ISSN: 0975-833X

Available online at http://www.journalcra.com

International Journal of Current Research

Vol. 16, Issue, 10, pp.*30216-30222*, October, 2024
DOI: https://doi.org/10.24941/ijcr.47899.10.2024

INTERNATIONAL JOURNAL
OF CURRENT RESEARCH

RESEARCH ARTICLE

ENHANCED HUMAN FACIAL ANTI-SPOOFING THROUGH TEXTURE ANALYSIS LEVERAGING CONVOLUTIONAL NEURAL NETWORKS

Logeswari Saranya, R. and Umamaheswari, K.

[1]Research Scholar, Departement of Information Technology, PSG College of Technology, Coimbatore, Tamil Nadu, India; [2]Professor, Departement of Information Technology, PSG College of Technology, Coimbatore, Tamil Nadu, India

ARTICLE INFO

Article History:
Received 24th July, 2024
Received in revised form
17th August, 2024
Accepted 29th September, 2024
Published online 30th October, 2024

Key Words
Face Anti-Spoofing, Texture Analysis, Local Binary Pattern, Convolution Neural Network, VGG Architecture, Mobilenet Architecture.

****Corresponding author:***
Logeswari Saranya, R.

ABSTRACT

Face recognition is a widely utilized authentication method in biometric systems. However, a critical challenge faced by these systems is the occurrence of false acceptance, where unauthorized individuals or attackers gain access. This type of security breach, known as a presentation attack, exposes the system to significant vulnerabilities. The focus of this research is on a specific type of presentation attack called spoofing, where attackers employ artifacts such as photographs, masks, or pre-recorded videos to replicate the appearance of legitimate users and gain fraudulent access. The primary objective of this study is to develop a neural network model capable of accurately detecting whether a user is genuine or a spoof, thereby mitigating the risk of biometric system compromise.

Citation: Logeswari Saranya R.and Umamaheswari, K. 2024. "Enhanced Human Facial Anti-Spoofing through Texture Analysis Leveraging Convolutional Neural Networks". *International Journal of Current Research*, 16, (10), 30216-30222.

INTRODUCTION

Face anti-spoofing (FAS) has lately attracted increasing attention due to its vital role in securing face recognition systems from presentation attacks (PAs). As more and more realistic PAs with novel types spring up, traditional FAS methods based on handcrafted features become unreliable due to their limited representation capacity. Face spoofing attack is a process in whicha fraudulent user can attack face recognition system by impersonating a registered user and thereby gaining unauthorized access and advantages. Spoofing is often done by employing a photo, video, or a special substitute for a licensed person's face. Face recognition systems are vulnerable to photographs called print attacks in which the image is printed or displayed on a digital device. This is the foremost common sort of attack since most people have facial pictures available on the web or photos might be obtained without permission. The video spoof attacks may be more sophisticated to trick the system, which usually requires a looped video of a victim's face. In addition to natural facial movements, it enables ways to deceive some extra layers of protection such as depth Adding a spoofing attack detection module to the existing face recognition systems can be the solution to this problem. In this work, LBP-SVM Classifier based approach, CNN-based machine learning approach with Visual Geometry Group (VGG) architecture, and MobileNet architecture has been implemented todetect the face spoofing attack. The features of the face images are extracted using texture analysis since printed faces produce certain texture patterns that do not exist in real ones. The extracted features are then used to train the model to detect the spoofing attack.

Motivation: Face anti-spoofing is the task of preventing false facial verification by using a photo, video, or mask for an authorized person's face. There are a few attacks like the Replay attack, print attack, and 3D mask attack which result in the vulnerability of the biometric system. Our objective is to identify whether the recognized face images are real or spoofed using neural networks and effective machine learning algorithms.

