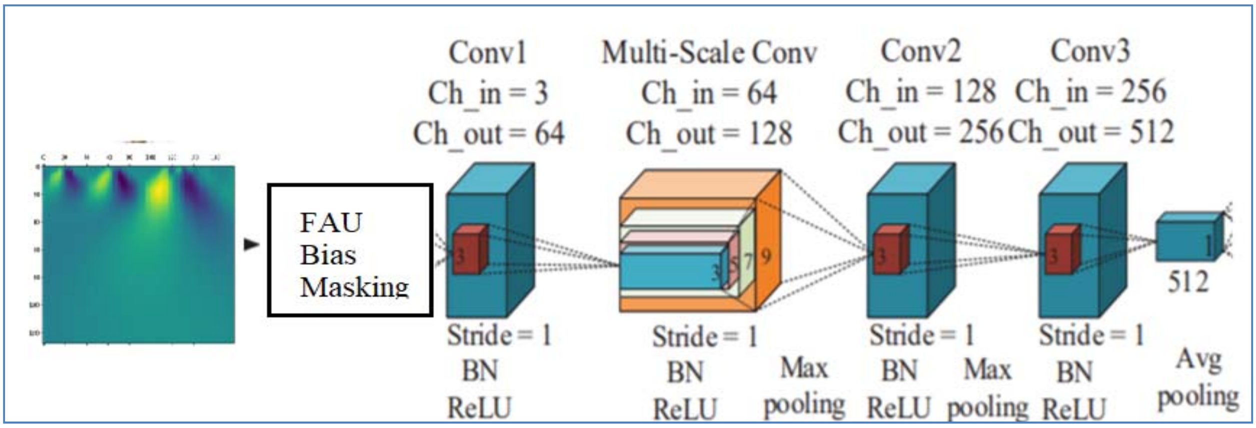

Fig. 3. FAU based CNN

Each of the stages of training and liveliness detection are detailed in below subsections.

3.1. *Training*

In the training stage, the facial responses of the different users when presented with sequence of random images of different emotions are captured. The facial response video is processed frame by frame. Face region is detected in frame using Voila Jones method [26]. Discriminative response map fitting [27] is used to find the 68 different landmarks. Based on the emotions convened in each image a activation set of facial landmarks $\{L_1, L_2, \dots, L_n\}$ in ascending order is found by mapping emotions to activation landmarks in Table 1. A square patch of size m is created for each of activation set of facial landmarks and this becomes the ROI set.

$$ROI = \{S(L_1, m), S(L_2, m), \dots S(L_n, m)\}$$

Where S is the square with size m and center point as L_x. The change in the mean intensity values over these ROI regions are collected and shown in Fig 4 below

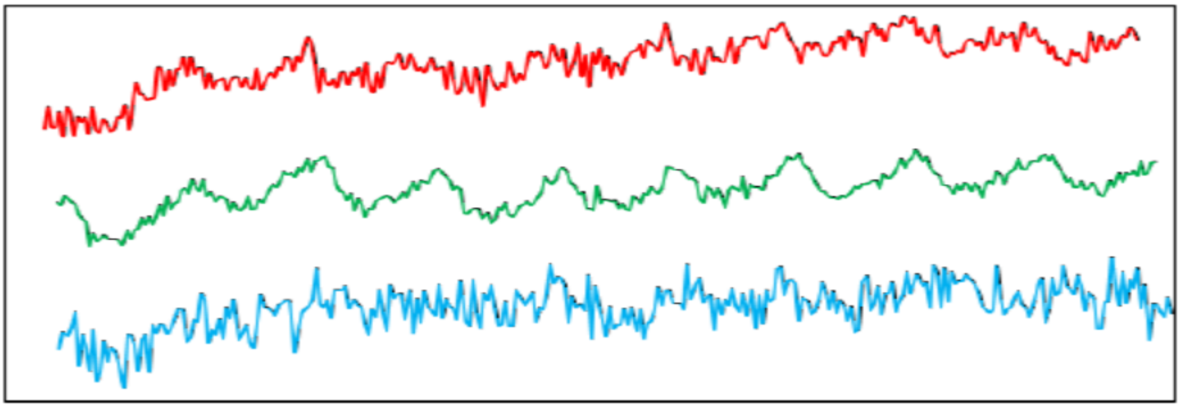

Fig. 1. Intensity plot for the ROI regions

The intensity values in time domain are preprocessed using three temporal filters. De-trending filter is first applied to reduce the slow or non-stationary trend of the signal. Moving average filter is applied to remove random noises. Finite impulse bandpass filter is then applied with cut off frequency of [0.7, 4] Hz. After preprocessing wavelet transform is applied. Wavelet transform helps analysis of signals with dynamic frequency spectrum. Wavelet transforms are of two categories: continuous and discrete. Continuous wavelet transform is given as

$$X_w(a, b) = \frac{1}{|a|^{1/2}} \int_{-\infty}^{\infty} x(t)\varphi(\frac{t-b}{a})dt$$

