Intelligent Automation & Soft Computing
DOI: 10.32604/iasc.2023.030763
Article

Tech Science Press

Spoofing Face Detection Using Novel Edge-Net Autoencoder for Security

Amal H. Alharbi[1], S. Karthick[2], K. Venkatachalam[3], Mohamed Abouhawwash[4,5] and Doaa Sami Khafaga[1,*]

[1]Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia
[2]Department of Electrical Electronics and Communication Engineering, GITAM School of Technology, GITAM Deemed to be University, Bengaluru Campus, Karnataka, 560065, India
[3]Department of Computer Science and Engineering, CHRIST (Deemed to be University), Bangalore, 560074, India
[4]Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt
[5]Department of Computational Mathematics, Science, and Engineering (CMSE), Michigan State University, East Lansing, MI, 48824, USA
*Corresponding Author: Doaa Sami Khafaga. Email: dskhafga@pnu.edu.sa

Received: 01 April 2022; Accepted: 05 May 2022

Abstract: Recent security applications in mobile technologies and computer systems use face recognition for high-end security. Despite numerous security techniques, face recognition is considered a high-security control. Developers fuse and carry out face identification as an access authority into these applications. Still, face identification authentication is sensitive to attacks with a 2-D photo image or captured video to access the system as an authorized user. In the existing spoofing detection algorithm, there was some loss in the recreation of images. This research proposes an unobtrusive technique to detect face spoofing attacks that apply a single frame of the sequenced set of frames to overcome the above-said problems. This research offers a novel Edge-Net autoencoder to select convoluted and dominant features of the input diffused structure. First, this proposed method is tested with the Cross-ethnicity Face Anti-spoofing (CASIA), Fetal alcohol spectrum disorders (FASD) dataset. This database has three models of attacks: distorted photographs in printed form, photographs with removed eyes portion, and video attacks. The images are taken with three different quality cameras: low, average, and high-quality real and spoofed images. An extensive experimental study was performed with CASIA-FASD, 3 Diagnostic Machine Aid-Digital (DMAD) dataset that proved higher results when compared to existing algorithms.

Keywords: Image processing; edge detection; edge net; auto-encoder; face authentication; digital security

1 Introduction

At present, face biometrics has been popularly used in various intelligent products and authentication systems, including all entrance guard systems, Automated Teller Machines (ATM)s, and mainly mobile phones. People no longer necessitate remembering their credit card's complex password; instead, they may pay for things simply by scanning their faces. Face locks are now supported by several Android smartphones, including Huawei's honor series. Customers place their faces in front of the inbuilt camera to unlock the phone. However, certain illegal persons or criminals may easily breach these biometric systems with a single face photo or captured video and steal the money and information. At the University of Hanoi, the security and vulnerability team said that it is efficiently bypassing these methods with legitimate users' spoofed face images. Hence, anti-spoofing of the face is an essential requirement for the face security system. Existing facial biometric devices are pretty sensitive to spoofing attacks of many kinds.

A person's identification can be fabricated by showing a video, photograph, or 3D mask of a legitimate individual to the sensor; a person can also apply makeup and undertake any surgery to impersonate a real user. A sample flow diagram of a spoofing attack is depicted in Fig. 1. The research becomes very challenging, and several conferences have recently been held to promote the advances made in this field. Deep learning-based algorithms [1,2] have recently adopted cross-entropy with binary for loss and are considered as a direct supervisor in the anti-spoofing of faces.

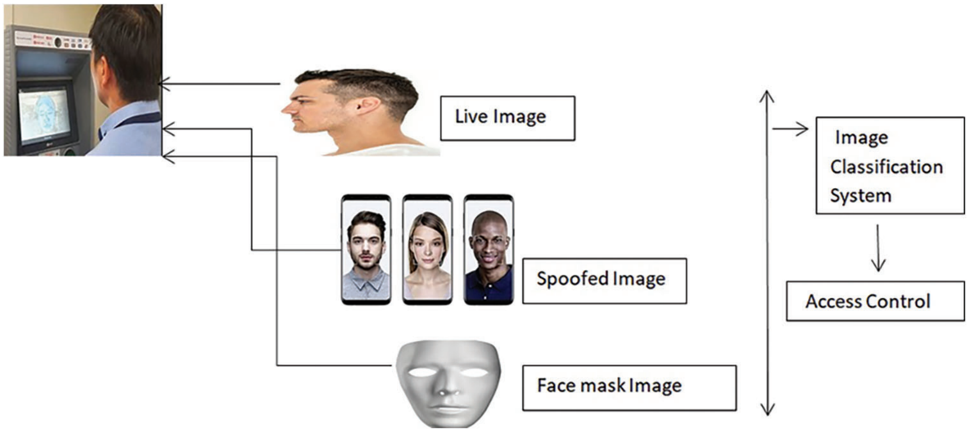

Figure 1: Spoofing attacks

Nevertheless, Convolutional Neural Network (CNN)-based approaches concentrate on deeper semantic aspects, which are ineffective at defining fine-grained data between live and spoof images and are readily ineffectual when the ambiance changes which means various lighting effects.

Anti-spoofing is based on features of picture and motion quality. It is classified into two categories: dynamic and static. The static technique is a single static snapshot being analyzed. On the other hand, the dynamic approach analyzes the temporal and spatial characteristics of a sequence of input frames. Computer vision projects are highly encouraged to use deep learning techniques for accurate outcomes. Most modern deep learning-based FAS approaches, such as Visual geometry group (VGG) [3], Residual Network (ResNet) [4], and DenseNet Network (DenseNet) [5], are based on image classification task-based backbones [6,7].