LITERATURE SURVEY

Youngjun Moon proposed a face anti-spoofing method using color texture segmentation on FPGA in which RGB face images are converted into the YCbCr and HSV color spaces, and the spoofing images are classified by applying an LBP to each color space. This proposed scheme uses less memory with fewer feature dimensions, thus enabling high-speed processing. (1). Rui Shao, Xiangyan Lan, Jiawei Li, and Pong C proposed a technique called Multi-adversarial Discriminative Deep Domain Generalization in which one feature generator is trained to compete with multiple domain discriminators simultaneously so that the generalized feature space can be automatically and adaptively learned. Dual-force Triplet-mining Constraint is used in print and video relay attacks, the negative is likely to be more similar than the positive for each subject in both intra and cross- domains. This constraint attempts to solve this problem by minimizing intra-class distance while maximizing inter-class distance in both intra and cross-domains. (2)

An RCTR-based face presentation attack detection method is being adopted by Yuting Du and Tong Qiao, in which the residual image is transformed from RGB into another color space (e.g., YCbCr). Subsequently, texture descriptors are applied to extract rich texture information, in which a comprehensive representation is constructed by combining optimal descriptor feature vectors, namely, RCTR. An ensemble classifier with the effective strategy of probabilistic voting decision is designed which can complete the task of face presentation attack detection (3). Zitong Yu and Xiaobai Li introduced a method called Bilateral Convolutional Networks (BCN) that utilizes bilateral filtering. The bilateral filtered frames are taken as network inputs instead of the original RGB frames. The bilateral filter is utilized to smooth the original frame while preserving its main edges. BCN is mainly adopted here to integrate traditional Bilateral filtering with deep networks properly (4). The different methods adopted for face anti-spoofing have been reviewed where methods based on image quality, depthinformation, and feature integration are present. With the extensive application of deep learning in face anti-spoofing, more methods have been proposed. However, these methods are usually limited to the detection of known spoofing attacks, and there is blindness to unknown spoofing. To improve the generalization ability of detection methods under "invisible" attacks, methods based on Domain Generalization has been developed (5).

The model generated by Ying Huang, Wenwei Zhang, and Jinzhuo Wang used frequent spatial temporal architecture in which the main aim is to guide the deep network to focus on the obvious spoof patterns across a different domain, automatically make a trade-off in selecting the high level semantic deep features and low-level frequent expression. Here multi-frame images are fed as input to the network. A technique called the Fourier transform is employed to get the spectral image, which keeps the whole input information but expresses it differently (6). Zitong Yu and Xiaobai Li proposed a novel pyramid supervision, which guides deep models to learn both local details and global semantics from a multi-scale spatial context. Extensive experiments are performed on five FAS benchmark datasets to show that, without bells and whistles, the proposed pyramid supervision could not only improve the performance beyond existing pixel-wise supervision frameworks but also enhance the model's interpretability (i.e., locating the patch-level positions of PAs more reasonably). Furthermore, elaborate studies are conducted for exploring the efficacy of different architecture configurations with two kinds of pixel-wise supervisions (binary mask and depth map supervisions), which provides inspirable insights for future architecture/supervision design (7). Mingxin Liu and Jiong Mu proposed a framework to improve the generalization ability of face anti-spoofing in two folds: a generalized feature space is obtained via aggregation of all live faces while dispersing each domain's spoof faces, and a domain agnostic classifier is trained through low-rank decomposition.

Specifically, a Common Specific Decomposition for a Specific (CSD-S) layer is deployed in the last layer of the network to select common features while discarding domain-specific ones among multiple source domains. The two mentioned components are integrated into an end-to-end framework, ensuring the generalization ability to unseen scenarios. The extensive experiments demonstrate that the proposed method achieves state-of-the-art results on four public datasets, including CASIA-MFSD, MSU-MFSD, Replay-Attack, and OULU-NPU. (8)

Inference

From the above study, it is analyzed that the main objective of any ML model is determined by how well it can generalize on unseen data as opposed to how well it performs on the training data. It is preferred to have different model adaptations for trainingand inference purposes. During training, deep and complex modelsare needed to train upon a large amount of training data but during inference, we just need a lighter model that generalizes well on any unseen data. The lighter model has a good performance during inference time in production. So the context setting for this article is to see if there is an effective way to distill this generalized knowledge to a lighter model to get the best of both worlds. The proposed Convolutional Neural Network (CNN) approach with Visual Geometry Group (VGG) and Local Binary Pattern with SVM classifier for the former case and MobileNet for the latter will improve the performance of face anti-spoofing detection.

