www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 04, Issue 11, November 2024, pp : 2540-2547

e-ISSN : 2583-1062
Impact Factor : 7.001

FACE LIVENESS DETECTION IN BIOMETRIC SYSTEMS

Ratti Pallivi[1], Dr. Sreejyothsna Ankam[2]

[1,2]GMR Institute of technology, Rajam

ABSTRACT

Face biometric systems are now widely employed across governments, commercial sectors, and social media platforms, but this widespread adoption has made them susceptible to various attacks, including face spoofing, Distributed Denial of Service (DDoS), and phishing. These threats pose significant security challenges, underscoring the urgent need for more robust protective measures to safeguard these systems. The widespread adoption also comes with serious security concerns and underlines the necessity for advanced protective measures. In this regard, the paper proposes an enhanced face spoofing detection framework, which may effectively cover these vulnerabilities with improved accuracy. The proposed framework builds on deep learning advancements and integrates pre-trained convolutional autoencoders for feature extraction and dimensionality reduction, followed by a softmax classifier for classification. It is also further strengthened by advanced architectures such as ResNet, which uses deep residual learning, with Convolutional Block Attention Modules to enhance feature focus. Besides, the hybrid models combining CNNs with LSTM networks are used to extract both spatial and temporal features, which are very essential in detecting spoofing either in images or videos. Comprehensive experiments on benchmark datasets like Idiap Replay Attack, CASIA-FASD, and 3DMAD demonstrate that an extended framework can at least be as good as, and often better than, state-of-the-art methods, ensuring better generalization across various spoofing scenarios. This fundamentally increases the security and reliability of biometric systems through a view to offer a robust solution against sophisticated spoofing attacks.

Keywords: Face Spoof Detection, Deep Learning, Convolutional Autoencoders, Pre trained Weights

1. INTRODUCTION

Face spoofing attacks are a serious threat against facial recognition systems. The increase in the implementation of facial recognition systems across financial transactions, verification purposes, and access controls underlines the use of such a system as a biometric authenticating agent. However, with such wide application, these systems have also become a significant target for spoofing attacks. Methods to defeat facial recognition algorithms exist and include printed photos, video replays, or 3D masks, which then grant the attacking agents the ability to counter these applications at sensitive points. An attacker seeks to pretend to be an authorized user by posing with photos, videos, or even 3D masks. Available techniques address the problem of spoofing to some extent but are inadequate to handle complex and various attack scenarios .

These attacks are increasingly sophisticated, breaking the boundaries of what regular liveness detection can cope with, since traditional liveness detection is reliant on simplistic cues or restrictive datasets. Eliminating these vulnerabilities is what will seal the gaps and build confidence in biometric systems while maintaining authentication strength in high-stakes environments.

Fig-1: Face Spoofing

Deep learning has been the powerhouse tool against spoofing attacks. Convolutional neural networks have so far achieved remarkable success in extracting detailed spatial features from images, capturing subtle differences between genuine and spoofed faces. Hybrid models that combine CNNs with recurrent architectures such as LSTMs also enable the analysis of temporal dynamics in video sequences, which can be extremely important in order to detect replay attacks. More contemporary breakthroughs like the attention mechanisms and densely connected networks extend feature extraction through focussed discriminative regions and the optimization in features reuse. These developments would be auspicious to more accurate and efficient detection frameworks of liveness. However, despite such breakthroughs, generalization across diverse datasets and real-world conditions poses significant challenges, especially lighting, angles, and spoofing types. Lightweight architectures also play an important role for resource-constrained environments such as mobile devices where a high efficiency of computation is particularly crucial. This paper applies a convolutional autoencoder for dimensionality reduction in that redundant features are discarded and enhances the capacity of generalization within the framework. It effectively obtains features and classifies the reconstructed facial

images by means of pre-trained encoder weights, which leads to improved performance. Introducing these features with the proposed architecture of a convolutional neural network, those features can automatically be learned to distinguish live faces from spoofed ones. The softmax classifier is utilized to classify into real or spoofed categories. Besides, it has streamlined the processing of high-dimensional data and ensured higher accuracy in face liveness detection. With this framework using pre-trained weights and getting the autoencoder to reconstruct meaningful facial features, it gains great robustness and efficiency to make it perfect for real-world applications requiring reliable detection of face liveness in biometric systems.Results obtained on three popular benchmarks i.e Idiap Replay Attack, CASIA-FASD and 3DMAD are comparable to many state-of-the-art methods used for detection of spoofed faces.