$\varphi(t)$ is the mother wavelet with scale factor of a and translation factor of b. The application of continuous wavelet transform on a Intensity values results in 2D scaleogram which provides the detailed information about the state space of the system. This scaleogram can be used to understand the dynamical behavior of the system and to distinguish different types of emotion spread over the facial landmarks. In this work continuous wavelet transform with Gaussian as mother wavelet is applied onto the preprocessed mean intensity signal to generate the 2D scaleogram. A sample scaleogram of a mean intensity signal is given in Fig 5.

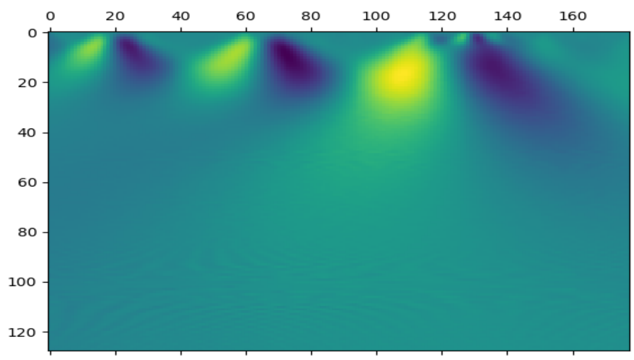

Figure 2 Scaleogram for facial response video

Scaleogram is created for each of the random emotion image sequence in the training set. The scaleogram is given as input to a Deep learning convolutional neural network to provide features of dimension 512. The scaleogram image 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 novel facial action unit (FAU) biased CNN architecture used for feature extraction is given in Fig 3. In this architecture, a facial action unit biased mask is applied to image to learn convolutional features only in the FAU activation regions skipping the irrelevant regions. By this discriminative ability of convolutional features in recognizing emotions increases. Also this reduces the time of matching. A dataset of 2D scaleogram images for each of the emotions of: happiness, sad, fear, angry, surprised, disgusted, awed, appalled and hatred. The scaleogram images for particular emotions are grouped. For each group, a facial action unit bias mask is constructed. The procedure for binary mask construction is given as Algorithm 1. The images in the group are binary OR to a result image. This image is binary thresholded using OSTU in three channels of R, G, B separately and the result is binary OR to get the mask. Mask is prepared for each emotion class separately and a separate FAU biased CNN is trained for each emotion separately.

Algorithm 1: FAU Bias Mask construction
Input: Scaleogram images for particular emotion
Output: Binary mask
1. *res* ← *zero matrix*
2. *for i=1 : no of scaleogram images*
res=bres OR image(i)
end
3. *R_I* ← *OSTU thresholding of R channel of res*
4. *G_I* ← *OSTU thresholding of G channel of res*
5. *B_I* ← *OSTU thresholding of B channel of res*
6. *M_I* ← *R_I OR G_I OR B_I*
7. *return M_I*

The algorithm for extracting FAU bias CNN features is give as Algorithm 2. In this algorithm, image belonging to each emotion class is passed as input. A FAU Biased CNN is created for each emotion class. Image of particular class is passed to its corresponding FAU Biased CNN with binary mask of that class and the image as input. The results features for each of the image in the class are extracted at the last average pooling layer of the CNN as shown in Figure 3. The features belonging to same class are added into a list and mapping between the emotion class label and feature list is kept in a feature map.

Algorithm 2: Feature extraction
Input: Images of each class, binary mask(B) of each class
Output: Feature map
1. *FVMap={}*
2. *for each class K*
Fvlist = []
3. *for each image I*
. *Fv* ← *Invoke_FAU_Bias_CNN(K,B,I)*
. Fvlist.append(Fv)
End

. FVMap.put(K,Fvlist)
End
4. return FVMap

The feature vector for each of the image belonging to the class is learnt using a FAU Bias CNN. The FAU Bias CNN takes the input 2D scaleogram image of size 64×64. Local binary pattern of the image is computed. This local binary pattern is logical AND with the binary mask of the corresponding class. This masked output is passed to next stage of convolution with 7×7 kernel to get the feature map.

$$C(q) = \sum_{m=1}^{M} \sum_{n=1}^{N} AND\big(LBP(q), mask(m)\big).K\,(j)$$

Where N is the number of kernels and M is the number of times mask to be applied onto input image q. This resulting feature map is passed to subsequent stages of the CNN. The final feature of dimension 1×512 is extracted at the average pooling layer.