PROPOSED SYSTEM

Face spoofing can be detected by major three methodologies - liveness based, texture-based, and 3D geometric- based. The liveness-based method requires lots of video samples with a longer duration to train the model effectively. The 3D geometric-based method is not effective for mask attacks. Hence, this work uses the texture-based method. As the texture or depth of the real and spoof is different, building the model using these texture features proved to give better results.

Proposed Work Flow: The various phases in face spoofing detection are represented in Fig 1. Image frames are captured from the video dataset (CASIA-FASD). The face is detected from the images using the Haar-cascade face detection algorithm. The detected face is then cropped into image frames and all the preprocessing is done at this stage. The pre-processed image is fed into a spoofing detection model. If the final output value is >0.5 the subject is identified to be a real image, else it is considered a spoof.

Steps in Preprocessing

Median Filter: Median filtering is a nonlinear process useful in reducing impulsive, or salt-and-pepper noise. The median filter is also used to preserve edge properties while reducing the noise. Smoothing techniques, like Gaussian blur, are also used to reduce noise but they can't preserve the edge properties. The median filter is widely used in digital image processing just because it preservesedge properties.

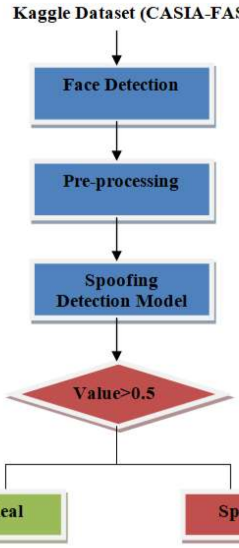


Fig 1 Proposed Work Flow for Face Anti-Spoofing

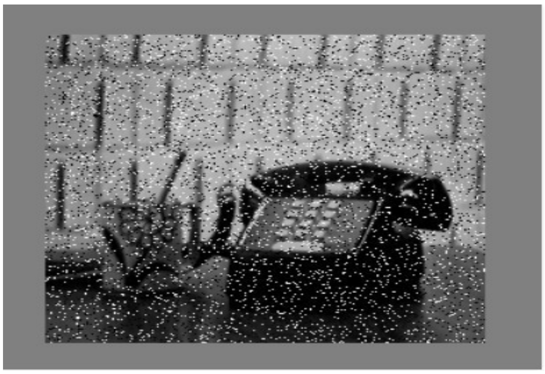
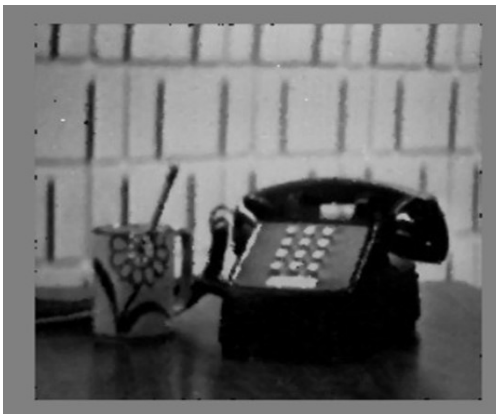


Fig 2 Before Applying Median Filter



Fig, 3. After Applying Median Filter

Work Flow of Median Filter:

- Store the pixel values of the input image in an array.
- For each pixel value store all the neighboring pixel values including that cell in a new array (called a window).
- Sort the window array.
- The median of the window array is used to store outputimage pixel intensity.
- If N is an even number then the filter calculates an arithmetic mean between the two central elements of the sorted array

$$MedianFilter = a\left(\frac{N+1}{2}\right)$$

- If N is an odd number then the filter takes a centralelement of the sorted array

$$MedianFilter = \frac{a\left(\frac{N}{2}\right)+a\left(\frac{N}{2}+1\right)}{2}.$$

Texture Analysis: Texture analysis refers to the characterization of regions in an image by their texture content. Texture analysis attempts to quantify intuitive qualities described by terms such as rough, smooth, silky, or bumpy as a function of the spatial variation in pixel intensities. In this sense, the roughness or bumpiness refers to variations in the intensity values, or gray levels. Texture analysis can be used to find the texture boundaries, called texture segmentation. Texture analysis can be helpful when objects in an image are more characterized by their texture than by intensity, and traditional thresholding techniques cannot be used effectively.