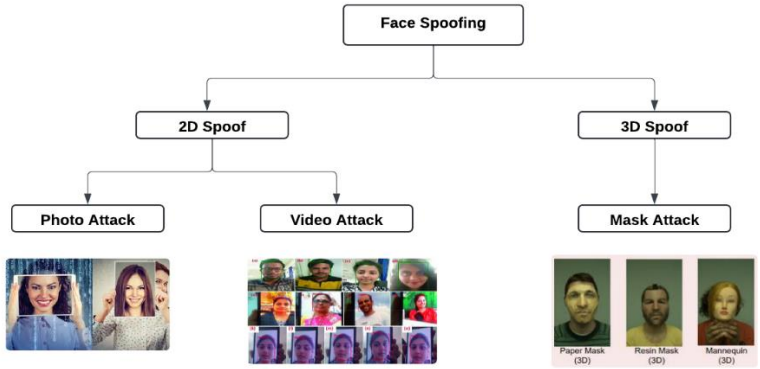

Fig. 2 Types of Attacks in face spoofing

2. RELATED WORK

Zuo, Gao, and Wang [1] designed a light-weight and efficient deep network for face liveness detection. They use activity-based indicators such as blinks, mouth movements, and head gestures against static image attacks. Similarly, facial structure tensors combined with gradient-based cues were used by Koll Reider et al. [7] to differentiate between live faces and spoofs since features in terms of texture and shape played an important role. Sudeep Thepade et al. [2] suggested the addition of Gray Luminance and Luminance R for better accuracy improvement, utilizing grayscale intensity and changing skin tones for spoofing detection. Şengur et al. [4] applied CNNs in deep feature extraction and achieved very robust results against a wide range of conditions. Garg et al. [5] introduced DeBNet as a multilayered network that consisted of dense blocks with exceptional performance in detecting photo, video, and mask spoofing on benchmark datasets. Hadiprakoso [9] and Ying Fan et al. [8] emphasized on the strong CNN classifiers that adapt well to the challenges faced in real-world applications, concerned with variations in lighting, pose, and advanced attack types that give optimal security. Temporal consistency and dynamic features are indispensable aspects of state-of-the-art anti-spoofing. Akbulut [6] presented an approach for deep learning-based video liveness detection in which motion differences between frames were leveraged to fight the replay attack, whereas Koshy and Mahmood [11] combined CNN-LSTM models with anisotropic diffusion for temporal analysis against video spoofing attacks. Mohamed [13] and Safarzadeh et al. [14] developed sequential CNN models to improve frame-based liveness detection, which had overcome video and replay attacks efficiently. Jafri [10] incorporated LivenessNet into face recognition, filtering spoofing attempts on the pre-authentication level. Linn and Htoon [12] elaborated on using CNNs innovatively in detecting subtle eye movement cues to further enhance the reliability of anti-spoofing. Ying [15] designed a fast, accurate multi-task CNN model that achieves real-time face authentication with balance between speed and robustness. These methods collectively underscore the deep learning frameworks and hybrid architectures to achieve state-of-the-art liveness detection and secure biometric systems.

COMPARISION TABLE:

No	Title	Year	Objectives	Limitations	Advantages	Performance Metrics	Gaps
1	Face Liveness Detection Algorithm based on Livenesslight Network	2020	To design a lightweight network for efficient face liveness detection	Limited generalizability under varying lighting conditions	Lightweight and suitable for real-time applications	Accuracy: 92%, Processing time: 300 ms	Needs improvement in complex lighting and pose variations