The networks have often overseen a loss in binary cross-entropy, which gets arbitrary patterns like screen bezels rather than the character of spoofing practices. Deep learning principles have outperformed other architectures when interacting with complicated computer vision challenges and image-related issues. We present a robust approach that equals state-of-the-art outcomes in detecting the assaults in this environment. This research contributed to this paper as follows:

- Edge detection is applied to differentiate the life and spoofed image. This technique is used for the first time in face anti-spoofing supervising data. Its 3D structure is used for finding dissimilarities between spoof and live faces.
- The novel effective Edge-Net auto-encoder model is proposed, which employs the detecting region of discontinuities in the images, enhancing the anti-spoofing in the authentication system.
- Autoencoders for dimensionality reduction, faces are pre-trained with local and fixed regions to improve the framework performance and convergence speed.
- All Weight is acquired from the autoencoder, and the weights of pre-trained encoders are used to classify the spoofed and real images.
- This paper investigates the application of a convolutional autoencoder to downscale the input image dimensionalities. It supports by avoiding unnecessary features, thus enhancing the framework with generalized capability.
- The layers automatically learn Liveness features in the architecture. Then images are classified as real and spoofed are achieved with a softmax classifier.
- This Edge-Net model obtains state-of-the-art outcomes on benchmark datasets with intra-dataset testing.

1.1 Definitions

Deep convolutional neural network (CNN): This neural network is constructed for structured array processing of data like text and images. Computer vision mostly uses CNN for effective results. Several visual applications like image classification depend on CNN for a perfect training and matching images. It also established successful text classification models using natural language processing.

Edge detection: This approach is used in image processing for edges identification of objects within images, and it processes data by identifying the discontinuities in brightness. Detection is performed for data extraction in image processing and image segmentation, machine vision, and computer vision.

1.2 Motivation

By spoofing attacks, attackers can gain the personal identity of others, money theft, spread malware via infected attachments or links, and bypass the control of network access. This problem motivates us to find a solution through Liveness detection for face recognition in biometrics, which is the capability of a computer system to determine if the person in front of the camera is a spoofed image or a natural person's image.

1.3 Organization of the Paper

This paper is structured as follows from Section 2. In Section 2, we go over similar research that has been done by many researchers utilizing various methodologies and algorithms. The architecture of the proposed work is explained in Section 3. The proposed model is implemented in Section 4 using the tools and datasets that were used. Section 5 discusses and captures the outcomes of individual approaches by incorporating various test levels of the model, and Section 6 concludes our paper.

2 Structures

A paper for a literature survey is extensively done on face detection techniques. Spoofing attacks use different methods, which are classified into many groups based on (a) motion (b) texture (c) multi-cues fuse (d) image quality analysis. Fundamentally, the texture of real face skin is varied from spoofing attacks in mediums. Spoof materials are like Liquid Crystal Display (LCD) and soft plastic can be identified by the texture analysis concept. These techniques dive image to greyscale and colors. The End to End [8] spoofing on face uses a deep learning algorithm for accurate face spoofing detections. New descriptors are used to represent the appearance of images and structures of both original and spoofed features. Spoofing assaults in [9] are trained the ensemble types for improving the ability to exist methods and find the unseen face attacks easily.

In paper [10] Investigated the potential impact of auxiliary supervision on deep learning approaches in the PAD problem. Another model uses CNN and Recurrent Neural Network (RNN) networks to measure the face depth and Remote Photoplethysmography Pulse (RPG) signal of the video. Fusing the estimated depth and rPPG signal differentiates attack accesses from actual accesses.

The Framework with multiclass features is adopted in a few types of research to plan the Presentation Attack Detection (PAD) problem to prevent spoofing attacks. However, the two-class method has many limitations, including (a) hardship in determining an efficient decision limit to discern actual from spoofing samples [11], (b) worse novel attacks generalization [12] (c) the over-fitting problems, limited size dataset training problems [13].

Face spoofing using binary representation has disadvantages like detection formulation, many authors discussed one class (anomaly) based algorithms. It provides attractive quality when compared with its two-class counterpart. The training data can be elongated to original data for training the anomaly classifier. It consists of huge generalization due strongest to the nature of new attacks by face spoofing. Recent research has established the advantages of designing a client-based anomaly face anti-spoofing system [14]. In the article [15] implement a new client-specific anomaly Stacking ensemble for detecting spoofing face detection. It tends to find two-class demerits with learners as an unseen scenario attack and develop a classifier pool for Stacking. They have taken three deep CNNs, facial region, and three anomaly experts making 63 spoofing detectors. In the article [16] 3D point cloud (3DPC) is employed as data supervision for face anti-spoofing. It proposes an effective network for encoder-decoder (3DPC-Net) that only depends on 3DPC supervision.

The standard operator known as Central Differences Convolution (CDC) [17] proposes anti-spoofing detection of the face. Based on CDC, a Central Difference Convolutional Network is constructed. It proposes CDCN++, which contains a CDC backbone search and multi-scale attention model with fusion. The article [18], introduces a convolutional autoencoder to minimize the dimensionality of images. Moreover, classification and feature extraction of spoofed and natural is done using pre-trained weights. The architecture applied in [19], uses two convolutional neural network streams that work on two spaces, which are multi-scale retinex space and Red, Green, and Blue (RGB) color space. RGB space provides detailed data about the face texture, and the latter high-frequency data is captured in the face for discrimination. Another approach [20] fused the Simplified encoded CNN called Weber Local Descriptor and Local Binary pattern models.

The extracted features were integrated with preserving the data intensity and edge orientation. Further classifier Support Vector Machine (SVM) was applied to find spoofed images or real images. Then [21–37] detail map is used for training the fully convolutional network (FCN), which achieves good results than the traditional CNN approach.

3 Proposed Methodology: EdgeNet

This work proposes a novel Edge-Net deep learning structure for spoofing detection in the faces while getting access through biometric authentication. It would guarantee more efficient learning of spoofed features based on the below features

- The first step in the proposed work, the edges are detected in the input images to find the discontinuities in the pictures, which helps identify the spoofed image from the live images.

- The dimensionality reduction phase helps to reconstruct the input images using an increased layer of the convolutional autoencoder. The weights from the autoencoder are given as input to another model consisting of a fully connected layer and the flattened layer comprising 1024 neurons.
- Further, a softmax is applied to classify whether the input is real or spoofed.

The general structure of the proposed Edge-Net model is shown in Fig. 2.