Local Binary Pattern: Local Binary Pattern (LBP) is a simple yet very efficient texture operator which labels the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number. Due to its discriminative power and computational simplicity, the LBP texture operator has become a popular approach in various applications. It can be seen as a unifying approach to the traditionally divergent statistical and structural models of texture analysis. Perhaps the most important property of the LBP operator in real-world applications is its robustness to monotonic gray-scale changes caused, for example, by illumination variations. Another important property is its computational simplicity, which makes it possible to analyze images in challenging real-time settings. In the LBP approach for texture classification, the occurrences of the LBP codes in an image are collected into a histogram. The classification is then performed by computing simple histogram similarities. However, considering a similar approach for facial image representation results in a loss of spatial information, and therefore one should codify the texture information while retaining also their locations. One way to achieve this goal is to use the LBP texture descriptors to build several local descriptions of the face and combine them into a global description. Such local descriptions have been gaining interest lately which is understandable given the limitations of the holistic representations. These local feature-based methods are more robust against variations in pose or illumination than holistic methods. This histogram effectively has a description of the face on three different levels of locality: the LBP labels for the histogram contain information about the patterns on a pixel level, the labels are summed over a small region to produce information on a regional level and the regional histograms are concatenated to build a global description of the face as shown in Fig 4.

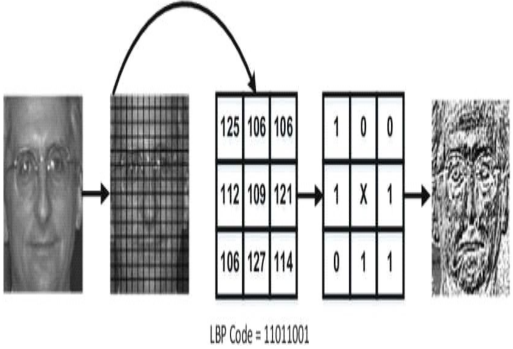


Fig. 4. Work Flow of LBP

A support vector machine (SVM) is a type of deep learning algorithm that performs supervised learning for the classification or regression of data groups. In AI and machine learning, supervised learning systems provide both input and desired output data, which are labeled for classification.

The advantages of support vector machines are

- Effective in high-dimensional spaces.
- Still effective in cases where the number of dimensions is greater than the number of samples.

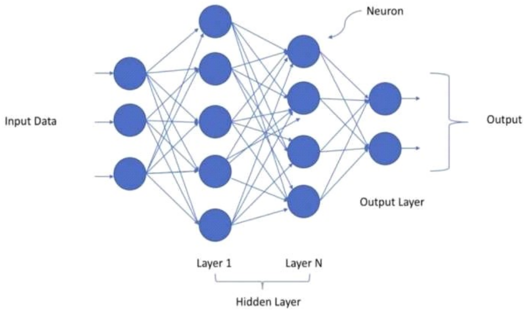


Fig 5 Structure of Neural Network Topology

- Uses a subset of training points in the decision function (called support vectors), so it is also memory efficient.
- Versatile: different Kernel functions can be specified for the decision function. Common kernels are provided, but it is also possible to specify custom kernels.

The disadvantages of support vector machines include:

- If the number of features is much greater than the number of samples, overfitting occurs.
- SVMs do not directly provide probability estimates, these are calculated using an expensive five-fold cross-validation.

Convolutional Neural Networks: Deep learning is an artificial intelligence function that imitates the working of the human brain in processing data and creating patterns for use in decision-making. Deep learning is a subset of machine learning in artificial intelligence that has networks capable of learning unsupervised from data that is unstructured or unlabeled. Deep learning requires large amounts of labeled data and substantial computing power. Deep learning models can achieve better accuracy, sometimes exceeding human- level performance. Models are trained by using a large set of labeled data and neural network architectures that contain many layers. The computer model learns to perform classification tasks directly from images, text, or sound. Neural networks are statistical models based on biological neural networks. They are capable of modeling and processing nonlinear relationships between inputs and outputs in parallel. Neural networks are characterized by containing adaptive weights along paths between neurons that can be tuned by a learning algorithm that learns from observed data to improve the model. The cost function is used to learn the optimal solution to the problem being solved. A neural network is modeled using layers of artificial neurons, or computational units able to receive input and apply an activation function along with a threshold to determine if messages are passed along. The first layer is the input layer, followed by one or more hidden layers, and the last is an output layer as shown in Fig 5. Each layer can contain one or more neurons. Neural networks have many learning algorithms. For the face mask detector model the learning algorithm is Convolutional Neural Networks.