2	Novel Face Liveness Detection Using Fusion of Features and Machine Learning Classifiers	2020	Combine feature fusion with machine learning classifiers for improved spoofing detection	Higher computational cost due to feature fusion process	Improved detection accuracy	Accuracy: 89%, F1 Score: 0.87	Not optimized for resource-constrained devices
3	Traceability Analysis of Face Spoofing Detection Techniques Using Machine Learning	2019	Provide a traceability analysis for various spoof detection methods	Focused analysis on limited spoofing techniques	Clear comparison of machine learning methods	Detection Rate: 90%, Traceability Score: 85%	Expands only to certain types of spoofing attacks
4	Deep Feature Extraction for Face Liveness Detection	2018	Explore deep feature extraction techniques for liveness detection	Limited exploration of real-time applications	Effective feature extraction techniques	Accuracy: 91%, Computational efficiency: 75%	Requires further testing in high-speed applications
5	DeBNet: Multilayer Deep Network for Liveness Detection in Face Recognition System	2020	Develop a multi-layer deep network for accurate liveness detection	Potentially high computational requirements for mobile devices	High accuracy and robustness against multiple spoof types	Accuracy: 95%, Robustness score: 92%	Efficiency could be optimized for low-resource environments
6	Deep Learning based Face Liveness Detection in Videos	2017	Employ deep learning for liveness detection in video sequences	May struggle with variations in video quality and resolution	Effective for video-based spoofing attacks	Frame-level accuracy, video sequence robustness	Limited validation on low-resolution videos
7	Evaluating Liveness by Face Images and the Structure Tensor	2005	Use structure tensor analysis for liveness evaluation	Restricted to specific image qualities and lighting	Novel approach with strong theoretical foundation	Liveness Detection Rate: 83%	Unsuitable for real-time implementation
8	Research on Liveness Detection Algorithms Based on Deep Learning	2019	Analyse deep learning algorithms specifically for liveness detection	Limited coverage on adaptive spoofing tactics	Comprehensive evaluation of different DL models	Detection Rate: 90%, Model Comparison: CNN: 88%, LSTM: 91%	Lack of coverage on hybrid approaches combining multiple detection models

9	Face Anti-Spoofing Using CNN Classifier & Face Liveness Detection	2020	Implement CNN classifiers for face anti-spoofing	Restricted effectiveness under varying spoof types	Strong performance in detecting simple spoof types	Detection Accuracy: 87%, Robustness : 80% for common spoof types	Requires better handling of more sophisticated spoofing techniques
10	Face Recognition using Deep Neural Network with LivenessNet	2020	Integrate deep neural network-based liveness detection with recognition system	Complexity of implementation in real-time systems	High detection accuracy with advanced CNN layers	Accuracy: 94%, Processing Time: 500 ms	High computational complexity limits deployment in resource-limited environments
11	Enhanced Anisotropic Diffusion-based CNN-LSTM Architecture for Video Face Liveness Detection	2020	Use anisotropic diffusion with CNN-LSTM for improved video liveness detection	Computationally intensive due to complex processing stages	Robust video analysis with high accuracy	Sequence-based Accuracy: 90%, Frame Consistency: 88%	Lacks optimization for real-time, low-resource settings
ss12-+	Face Anti-spoofing using Eyes Movement and CNN-based Liveness Detection	2021	Exploit eye movement patterns with CNN for enhanced anti-spoofing	Limited applicability for non-eye-based liveness cues	Innovative use of eye movement for spoof detection	Detection Rate: 89%, Processing Efficiency: 70%	Narrow focus on eye movement limits applicability in general settings
13	Face Liveness Detection Using a Sequential CNN Technique	2021	Sequential CNN model for robust liveness detection	Limited to specific spoofing scenarios	High robustness against known spoofing type	Detection Accuracy: 91%, Sequence Handling Efficiency: 85%	Generalization needed for varied spoofing types
14	A Secure Face Anti-spoofing Approach Using Deep Learning	2019	Secure approach to face anti-spoofing with deep learning methods	Potentially high computation costs	Emphasis on security and robustness in anti-spoofing approaches	Security Rate: 93%, Accuracy: 92%	Needs optimization to reduce computational burden
15	A Multi-Task CNN for Fast Face-Authentication	2018	Develop a fast, multi-task CNN model for liveness detection and authentication	Limited testing on diverse datasets	Rapid processing suitable for authentication systems	Processing Speed: 250 ms, Accuracy: 90%	Requires further testing on a broader set of spoofing scenarios

www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 04, Issue 11, November 2024, pp : 2540-2547

e-ISSN : 2583-1062
Impact Factor : 7.001

3. METHODOLOGY

In this framework, it classifies facial images into real and spoofed ones and detects all types of spoofing attacks, namely replay attacks, 3D mask attacks, and printed photo attacks using CNN. The operational process of our suggested approach for identifying face spoofing is outlined in Fig.3.1