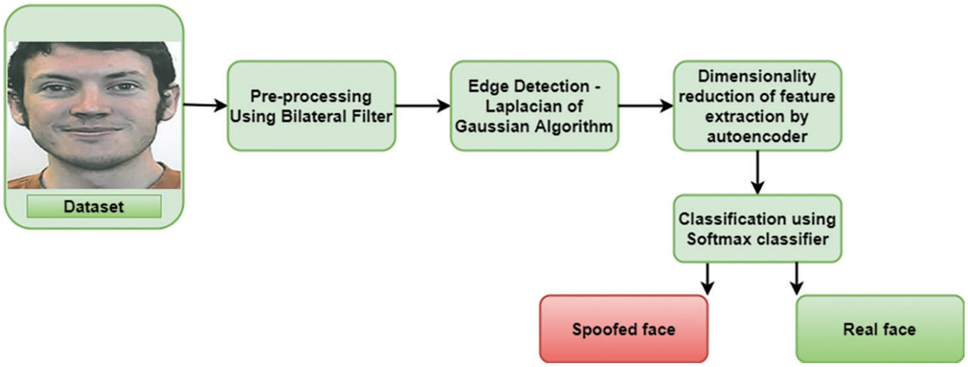
Dataset
Pre-processing Using Bilateral Filter
Edge Detection - Laplacian of Gaussian Algorithm
Dimensionality reduction of feature extraction by autoencoder
Classification using Softmax classifier
Spoofed face
Real face

Figure 2: The general structure of the proposed Edge-Net model

3.1 Pre-Processing

Pre-processing is needed to classify the spoofed and actual image, enhancing the image for further processing. The steps in pre-processing involve three stages: removing noise, resizing the image, and normalization.

3.1.1 Resize Image

All images are not the same in size in the dataset and processing multiple-size data does not provide standard results. Every image is resized as 256 × 256 and sent for the next step of processing. Downsampling and upsampling methods are applied to resize the input image.

3.1.2 Removal of Noise

The noise is eliminated from the original input face image by applying the bilateral filter to upgrade the effectiveness in classifying spoofed and authentic images. The bilateral filter applies an Gaussian Filter, but it possesses one more multiplicative element, an operation of pixel intensity variance. It makes sure that only pixel intensity parallel to that of the principal pixel is incorporated in calculating the blurred intensity value. This filter conserves the edges of the image, which helps us clearly fr the edge detection practical filter can be constructed as follows:

$$BiF[I]_p = \underbrace{\frac{1}{W_p}}_{\text{Normalization factor}} \sum_{q\in S} \underbrace{G_{\sigma_s}(||p-q||)}_{\text{Space Weight}} \underbrace{G_{\sigma_r}(|I_p - I_q|)}_{\text{Range Weight}} I_q \tag{1}$$

Here $BiF[I]_p$ ans pixel outcome the of p, and the RHS is sum of pixels q weighted using Gaussian function. I_q is pixel q in normalization range weight ranges a is tine with added terms to the previously derived equation. σ_s denotes the kernel spatial extent. The neighborhood size, and σ_r shows the edge with low amplitude. It confirms that pixels with that value intensity equal to pixel center value are assumed for blurring, intensity sharpness changes are maintained

3.1.3 Normalization

To enhance the image's contrast by applying normalization, the contrast of the image depends on pixel intensity value. This proposed work's normalization process uses the RGB pixel technique [22] called compensation. The adaptive compensation illumination is used on the black pixel with histogram equalization.

3.1.4 Feature Extraction

In this paper, an autoencoder is applied to obtain the feature data from the input. Autoencoder, a neural network-based feature extraction technique, produces abstract elements from high-dimensional data with excellent success. However, they are officially learned using supervised learning techniques or "self-supervised". It is an unsupervised learning method. Generally, it trains the autoencoder as one of the vast models that reproduce the input.

3.2 Edge-Net Dataset Creation

In this paper, we adapt the Laplacian of Gaussian (LoGS) to mark the edges of the image to enhance the prediction level much better than the previous model. It works on the zero-crossing technique, i.e., when the second-order derivative comes across zero, that particular location coincides with a higher level It is called an edge area. Here the Gaussian operator minimizes the noise, and the Laplacian operator determines the sharp edges.

The formula defines the Gaussian functional:

$$G(x,\ y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \left(\frac{x^2 + y^2}{2\sigma^2}\right) \tag{2}$$

where, σ-is the standard deviation.

And the LoGS operator is calculated from

$$LoGS = \frac{\partial^2}{\partial x^2}\ G\ (x,\ y) + \frac{\partial^2}{\partial y^2} G(x,\ y) = \frac{x^2 + y^y - 2\sigma^2}{\sigma^4} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \tag{3}$$

The live face image doesn't have darkened edges, but the spoofed images have, as shown in Fig. 1. We can find out the significant differences among the images that enhance EdgeNet's prediction level.

3.3 Autoencoder

Autoencoder is kind of artificial neural network that encodes an image by applying the encoder $a_e\ (\theta)$ and rebuild an image X with the help of the decoder. An autoencoder contains two subparts, the encoder and the decoder, which can be interpreted as transitions ∅ and φ, such as

$$\varnothing :\ \aleph\ \rightarrow F \tag{4}$$

$$\varphi : F\ \rightarrow\ \aleph \tag{5}$$

$$\varnothing,\ \ \varphi = \underset{\varnothing,\varphi}{\arg min}\ ||א -\ \ (\varphi \circ \varnothing)\ \aleph\ ||^2 \tag{6}$$

The feature space F has low dimensionality than the, the feature vector $\varnothing(x)$ is represented as compressed of the input x. Autoencoding is an algorithm applied to compress the data. An autoencoder with trained faces would execute efficiently on face compression. The decompressed output data would be as lightly degraded as compared to given input images but can reduce the loss.