A map of sequence of emotions (from the random emotion image sequence) and the corresponding feature vector as kept in a challenge response map. This challenge response map is used for detection of liveliness.

3.2. *Detection of liveliness*

During authentication phase, a random index is selected from the challenge response map. A random set of images is selected matching the emotion sequence corresponding to selected random index. This is displayed to the user and the facial response video of user is captured. Scaleogram of the facial response video is created as described by the process mentioned in training. This Scaleogram image is passed to deep learning convolution neural network given in Figure 3 to extract the features. The distance between the extracted feature and the feature stored in the challenge response map for the selected random index is calculated using Euclidean distance. If the distance is less than a threshold (decided by configuration), then face input is classified as live and real. In case the distance is greater than threshold, the input is classified as spoofed.

4. Results

The performance of the proposed challenge response liveliness based PAD is first tested in terms of clustering analysis for the feature discriminating ability of FAU bias CNN features and time taken for feature extraction. The performance is compared against CNN model without FAU bias mask layer. Images belong to same emotion are kept as cluster. EmotioNet [24] dataset is used for experiment. Features are extracted using default CNN and and clustered based on emotions. Clustering analysis is conducted to measure following standard metrics of: Average cohesion, Average separation and Silhouette coefficient [30].

The average cohesion value must be less compared to average separation between clusters for a good cluster. Silhouette coefficient value ranges from 0 to 1 and when the value is towards 1, it indicates good quality clusters.

Emotion	Average cohesion	Average separation	Silhouette coefficient
Clustering results for FAU bias CNN Features			
Happiness	120	341	0.8
Sad	135	421	0.8
Fear	96	452	0.81
Angry	87	321	0.84
Surprised	121	370	0,75
Disgusted	134	351	0.77
Awed	111	372	0.76
Appalled	121	410	0.72
Hatred	125	450	0.73
Average	111.6	387.55	0.77
Clustering results for CNN without FAU bias layer			
Happiness	132	312	0.72
Sad	146	405	0.71
Fear	118	418	0.69
Angry	101	311	0.71
Surprised	134	338	0,72
Disgusted	142	351	0.71
Awed	131	352	0.71
Appalled	141	390	0.69
Hatred	135	414	0.71
Average	131.11	365.66	0.706

Table 1 Clustering analysis results

The results are clustering analysis is given in Table 2.The average cohesion in FAU Bias CNN has decreased by 17.4 % compared to CNN without FAU Bias layer. The average separation in FAU Bias CNN has increased by 5% compared to CNN without Bias. The average SC has increased by 8% in FAU Bias CNN compared to CNN without bias. FAU bias mask has increased the discriminating ability of features. The features have higher significant to the emotion labels.

This has reduced the average cohesion, increased the average separation and SC compared to CNN without FAU bias mask layer. The results reveal that proposed feature extraction method is able to clearly separate each emotion. The cohesion value is low compared to the separation. Also Silhouette coefficient is greater than 0,7 indicating the feature distance are close by for same emotions. This indicates the effectiveness of proposed feature extraction algorithm in clearly separating the emotions. Due to clean separation of multi class emotions with proposed FAU bias CNN, the accuracy of challenge response is increased. The time for feature extraction is compared between FAU bias CNN and CNN without FAU bias layer and the result is given in Fig 6.