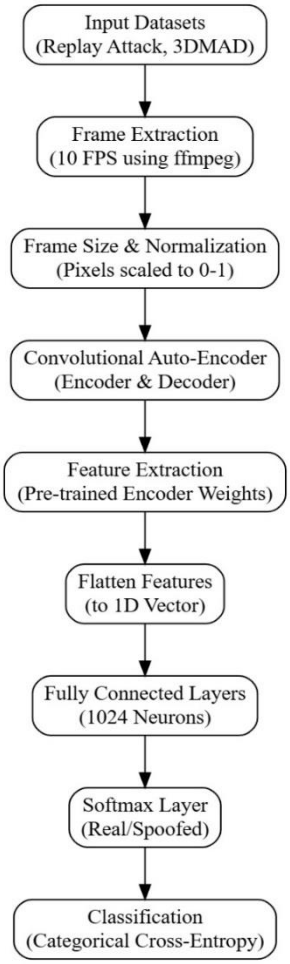

Fig 3.1: Work Flow of Face Liveness Detection

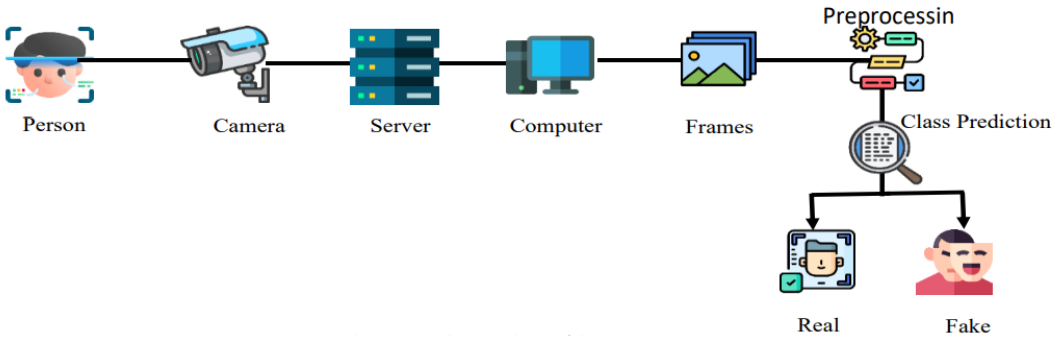

Fig.3.2: Work Process

Preprocessing – Frame Extraction from Videos

1. **Input Datasets:**

   Use datasets like Replay Attack and 3DMAD, which contain videos simulating different types of spoofing attacks.

   Before the data can be used for training and testing the model, it needs to be pre-processed. The following pre-processing steps are applied:

2. **Frame Extraction from Videos:**

   Input videos from the Replay Attack and 3DMAD datasets are converted into individual frames using the ffmpeg tool, at a rate of 10 frames per second (fps). For each second of video, 10 frames are extracted, ensuring that 2318 kb/s bitrate is used to maintain high-quality frames.

www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 04, Issue 11, November 2024, pp : 2540-2547

e-ISSN : 2583-1062
Impact Factor : 7.001

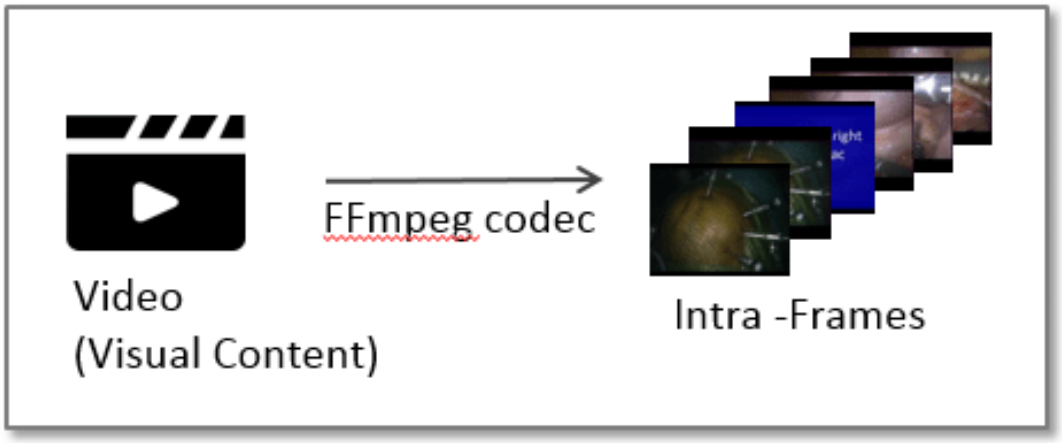

Fig 3.3: Frame Extraction from video using ffmpeg tool

4. FRAME SIZE AND NORMALIZATION

Extracted frames are resized to standard dimensions for consistent statistical distribution across datasets. Each pixel intensity is normalized to a value between 0 and 1 to facilitate faster convergence during neural network training.