A convolutional neural network is a class of deep neural networks, most commonly applied to analyzing visual imagery. Convolutional networks use convolution in place of matrix multiplication in at least one of their layers. A convolutional neural network consists of an input and an output layer and multiple hidden layers. The hidden layers of a CNN consist of a series of convolutional layers that convolve with multiplication. The activation function is commonly a RELU layer and is subsequently followed by additional convolutions, pooling layers, fully connected layers, and hidden layers where inputs and outputs are masked by the activation function and final convolution.

Proposed VGG Architecture: VGG is a linear model and is one of the most used image-recognition architectures. VGG is an innovative object-recognition model that supports up to 19 layers. Built as a deep CNN, VGG also outperforms baselines on many tasks and datasets outside of ImageNet.

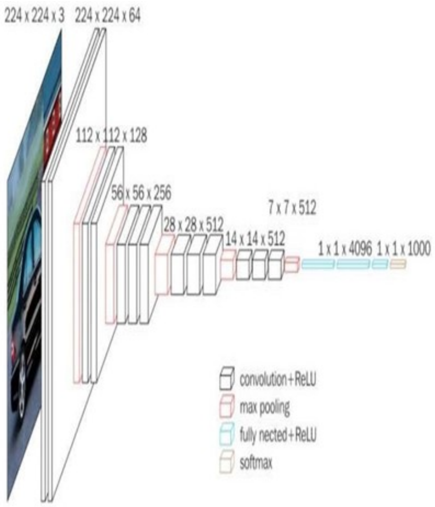


Fig 6. Proposed Structure of VGG Architecture

This network is characterized by its simplicity, using only 3×3 convolutional layers stacked on top of each other in increasing depth as shown in Fig. 6.

- The input to the first convolutional layer is of fixed size 224 x 224 RGB image.
- The image is passed through a stack of convolutional layers, where the filters were used with a very small receptive field: 3×3 (which is the smallest size to capture the notion of left/right, up/down, and center). One of the configurations also utilizes 1×1 convolution filters, which can be seen as a linear transformation of the input channels (followed by non-linearity).
- The convolution stride is fixed to 1 pixel. The spatial padding of convolutional layer input is such that the spatial resolution is preserved after convolution, i.e. the padding is 1 pixelfor 3×3 convolutional layers.
- Spatial pooling is carried out by fivemax-pooling layers, which follow some of the convolutional layers(not all the convolutional layers are followed by max-pooling).
- Max-pooling is performed over a 2×2 pixel window, with stride 2.
- Three Fully-Connected (FC) layers follow a stack of convolutional layers (which has different depth in different architectures).
- The final layer is the soft-max layer. The configuration of the fully connected layers is the same in all networks. All hidden layers are equipped with the rectification (ReLU) non-linearity.

There are two major drawbacks to VGGNet:

- It is too slow to train.
- The network architecture weights themselves are quitelarge (in terms of disk/bandwidth).

Due to its depth and number of fully-connected nodes, VGG requires more memory. This makes deploying VGG a tiresome task. VGG is used in many deep-learning image classification problems. However, smaller network architectures are often more desirable.

Proposed MobileNet Architecture: CNNs are becoming deeper and increasingly complex. So to achieve a higher degree of accuracy Mobilenet architecture is chosen. MobileNet is an efficient and portable CNN architecturethat is used in real-world applications. MobileNets primarily use depthwise separable convolutions as shown in Fig 7 in place of the standard convolutions used in earlier architectures to build lighter models.

- MobileNets introduce two new global hyperparameters (width multiplier and resolution multiplier) that allow model developers to trade off latency or accuracy for speed and low size depending on their requirements.
- MobileNets are built on depth-wise separable convolution layers. Each depth-wise separable convolution layer consists of a depthwise convolution and a pointwise convolution.
- Counting depthwise and pointwise convolutions as separate layers,a MobileNet has 28 layers.
- A standard MobileNet has 4.2 million parameters which can be further reduced by tuning the width multiplier hyperparameter appropriately.