3.4 Proposed EdgeNet Architecture

The convolutional autoencoder consists of convolution and max-pooling layers. The proposed architecture of the EdgeNet autoencoder is shown in Tab. 1. In the next step, the CNN network computes by forwarding pass, which applies the weight of the re-trained encoder in a flatter section. Then the images trained are labeled as original and spoofed. As a consequence of a fully connecting layer, the whole network is designed to train the above images. The images received after reducing the dimension are given as input to the CNN's fully connected layers with a softmax classifier which finds the given facial image is original or spoofed. This architecture is trained for 75 epochs. The proposed work uses an edge detected image of the live and spoofed face image, shown in Figs. 3 and 4.

Table 1: Dataset description

Dataset	Data description	Attack type
Replay attack dataset	Short video	Captured videos in various environment
3D mask attack dataset	Short video	Creation of Mask based on the face image
CASIA FASD dataset	Images and short video	Video, cut attack, and warped photo
Edge-Net dataset	Images	Edge detected images of CASIA-FASD

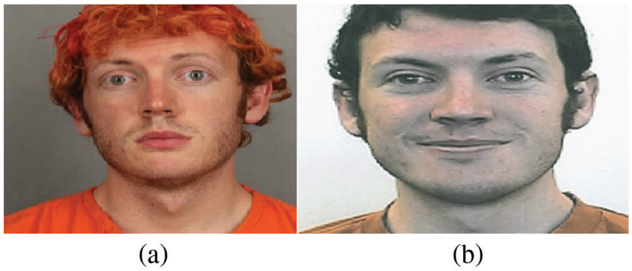
(a) (b)

Figure 3: (a) Live face. (b) Spoof face

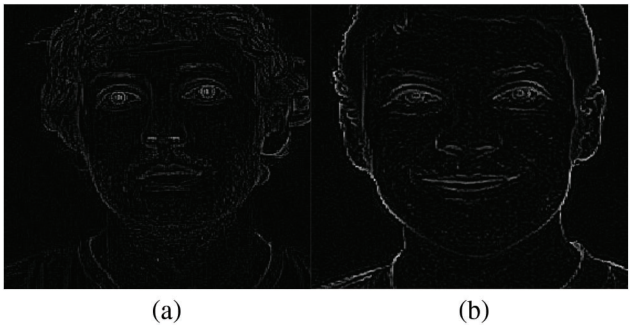
(a) (b)

Figure 4: a. Edge detected live face image; b. Edge detected spoofed face image

In classifying spoofed images and real images, this research proposes an EdgeNet autoencoder structure in Fig. 5. Pre-processed and edge detected image is given as input to the encoder. The autoencoder further reduces the image's dimension [38,39] but increases the extraction of the features. Then feature of the given image is learned by the encoder and decoder. The autoencoder consists of a convolutional and max-pooling layer. The pooling layer is used to downsampling and upsampling the image at the encoder and decoder end.

Then forward a Convolutional Neural Network bypass, which uses the pre-trained weight of encoder in a flatten area. The input images are reconstructed to be trained and it is labelled as spoof or real. This model uses weights of pre-trained encoder. Therefore, these layers are not re-trained. After dimensionality reduction completion, the images are directly fed into the CNN fully connected layer [40,41]. The softmax classifier classifies whether the image is real or spoofed. The Edge-Net autoencoder model algorithm is given as a flow chart in Fig. 6.