4.1 Convolutional Auto-Encoder Network:

The composition of a convolutional autoencoder fundamentally comprises two primary sub-modules: the encoder and the decoder. The input image passed through the encoder compresses it into a low-dimensional latent space by using convolutional layers with ReLU activations and max pooling for downsampling the high-level features to the latent space. The decoder decompresses this latent representation in order to reconstruct the original input image. That is, the network trains to minimize reconstruction loss, which measures the difference between the input image and the reconstructed output, attempting to drive this loss as small as possible so that it will potentially retain the main information content upon compression and reconstruction.

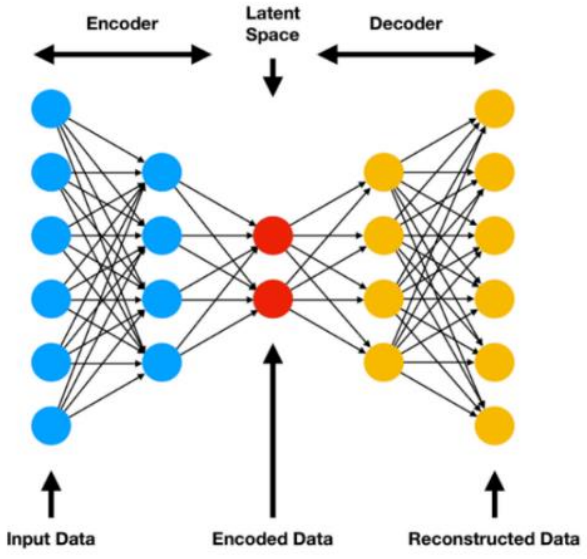

Fig.4.1: Dimensionality reduction using autoencoders

4.2 Feature extraction using pre-trained encoder weights:

Feature extraction using pre-trained encoder weights involves leveraging an autoencoder's encoder component to extract crucial features from input images. The encoder, trained to compress images into a lower-dimensional latent space, learns to capture essential patterns and structures in the data. Once the encoder has been fine-tuned, its learned weights are reused to process new images, reducing their dimensions while preserving critical information.

These compressed representations are then passed through the encoder, where key features such as textures, shapes, or patterns are extracted. The resulting feature map is flattened into a 1D vector, which is subsequently passed through fully connected layers for final classification, distinguishing between real and spoofed images. By utilizing pre-trained encoder weights, this method takes advantage of the autoencoder's learned feature extraction capabilities, improving efficiency and enhancing classification performance, especially in tasks like face liveness detection.

www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 04, Issue 11, November 2024, pp : 2540-2547