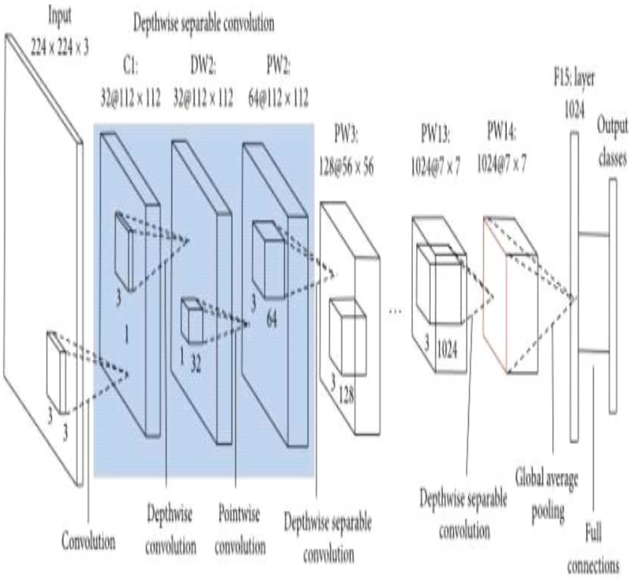


Fig. 7. Proposed Structure of MobileNet Architecture

IMPLEMENTATION AND RESULTANALYSIS

The neural network models for face anti-spoofing are built using the architectures mentioned in the above section. These models are then trained with the preprocessed dataset for about 20 epochs using the computing service provided by the GPU of the Google Colab.

Data Collection: The dataset used in the project is CASIA Face Anti- Spoofing Database (CASIA-FASD). The data consists of real videos and spoof videos. It contains 50 subjects, and 12 videos for each subject under different resolutions and light conditions. This dataset is designed for three different spoof attacks: replay, warp print, and cut print attacks. It contains 600 video recordings, and each video ranges between 5 to 15 seconds.

Data Preprocessing: The real and spoofed videos are split into separate folders. The image frames are extracted from these videos at the rate of 30 frames per second. The new dataset consists of real and spoof images approximately 5000 images for each. From these images, the face of the person is alone cropped using the haar cascade algorithm. Then the images are resized into 224x224 pixels. The salt and pepper noise is removed using the median filter and images are stored in new folders. The processed dataset is split into testing and training sets. The split data is converted into NumPy arrays with the labels 0 and 1, 0 for real, and 1 for thespoof.

Model Building

LBP Model: The facial texture features are extracted from the image using the LBP method described in section 3.1.3. Extracted features are then fed to the SVM classifier. It is a supervised classification algorithm where it draws a hyperplane between two different categories to differentiate between them. The real features are labeled as 0 and spoof images are labeled as 1. The SVM classifier returns the value 0 if it is real, or the value 1 if it isspoofed.

VGG Model: CNN VGG model is built with the layers described in section 3.1.4.1 with the help of the Keras library. The first layer is the input layer which takes the 224x224 image as input. Followed by the convolutions, pooling, and activation layers. The output layer with sigmoid activation returns a single value. If the value is greater than 0.5 then the image is real else it is spoofed.

MobileNet Model: In the MobileNet model, depthwise convolution is applied to the input image to form the channel-wise spatial convolution. Pointwise convolution is then applied to change the dimension of the image. The sigmoid activation for the output image returns 2 values i.e., the probability of the real image and the probability of the spoof image. By using argmax, the class with maximum probability is identified as output.

Training Phase: Around 19,000 images have been preprocessed and trained for prediction and a model is developed using the Convolutional Neural Networks-Deep learning algorithm toclassify whether the image is real or fake.

Analysis of Parameter Metrics

Learning Rate: The learning rate controls how quickly or slowly a neural network model learns a problem. If the initial learning rate is minimum, the loss will be also minimum. For the face-anti spoofing model, the initial learning rate value is set to 0.001.

Epochs: The number of epochs defines the number of times thatthe learning algorithm will work through the entire training dataset. For the face-anti spoofing model, the epoch is set to 20.