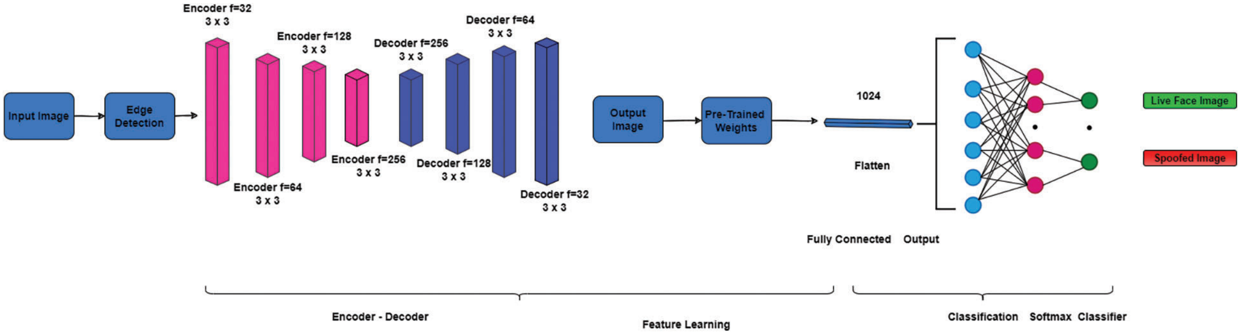

Figure 5: Edge-Net deep convolutional neural network model

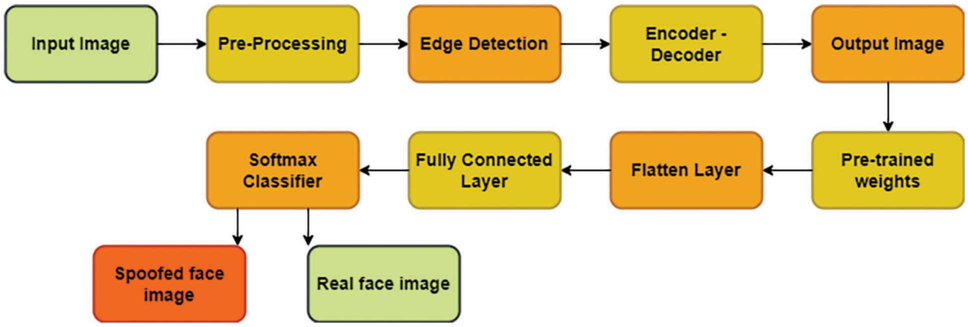

Figure 6: Flow chart for the EdgeNet algorithm

Algorithm 1: Edge-Net Autoencoder for Spoofed and Real image Classification

```
Step 1: Given the input image $I_{in}$
Step 2: Pre-processing the image $I_{in}$,  $BiF[I_{in}]_p = \frac{1}{W_p}\sum_{q \in S} G_{\sigma_s}(||p-q||)G_{\sigma_r}(|I_p - I_q|)I_q$
Step 3: Then $BiF[I_{in}]_p$ feed in to Laplacian of Gaussian model $LoGS = \frac{\partial^2}{\partial x^2} G(x, y) + \frac{\partial^2}{\partial y^2} G(x, y) = \frac{x^2+y^y-2\sigma^2}{\sigma^4}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$, to detect the edge of the images.
Step 4: Then the LoGS image given as input to the autoencoder $\varnothing$,  $\varphi = \underset{\varnothing,\varphi}{\arg min} ||\aleph - (\varphi \circ \varnothing) \aleph ||^2$ for dimensionality reduction and feature extraction.
Step 5: Find the weights from the pre-trained model (receive weights ($W_x$ and bias$B_x$)
Step 6: Convert the data into single dimensional array.
Step 7: Then fully connected layer ($I_{in}$ – (FS – 1)) is determined
Step 8: Softmax classifier performs the classification on input image spoofed and real i = 1,
        $S(i_n) = \frac{e^{in}}{\sum_{m=1}^{l} e^{xm}}$    (2)
Step 9: The output image is classified.
```

This edge detection algorithm helps to improve the accuracy of fake faces in the digital security system. Laplacian of Gaussian model with autoencoder is first time implemented in fake face detection strategy. The result computation is discussed in upcoming sections.

4 Experimental Setup

4.1 Datasets Used

The proposed algorithm is evaluated with three benchmark datasets CASIA FASD [23], 3DMAD with Replay-attack IDIAP dataset, and our Edge-Net dataset.

4.1.1 D Mask Attack Dataset (3DMAD)

The Mask Attack 3D Database (3DMAD) is a human face spoofing dataset. It occupies 76510 frames of 17 persons, recorded using the Kinect for authenticated access and spoofing attack. Each structure has:

Image depth of (640 × 480 pixels – 1 × 11 bits).

The RGB corresponding images (640 × 480 pixels – 3 × 8 bits).

Eye position is manually annotated (concerning the RGB image).

The data in 3 types is collected in all subject session, and for every session, five videos of 300 frames are recorded. The recordings done in controlled condition, with front view and neutral expression. A sample of the 3DMAD images is shown in Fig. 7. The eye positions are manually labelled for every 1st frame, 61st frame, 121st frame, 181st frame, 241st frame, and 300th frame in each video. They are interpolated linearly for the rest. The mask real size are obtained using "ThatsMyFace.com." The dataset also possesses the face images utilized to provide these masks and paper-cut masks created by the same service and employing the same images.

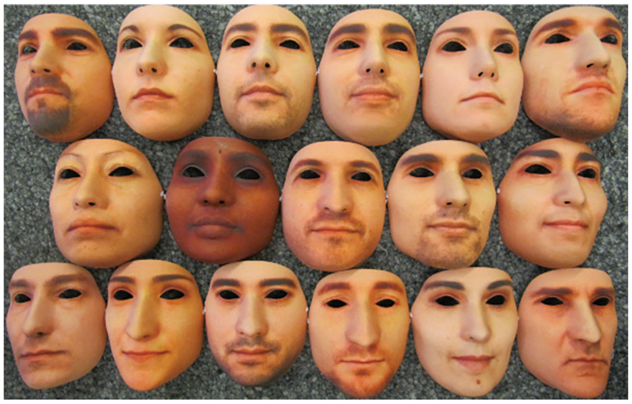

Figure 7: Images in 3DMAD dataset

4.1.2 Replay Attack Dataset

Itis given by the Idiap Research Institute contains spoofed and live videos of 50 various subjects. It has 1300 above videos and each.mov.movclip is nearly 9 s in duration. All generated videos with the help of webcam built in or recording video of themselves. High-resolution video is captured in same session by Canon PowerShot SX150 IS camera to do the attacks. Twenty attack short duration with videos are captured for every subject in two various modes such as adverse and controlled.

4.1.3 CASIA-FASD Dataset

This dataset includes three attack types like attack-cut, printed photograph, and wrapped photograph. To do the photo attack warped, both videos and images are captured with the help of a camera. The attacker purposely put effort to simulate facial motion and intact photo. There is no region cut-off in face, unlike in cut photo attack, in which eye portions are removed off, and the attacker hides back of the eye holes in this attack. Video attacks are done by the captured video in the camera. All the description of datasets is provided in Tab. 1.

4.2 PreProcessing (Video)

The videos of the Replay attack, 3D Mask Attack, CASIA-FASD Datasets are reformed into frames by applying clideo tool. It is an openly available video processing tool that converts frames of the videos. We can select the running speed of the video play of every stop moving frame. There are three modes: fast, medium, and slow. Then select a clip rate. It differs from 0.2 to 1.5 s. These setting regulates how repeatedly will get a frame from the video. The images are then given as input to the Edge-Net architecture for reduce the dimension and extract the feature.

4.3 Dimensionality Reduction

After completion of pre-processing the videos of three datasets applying the clideo tool, the experiments are done on these datasets. The images are remade using an autoencoder. Then the remade images are exposed to extract the feature and classify real or spoofed images.

4.4 Training

Fig. 8. depicts the accuracy gained while performing feature extraction of image input from the Edge-Net, 3DMAD, CASIA, Replay attack dataset for the first 15 out of 75 epochs. The loss value during the remake of images applying the autoencoder approach on images input from all three datasets. It can be recognized that the loss has been reduced during the reconstruction of images using autoencoders.