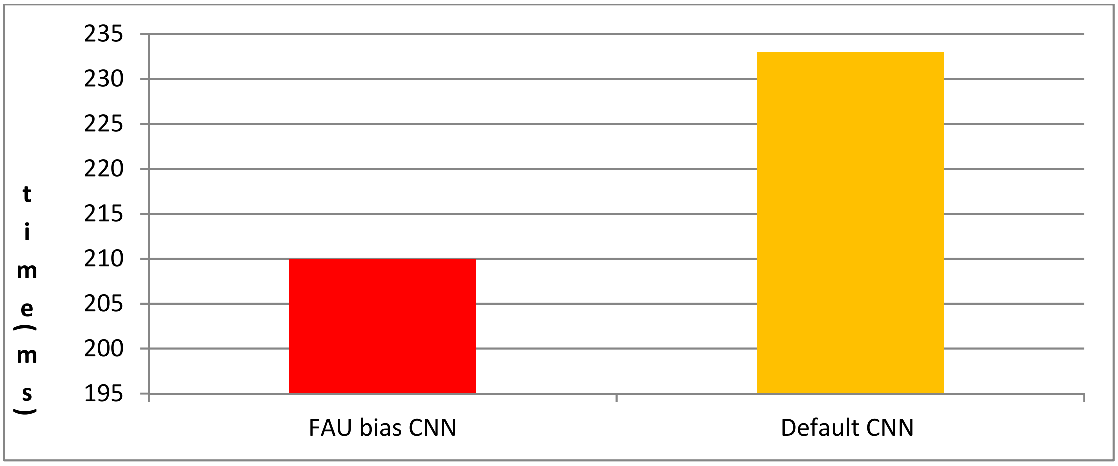

Fig. 3. Comparison of feature extraction time

The feature extraction time has reduced by 9.87% compared to CNN without FAU bias layer. The mask has reduced the feature extraction area and this has reduced the feature extraction in time in proposed solution. The performance of the proposed solution to detect face spoofing is tested against OULU-NPU dataset [28]. The dataset has 990 real face videos, 3,960 fake face videos. The performance of the proposed solution is compared against 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 3. The corresponding results in these environments are given in Table 4 to Table 7.

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

Table 2 Environment for testing

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
Proposed	0.62	0.51	0.32

Table 3 Env1 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
Proposed	0.64	0.56	0.41

Table 4 Env2 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
Proposed	0.71	0.61	0.58

Table 5 Env3 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
Proposed	1.81	1.61	2.36

Table 6 Env4 results

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. The ROC curve plot for the solutions is given in Fig. 7

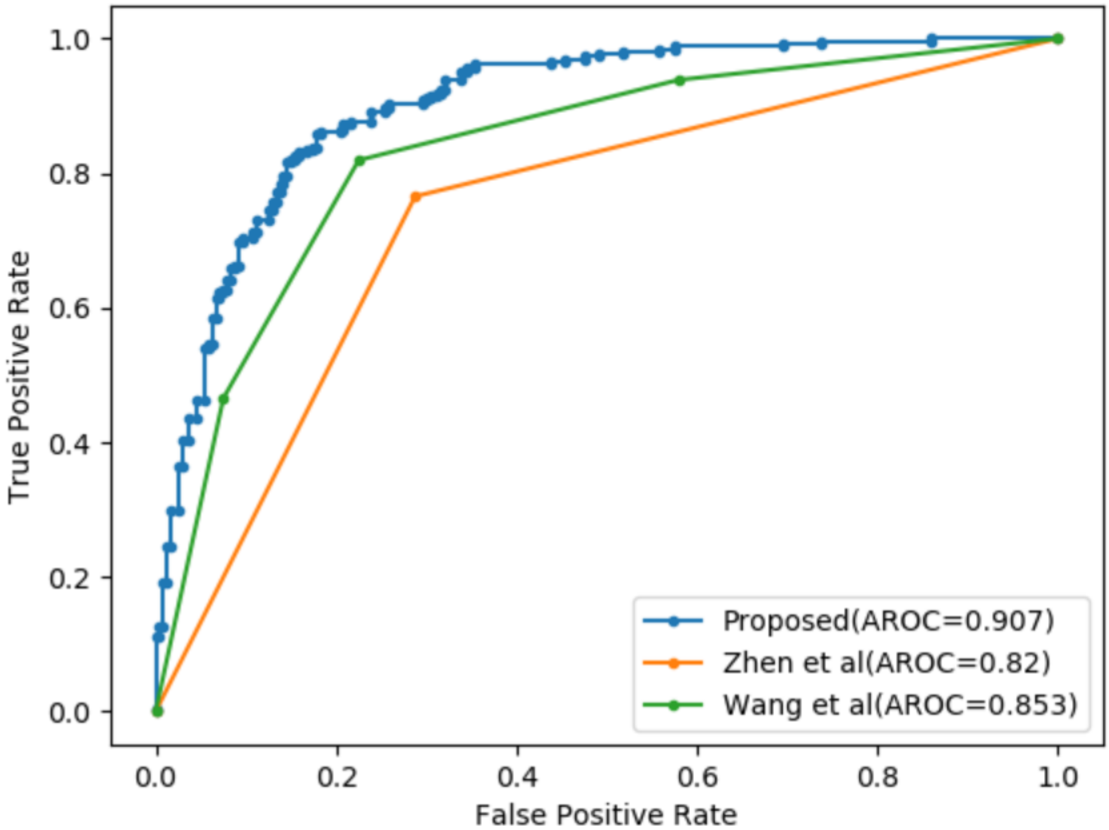

Fig. 4. ROC Curve

From the ROC curve, proposed solution is found to have better performance compared to existing works. The false positives are comparatively lower due to use of multiple random images and selection of ROI regions around the relevant face activation points in the proposed solution.