Batch Size: The batch size defines the number of samples to work through before updating the internal model parameters. A training dataset can be divided into one or more batches. For the face-anti spoofing model, the batch value is set to 32.

Testing Phase: After the model is created, testing is done by passing the data into the model. Using OpenCV, a live web camera is turned on. The image frames are captured from the video and are sent to the model as an input image. The model is then able to classify the image as real or spoofed.

RESULT ANALYSIS

The performances of the VGG, LBP-SVM, and MobileNet architectures are measured in terms of four evaluation metrics: accuracy, F-score, recall, and precision.

Table 1. Performance Evaluation of the Models Implemented

EVALUATION METRICS (%)	Performance		
	VGG Architecture	LBP-SVM Architecture	Proposed Mobilenet Architecture
Accuracy	99	97.75	99.75
Precision	98	99	99
Recall	100	97	100
F1 Score	99	98	100

The dataset was split as training to testing datasets in the ratio 80:20. The accuracy, recall, precision, and F-score values of VGG, LBP-SVM, and MobileNet Architectures are listed in Table 1, respectively. The CASIA-FASD dataset is used to implement three architectures. MobileNet gives better accuracy among all three models implemented. The accuracy is 99.75%.

CONCLUSION AND FUTURE ENHANCEMENT

The proposed texture analysis method is used to solve face spoofing detection effectively. The texture features were extracted from individual image channels to differentiate the features of real and spoof images. The proposed approach showed very promising generalization capabilities. The face anti-spoofing models developed using LBP-SVM, VGG architecture, and MobileNet architecture proved to perform better with an accuracy above 97%. Among these models, MobileNet has a greater accuracy of 99.75%. In the future, these models can be implemented and tested for different datasets like Msspoof (Multispectral spoof), print attacks, and replay attacks.

REFERENCES

1. Moon, Y., Ryoo, I., & Kim, S. (2021). Face antispoofing method using color texture segmentation on FPGA. Security and Communication Networks, 2021. https://doi.org/10.1155/2021/9939232
2. Shao, R. (2019). Xiangyuan Lan, Jiawei Li, and Pong C Yuen. Multi-adversarial discriminative deep domain generalization for face presentation attack detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 10023-10031). **DOI:** 10.1109/CVPR.2019.01026
3. Du, Y., Qiao, T., Xu, M., & Zheng, N. (2021). Towards Face Presentation Attack Detection Based on Residual Color Texture Representation. Security and Communication Networks, 2021. DOI:10.1155/2021/6652727
4. Yu, Z., Li, X., Niu, X., Shi, J., & Zhao, G. (2020, August). Face anti-spoofing with human material perception. In European Conference on Computer Vision (pp. 557-575). Springer, Cham. https://doi.org/10.1007/978-3-030-58571-6_33
5. Ming, Z., Visani, M., Luqman, M. M., & Burie, J. C. (2020). A survey on anti-spoofing methods for facial recognition with RGB cameras of generic consumer devices. Journal of Imaging, 6(12), 139. https://doi.org/10.3390/jimaging6120139
6. Huang, Y., Zhang, W., & Wang, J. (2020). Deep frequent spatial temporal learning for face anti- spoofing. arXiv preprint arXiv:2002.03723. https://doi.org/10.48550/arXiv.2002.03723
7. Yu, Z., Li, X., Shi, J., Xia, Z., & Zhao, G. (2021). Revisiting pixel-wise supervision for face anti- spoofing. IEEE Transactions on Biometrics, Behavior, and Identity Science.
8. Liu, M., Mu, J., Yu, Z., Ruan, K., Shu, B., & Yang, J. (2022). Adversarial learning and decomposition-based domain generalization for face anti-spoofing. Pattern Recognition Letters, 155, 171-177.