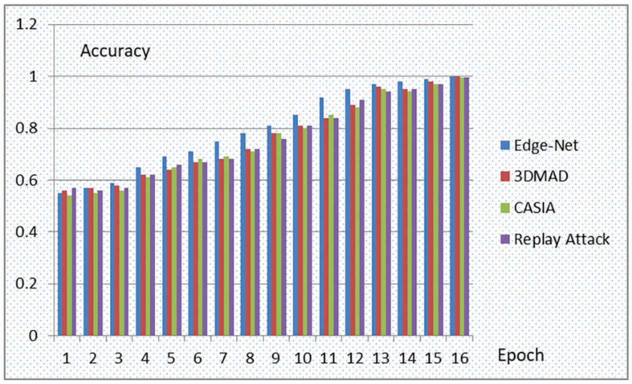

Figure 8: Accuracy level for 15 epochs

5 Results

We analyse the results obtained by the Edge-Net autoencoder on the Edge-Net datasets and the other three benchmark datasets with the results obtained using Convolutional autoencoders. Tabs. 2 and 3 depict the comparison based on performance measure like Accuracy, False Reject Rate (FRR), False Accept Rate (FAR), and Half Total Error Rate (HTER). Then the robustness of the Edge-Net autoencoder is validated using performance measure such as F1-score, Recall, and Precision. It is given in the Tabs. 4 and 5 as sample of 3DMAD and Edge-Net datasets.

Table 2: Results achieved using convolutional autoencoders

Dataset	Accuracy	HTER (%)	FRR (%)	FAR (%)
Edge-Net	99.03	0.045	00.032	3.85
3DMAD	100	0	0	0
CASIA	99.17	0.0176	0.0176	1.76
Replay attack	98.5	0.046	0.38	3.12

Table 3: Results achieved using proposed EdgeNet autoencoder

Dataset	Accuracy	HTER (%)	FRR (%)	FAR (%)
Edge-Net	100	0	0	0
3DMAD	100	0	0	0
CASIA	99.7	0.0112	0.0112	1.09
Replay attack	99.5	0.039	0.28	2.92

Table 4: Results obtained using proposed EdgeNet approach on the 3DMAD

Dataset	Precision	Recall	F1-score
Spoof image	1.0	1.0	1.0
Live image	1.0	1.0	1.0

Table 5: Results achieved using proposed EdgeNet architecture on the EdgeNet dataset

Dataset	Precision	Recall	F1-score
Spoof image	1.0	1.0	1.0
Live image	1.0	1.0	1.0

Tab. 3 encapsulate spoofed face detection results in dimensionality reduction using Edge-Net autoencoders and followed by feature extraction using fully connected layers with pre-trained encoder weights. According to this result, the proposed method performs well than other existing well-established techniques. The pre-trained weights and extra layers of encoder and decoder enhance the architecture's overall performance.

The accuracy level of the proposed EdgeNet with all four datasets is given in Fig. 9. According to the graph figure, the accuracy level reached above 99.5% for all the datasets, even for only 15 epochs. In Fig. 10. The accuracy level of EdgeNet for all four dataset sets is given. Fig. 11. Depicts the accuracy comparison between the proposed Edge-Net autoencoder and the existing model convolutional autoencoder. It shows EdgeNet produce 1% higher accuracy than the existing convolutional autoencoder model.