Half total error rate (HTER) is calculated for different number of emotion images in the challenge sequence and the result is give below:

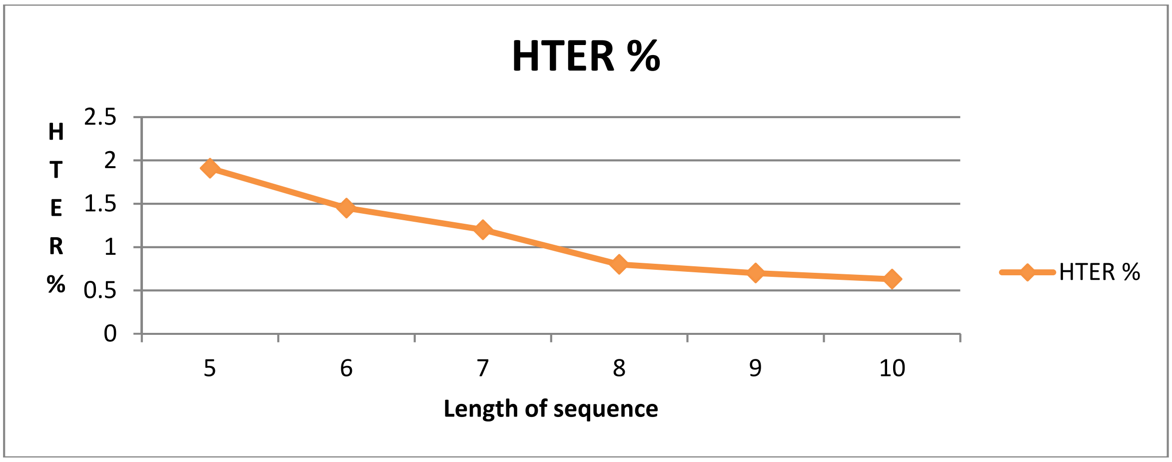

Fig. 5. HTER vs length of challenge sequence

The HTER % reduces with increase with length of challenge sequence.

5. Conclusion

Challenge response based liveliness detection for face recognition is proposed in this work. The work is based on observing the changes in region of interests around the facial landmarks for a random image conveying some emotions. Compared to other works which used other channels like speech, lighting captcha, the proposed system used the same channel for both acquisition and challenge response. The proposed system provided stronger defense against various spoofing attacks. Extending the work considering personalization in responses is in scope of future work.

6. Conflicts of Interest

The authors declare no conflict of interest.

References

[1] A. 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] A. 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] Mariswamy, Latha. (2020). Comparative Analysis of Feature Extraction methods for Kannada Bi-Syllable Words of Dysarthric Speech. Bioscience Biotechnology Research Communications. 13. 332-337. 10.21786/bbrc/13.13/51.

Authors Profile

Rohini B.R B.E,M.Tech in Computer Science & Engineering currently working as Assistant professor in Department of Computer Science & Engineering in Don Bosco Institute of Technology, Bengaluru. She has a teaching experience of 10 years. Areas of interest include Machine learning, Deep learning, Image processing algorithms and actively participated in research and academic activities in the institution.

Dr. Yogish H.K, Professor in Information Science & Engineering in Ramaiah Institute of Technology, Bangalore is a renowned Academician with 22 years of Teaching experience. His areas of Interest are Web Mining, Data Mining, and Cloud Computing.He has worked in several prestigious institutions. He has received Honor Fellow from Institution of Engineers. He has authored several journal and conference papers.

Dr. Deepa Y., working in the Department of Information Science & Engineering at Don Bosco Institute of Technology,Bangalore. Completed B. E in electronics and Communication Engineering from VTU Belagavi, M. Tech in Computer Science Engineering from VTU Belagavi and Ph.D. in Computer Science from VTU Belagavi.She has published and presented 18 papers at National/International Conferences/Journals among 5 papers were indexed by Scopus, and 1 Book Chapter. She has an extensive experience of 17+ years of Teaching Experience. She is a life time member of various Professional bodies – LMISTE, IAENG and Member of Institute of Engineer. Actively participated and won many prizes in sports activities. She filed two patent and reviewer for various conferences/journals. Collaborated with various industry for academic and industry collaboration. She is the proud owner of various awards in the fields of Teaching and Coordination.