https://doi.org/10.1016/j.patrec.2021.10.014Get rights and content

9. Wen, D., Han, H., & Jain, A. K. (2015). Face spoof detection with image distortion analysis. IEEE Transactions on Information Forensics and Security, 10(4), 746-761. **DOI:** 10.1109/TIFS.2015.2400395
10. Cai, T., Chen, F., Liu, W., Xie, X., & Liu, Z. (2022). Face anti-spoofing via conditional adversarial domain generalization. Journal of Ambient Intelligence and Humanized Computing, 1-14. https://doi.org/10.1007/s12652-022-03884-z ·
11. Chang, H. H., & Yeh, C. H. (2022). Face anti-spoofing detection based on multi-scale image quality assessment. Image and Vision Computing, 121, 104428. DOI:10.1016/j.imavis.2022.104428
12. Li, L., Xia, Z., Wu, J., Yang, L., & Han, H. (2022). Face presentation attack detection based on optical flow and texture analysis. Journal of King Saud University-Computer and Information Sciences, 34(4), 1455-1467. https://doi.org/10.1016/j.jksuci.2022.02.019
13. Boulkenafet, Z., Komulainen, J., & Hadid, A. (2015, September). Face anti-spoofing based on color texture analysis. In 2015 IEEE international conference on image processing (ICIP). **DOI:** 10.1109/ICIP.2015.7351280
14. Liu, S., Zhang, K. Y., Yao, T., Bi, M., Ding, S., Li, J.,... & Ma, L. (2021, October). Adaptive normalized representation learning for generalizable face anti-spoofing. In Proceedings of the 29th ACM International Conference on Multimedia (pp. 1469- 1477). DOI:10.48550/arXiv.2108.02667
15. Binti Ashari, N. N., Ong, T. S., Connie, T., Teng, J. H., & Leong, Y. F. (2021, September). Multi-Scale Texture Analysis For Finger Vein Anti-Spoofing. In 2021 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET) (pp. 1-6). IEEE. DOI:10.1109/IICAIET51634.2021.9574036
16. Rathgeb. C,Drozdowski.P, Fisher.D and Busch.C.,2020, “Vulnerability Assessment and Detection of Makeup Presentation Attacks”, IEEE: 8th International Workshop on Biometrics and Forensics (IWBF), June 2020. DOI:10.1109/IWBF49977.2020.9107961
17. Yuting Du, Tong Qiao., “Towards Face Presentation Attack Detection Based on Residual Color Texture Representation”, Journal of Security and Communication Networks, Article ID 6652727, March 2021. https://doi.org/10.1155/2021/6652727
18. Ketan Kotwal, Zohreh Mostaani, and Sebastien Marcel., ”Detection of Age-Induced Makeup Attacks on Face Recognition Systems Using Multi-Layer Deep Features", IEEE Transactions on Biometrics, Behavior, and Identity Science, Volume: 2, Jan 2020. DOI:10.1016/j.jksuci.2022.02.019
19. Rathgeb. C, Drozdowski. P and Busch.C., "Detection of Makeup Presentation Attacks based on Deep Face Representations," 2020 25th International Conference on Pattern Recognition (ICPR), 2021. **DOI:** 10.1109/ICPR48806.2021.9413347
20. Zhaoyang Sun, Feng Liu, Wen Liu, Shengwu Xiong, Wenxuan Liu.,2020, “Local Facial Makeup Transfer via Disentangled Representation”, Computer Vision – ACCV 2020, 2020. https://doi.org/10.1007/978-3-030-69538-5_28
21. Puspita Majumdar, Akshay Agarwal, Mayanak Vatsa, Richa Singh., "Facial Retouching and Alteration Detection", Springer 2021.
22. Karmakar D., Mukherjee P. & Datta M., “Spoofed Facial Presentation Attack Detection by Multivariate Gradient Descriptor in Micro-Expression Region”, Pattern Recognition and Image Analysis 31, June 2021.
23. Kips R., Gori P., Perrot M., Bloch I., CA-GAN., “Weakly Supervised Color Aware GAN for Controllable Makeup Transfer”, Computer Vision – ECCV 2020.
24. Anchieta, N.M., Mafra, A.L., Hokama, R.T., "Makeup and Its Application Simulation Affect Women’s Self-Perceptions”, 2021. DOI: 10.1007/s10508-021-02127-0
25. Yu Z., Li X., Niu X., Shin JZhao G., “Face Anti-Spoofing with Human Material Perception”, Computer Vision – ECCV 2020. https://doi.org/10.3390/electronics12102199

****