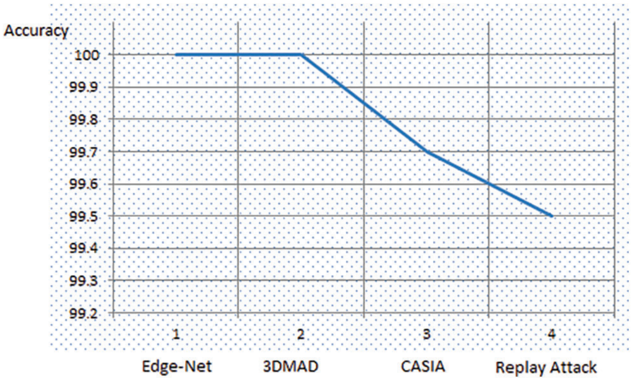

Figure 9: Accuracy level of Edge-Net architecture

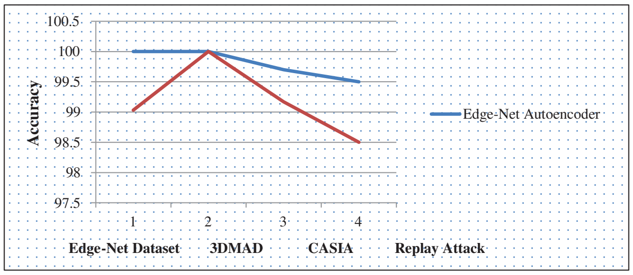

Figure 10: Accuracy comparison between Edge-Net and convolution autoencoder

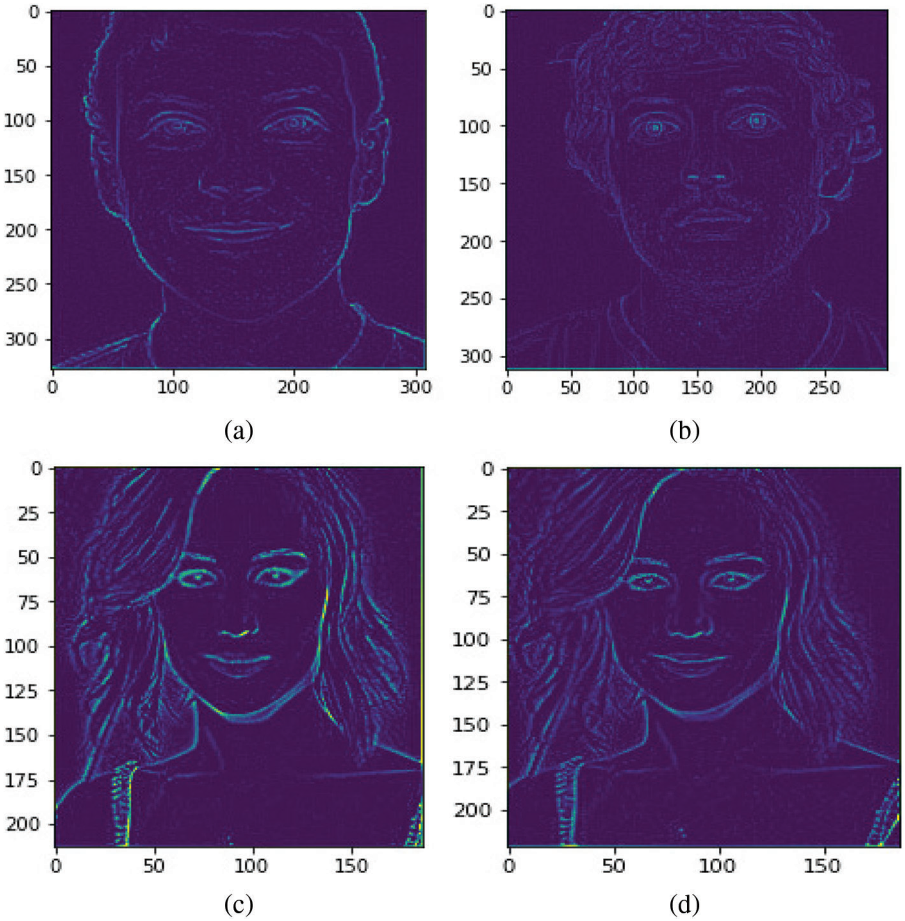

Figure 11: Output images-(a, c) spoofed images, (b, d) real images

Output Image

The output image of the Edge-Net architecture is given in Fig. 11. In the example output image, ‘a’ and ‘c’ images are spoofed images, ‘b’ and ‘d’ are the real images. The edges of the spoofed image are very sharp and darker than real images. Hence, this algorithm helps the machine identify and classify the spoofed and real images with the highest accuracy.

6 Conclusion and Future Work

The proposed EdgeNet model is helpful in the prediction of several sets of spoofing attacks in biometric systems which use faces for authentication. These attacks are photo attacks, 3D mask attacks, and replay attacks. This EdgeNet architecture has an extra layer of the encoder as well as decoder to reduce the dimensionality of the input images. It is followed by extraction of feature and classifies the images using pre-trained encoder weights and a SoftMax classifier. This proposed EdgeNet is more efficient and robust than the existing algorithm which is validated on three benchmark datasets by intra-dataset testing. The Edge-Net architecture provides the solution for overfitting. Performance of the architecture improved due to LoGS edge detection operator, pre-trained encoder weights, and extra encoder and decoder layers. The results are higher than most recent and advanced methods based on the performance measure, with accuracy higher than 99.5% in the case of intra-dataset testing. Thus, the Edge-Net architecture can predict the spoofed faces under numerous scenarios during authentication in a biometric system. In future the spoofing detection can be performed using machine learning classifiers for better accuracy.

Acknowledgement: Princess Nourah bint Abdulrahman University Researchers Supporting Project Number (PNURSP2022R120), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Funding Statement: Princess Nourah bint Abdulrahman University Researchers Supporting Project Number (PNURSP2022R120), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] J. Gan, S. Li, Y. Zhai and C. Liu, "Convolutional neural network based on face anti-spoofing," in *Proc. Int. Conf. on Multimedia and Image Processing (ICMIP)*, China, pp. 1–5, 2017.

[2] O. Nikisins, A. George and S. Marcel, "Domain adaptation in multi-channel autoencoder based features for robust face anti-spoofing," in *Proc. Int. Conf. on Biometrics (ICB)*, China, pp. 1–8, 2019.

[3] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv*, pp. 1409–1556, 2014.

[4] K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, India, pp. 770–778, 2016.

[5] G. Huang, Z. Liu, D. V. Maaten and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, India, pp. 4700–4708, 2017.

[6] X. Yang, W. Luo, L. Bao, Y. Gao, D. Gong *et al.,* "Face antispoofing: Model matters, so does data," in *Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition*, India, pp. 3507–3516, 2019.

[7] A. George and S. Marcel, "Deep pixelwise binary supervision for face presentation attack detection," in *Proc. Int. Conf. on Biometrics (ICB)*, India, pp. 1–8, 2019.

[8] X. Song, X. Zhao, L. Fang and T. Lin, "Discriminative representation combinations for accurate face spoofing detection," *Pattern Recognition*, vol. 85, pp. 220–231, 2019.

[9] A. Parkin and O. Grinchuk, "Recognizing multi-modal face spoofing with face recognition networks," in *Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition Workshops*, India, pp. 1–10, 2019.

[10] G. Heusch and S. Marcel, "Pulsebased features for face presentation attack detection," in *Proc. Int. Conf. on Biometrics Theory, Applications and Systems (BTAS)*, Chicago, USA, pp. 1–8, 2018.

[11] S. Fatemifar, S. R. Arashloo, M. Awais and J. Kittler, "Spoofing attack detection by anomaly detection," in *Proc. Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Brighton, UK, pp. 8464–8468, 2019.

[12] S. R. Arashloo, J. Kittler and W. Christmas, "An anomaly detection approach to face spoofing detection: A new formulation and evaluation protocol," *IEEE Access*, vol. 5, no. 4, pp. 13868–13882, 2017.

[13] O. Nikisins, A. Mohammadi, A. Anjos and S. Marcel, “On effectiveness of anomaly detection approaches against unseen presentation attacks in face anti-spoofing,” in *Proc. Int. Conf. on Biometrics (ICB)*, France, pp. 75–81, 2018.

[14] S. Fatemifar, M. Awais, S. R. Arashloo and J. Kittler, “Combining multiple one class classifiers for anomaly based face spoofing attack detection,” in *Int. Conf. on Biometrics (ICB)*, Crete, Greece, pp. 1–7, 2019.

[15] S. Fatemifar, M. Awais, A. Akbari and J. Kittler, “A stacking ensemble for anomaly based client specific face spoofing detection,” in *Int. Conf. on Image Processing (ICIP)*, Abu Dhabi, United Arab Emirates, pp. 1371–1375, 2020.

[16] X. Li, J. Wan, Y. Jin, A. Liu, G. Guo *et al.,* “3DPC-net: 3D point cloud network for face anti-spoofing,” in *IEEE Int. Joint Conf. on Biometrics (IJCB)*, USA, pp. 1–8, 2020.

[17] Z. Yu, C. Zhao, Z. Wang, Y. Qin, Z. Su *et al.,* “Searching central difference convolutional networks for face anti-spoofing,” in *Proc. Conf. on Computer Vision and Pattern Recognition*, Seattle, USA, pp. 5295–5305, 2020.