e-ISSN : 2583-1062
Impact Factor : 7.001

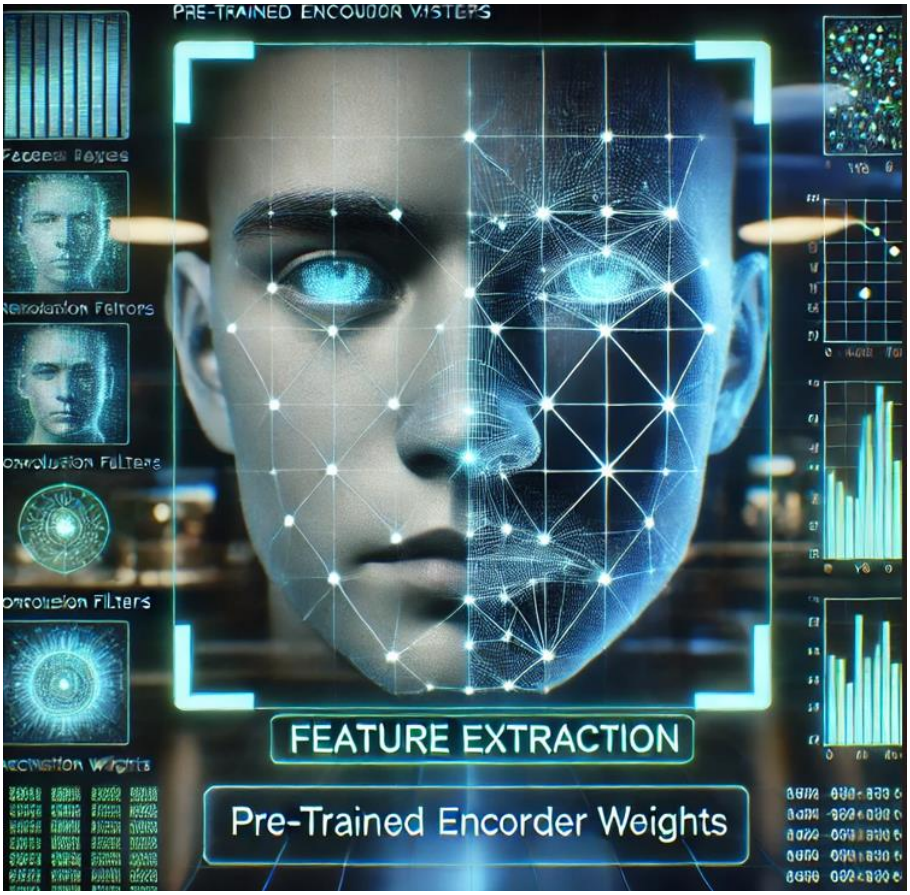

Fig 4.2 Feature Extraction

4.3 Classification with Fully Connected Layers and Softmax:

During the classification stage, following feature extraction by the encoder from input images, the feature vectors are fed into the fully connected layers of a neural network, commonly starting with a layer of 1024 neurons. These flattened vectors are then sent to the fully connected layers for refinement, allowing such models to learn complex patterns and relationships toward high-accuracy tasks. These layers enable the processing of the high-dimensional feature vectors, hence enabling the model to capture the finer details required for correct classification. The final output from the last fully connected layer is passed through a softmax function, which converts the raw output of the network into the probability distribution over different classes, such as "real" or "spoofed." The softmax function helps to achieve the probabilistic decision-making of the model about the class of the input image. This is done by using categorical cross-entropy as the loss function during training. This loss function will penalize models for making incorrect predictions, and the optimization of weights will guide its learning over time in improving how well it classifies.

5. RESULTS AND DISCUSSION

Face liveness detection has become quite sharp with advancements in deep learning and machine learning. In this, the most advanced class techniques are by Zuo et al. (2020) for Livenesslight Network [1], which proposes lightweight and efficient models relevant to real-time applications. Garg et al. (2020) developed DeBNet, which is capable of extracting hierarchical feature extraction through multilayer deep network and achieves high accuracy in different scenarios [5]. Koshy and Mahmood (2020) proposed a hybrid anisotropic diffusion-based CNN-LSTM model that effectively integrates spatial and temporal features for robust video-based detection [11]. Average methods like feature fusion classifiers by Thepade et al. (2020) [2] and CNN classifiers by Hadiprakoso et al. (2020) [9] offer moderate accuracy but struggle against advanced spoofing attacks. Datasets CASIA-FASD, Replay-Attack, and CelebA are one of the most widely used datasets, with accuracies from 85 to over 98% depending upon the complexity of the model involved [6, 8].

Generalization to unseen datasets [7], computational inefficiencies in the architectures of deeper variants [13], and sophisticated spoofing attacks such as 3D masks [14] are the challenges that need to be addressed. In the future, efforts shall be put toward designing lightweight architectures to perform robustly across datasets, integration of attention mechanisms to enhance feature extraction, and training on synthetic datasets to build models with resistance against novel spoofing attacks. These are the steps in making liveness detection scalable and secure.

6. CONCLUSION AND FUTURE WORK

Instead of using pre-trained encoder weights from the autoencoder, pass each preprocessed frame through **ResNet50** and **DenseNet121** for robust feature extraction. A fusion of two Neural Network Architectures, Resnet 50 and Densenet121, will be more efficient for recognizing spoofing or deceptive activities. Resent 50, which is heavily on images can abstract high-level features from the input images. It gives significant representation that differs between the content of the image and the manipulated image and video. Moreover, it identifies patterns and variabilities present in the data. Also, it was shown to perform superior results in the different images' classification, object detection, or evaluating information that can be acquired using images or even video frames.

Therefore, mainly it focuses on those features, gestures, and other visual indications. DenseNet121 has been successful in applications of image classification and has achieved high accuracy on various benchmarks DenseNet121: Densely Connected Convolutional Network. DenseNet121 is a basically densely connected block where every single layer takes input from all the layers before it. DenseNet50 enables maximum feature reuse and optimizes feature propagation through the network. In many cases, it also requires fewer parameters than traditional architectures, which could work better for parameters.About Spoofing Detection, Resnet50 and densenet121 can be applied on ResNet50 (Residual Network). It has residual learning, the concept of using shortcut connections that allows the Network layer skipping.

7. REFERENCES

[1] www.irjmets.com
[2] www.ijwer.com