[18] S. Arora, M. P. S. Bhatia and V. Mittal, “A robust framework for spoofing detection in faces using deep learning,” *The Visual Computer*, vol. 12, no. 5, pp. 1–12, 2020.

[19] H. Chen, G. Hu, Z. Lei, Y. Chen, N. M. Robertson *et al.,* “Attention based two stream convolutional networks for face spoofing detection,” *IEEE Transactions on Information Forensics and Security*, vol. 15, no. 3, pp. 578–593, 2019.

[20] M. Khammari, “Robust face antispoofing using CNN with LBP and WLD,” *IET Image Processing*, vol. 13, no. 11, pp. 1880–1884, 2019.

[21] Y. Atoum, Y. Liu, A. Jourabloo and X. Liu, “Face anti spoofing using patch and depth based CNNs,” in *IEEE Int. Joint Conf. on Biometrics (IJCB)*, USA, pp. 319–328, 2017.

[22] S. Mahajan, A. Raina, M. Abouhawwash, X. Gao and A. K. Pandit, “COVID-19 detection from chest X-ray images using advanced deep learning techniques,” *Computers, Materials and Continua*, vol. 70, no. 1, pp. 1541–1556, 2022.

[23] M. Abouhawwash, “Hybrid evolutionary multi objective optimization algorithm for helping multi criterion decision makers,” *International Journal of Management Science and Engineering Management*, vol. 16, no. 2, pp. 94–106, 2021.

[24] V. Kandasamy, P. Trojovský, F. Machot, K. Kyamakya, N. Bacanin *et al.,* “Sentimental analysis of COVID-19 related messages in social networks by involving an N-gram stacked autoencoder integrated in an ensemble learning scheme,” *Sensors*, vol. 21, no. 22, pp. 7582, 2021.

[25] M. AbdelBasset, N. Moustafa, R. Mohamed, O. Elkomy and M. Abouhawwash, “Multiobjective task scheduling approach for fog computing,” *IEEE Access*, vol. 9, no. 3, pp. 126988–127009, 2021.

[26] M. Abouhawwash and A. Alessio, “Develop a multi objective evolutionary algorithm for pet image reconstruction: Concept,” *IEEE Transactions on Medical Imaging*, vol. 40, no. 8, pp. 2142–2151, 2021.

[27] K. Deb, M. Abouhawwash and J. Dutta, “Evolutionary multi-criterion optimization: 8th international conference,” in *EMO 2015, Proc., Part II, Springer Int. Publishing*, Guimarães, Portugal, Cham, pp. 18–33, 2015.

[28] A. Nayyar, S. Tanwar and M. Abouhawwash, “Emergence of the cyber-physical system and IoT in smart automation and robotics: Computer engineering in automation,” *Advances in Science, Technology and Innovation*, Germany: Springer, 2021.

[29] A. Garg, A. Parashar, D. Barman, S. Jain, D. Singhal *et al.,* “Autism spectrum disorder prediction by an explainable deep learning approach,” *Computers, Materials & Continua*, vol. 71, no. 1, pp. 1459–1471, 2022.

[30] M. Abouhawwash and K. Deb, “Karush-kuhn-tucker proximity measure for multi-objective optimization based on numerical gradients,” in *Proc. of the 2016 on Genetic and Evolutionary Computation Conf. Companion*, Denver Colorado USA, ACM, pp. 525–532, 2016.

[31] A. H. El-Bassiouny, M. Abouhawwash and H. S. Shahen, “New generalized extreme value distribution and its bivariate extension,” *International Journal of Computer Applications*, vol. 173, no. 3, pp. 1–10, 2017.

[32] A. H. El-Bassiouny, M. Abouhawwash and H. S. Shahen, “Inverted exponentiated gamma and its bivariate extension,” *International Journal of Computer Application*, vol. 3, no. 8, pp. 13–39, 2018.

[33] M. Abouhawwash and M. A. Jameel, “KKT proximity measure versus augmented achievement scalarization function,” *International Journal of Computer Applications*, vol. 182, no. 24, pp. 1–7, 2018.

[34] H. S. Shahen, A. H. El-Bassiouny and M. Abouhawwash, "Bivariate exponentiated modified weibull distribution," *Journal of Statistics Applications & Probability*, vol. 8, no. 1, pp. 27–39, 2019.

[35] M. Abouhawwash and M. A. Jameel, "Evolutionary multi-objective optimization using benson's karush-kuhn-tucker proximity measure," in *Int. Conf. on Evolutionary Multi-Criterion Optimization*, East Lansing, Michigan, USA, Springer, pp. 27–38, 2019.

[36] M. Abouhawwash, M. A. Jameel and K. Deb, "A smooth proximity measure for optimality in multi-objective optimization using benson's method," *Computers & Operations Research*, vol. 117, no. 2, pp. 104900, 2020.

[37] M. Abouhawwash, K. Deb and A. Alessio, "Exploration of multi-objective optimization with genetic algorithms for PET image reconstruction," *Journal of Nuclear Medicine*, vol. 61, no. 1, pp. 572–572, 2020.

[38] W. Wang, X. Huang, J. Li, P. Zhang and X. Wang, "Detecting COVID-19 patients in X-ray images based on MAI-nets," *International Journal of Computational Intelligence Systems*, vol. 14, no. 1, pp. 1607–1616, 2021.

[39] Y. Gui and G. Zeng, "Joint learning of visual and spatial features for edit propagation from a single image," *The Visual Computer*, vol. 36, no. 3, pp. 469–482, 2020.

[40] X. R. Zhang, X. Chen, W. Sun and X. Z. He, "Vehicle reidentification model based on optimized densenet121 with joint loss," *Computers, Materials & Continua*, vol. 67, no. 3, pp. 3933–3948, 2021.

[41] W. Sun, G. C. Zhang, X. R. Zhang, X. Zhang and N. N. Ge, "Fine grained vehicle type classification using lightweight convolutional neural network with feature optimization and joint learning strategy," *Multimedia Tools and Applications*, vol. 80, no. 20, pp. 30803–30816, 2021.