International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 10-No.3 (2024) pp. 323-331
http://www.ijcesen.com

ISSN: 2149-9144

Research Article

A Novel Texture based Approach for Facial Liveness Detection and Authentication using Deep Learning Classifier

Khushboo JHA[1]*, Sumit SRIVASTAVA[2] and Aruna JAIN[3]

[1,2,3]Department of Computer Science and Engineering, Birla Institute of Technology Mesra, India.
* **Corresponding Author Email:** kjha.phd@gmail.com - **ORCID:** 0000-0003-1062-8128

[2]**Email:** sumit.srivs88@gmail.com - **ORCID:** 0009-0003-6880-2958
[3]**Email:** arunajain@bitmesra.ac.in

Article Info:

DOI: 10.22399/ijcesen.369
Received : 04 July 2024
Accepted : 26 July 2024

Keywords

Face authentication
Texture based feature
Liveness detection
Deep learning classifiers
Biometric system

Abstract:

In today's digital age, face authentication stands as a pivotal method for secure user verification, offering convenience and heightened security. Our approach addresses critical challenges like low illumination, pose variation, and spoofing attacks by integrating advanced facial feature extraction and liveness detection with deep learning classifiers. Texture based facial feature extraction technique is proposed by combining feature-level fusion of Global (Gabor Wavelets) and Local (Local Binary Patterns) features, termed as GW-LBP. Moreover, the proposed texture based approach is also utilized for liveliness detection to analyze temporal and spatial variations indicative that the facial image belongs to live face or photograph or video (spoof). Using Our Database of Faces (ORL) dataset, this approach is evaluated using three deep learning classifiers: Convolutional Neural Network, ResNet50 and Vision Transformers which achieved an accuracy of 96.5%, 97.2% and 97.9% respectively. Moreover, the proposed approach demonstrates significant improvements in several other performance measures and feature extraction techniques and surpasses current cutting-edge methods as a resilience solution for user authentication.

1. Introduction

In an era dominated by digital transformation and increasing reliance on secure authentication mechanisms [1], Face Authentication (FA) [2] has emerged as a pivotal technology with applications spanning from smartphone unlocking to border security and financial transactions. The ability to accurately and efficiently verify identities based on facial features not only enhances convenience but also strengthens security measures against unauthorized access and identity fraud. As such, the development of robust FA systems remains a critical area of research and innovation.

Moreover, conventional FA systems are susceptible to spoofing attacks through photos or videos because they lack features to prevent attackers. In these types of attacks, the attacker pretends to be a legitimate user in order to access the system without authorization. Face spoofing attacks employ various techniques, such as paper printouts of face images, mobile phone displays showing face photos, paper masks, and video footage of a person's face. To defend against these attacks, face liveness detection should be implemented. Face liveness detection is a technique used to distinguish between genuine and counterfeit face images. A live face exhibits natural movements, including blinking, head movements and smiling, influenced by physiological processes such as blood flow and breathing, which cause changes in skin tone and texture. Whereas, counterfeit faces lack such authentic reactions and physiological reactions, making them easily detectable by face liveness detection methods. For instance, a photograph or video of a face will not display any blinking or other facial movements, signifying that the image is not genuine. Likewise, a face mask or digital facial photograph will fail to accurately depict changes in skin tone or texture, suggesting that it is not a live person and real faces present different texture patterns in comparison with fake ones [3-7]. Thus, face liveness detection is crucial in FA system to ensure that only genuine images are processed for recognition. The detection process involves analyzing various indicators or presenting challenges to determine the authenticity of the subject [7]. Moreover, by integrating liveness detection mechanisms, which ensure that the presented facial image is from a live person rather than a static image or video, the methodology enhances security against spoofing attacks. This integration not only improves the system's reliability but also addresses critical concerns in biometric authentication system regarding vulnerability to presentation attacks. Advantages of facial liveness detection are:

- *Enhanced Security*: By ensuring that only live faces are authenticated, the system becomes robust against spoofing attacks, significantly enhancing security.
- *Improved Accuracy*: The integration of liveliness detection with global and local feature extraction reduces false positives and false negatives, leading to improved authentication accuracy.
- *Cost-Effectiveness*: Utilizing software based liveliness detection methods eliminates the need for additional hardware, making the system more cost-effective and easier to deploy.

Problem Formulation

Traditional methods often fail to adequately differentiate between genuine and spoofed inputs (through photos or videos), illumination variations, etc. due to their reliance on limited feature sets. To address these limitations, this research proposes a novel texture-based approach that integrates Gabor Wavelets (GW) [6] and Local Binary Patterns (LBP) [7] at feature level as GW-LBP for facial liveness detection as well as for feature extraction. The GW feature is well-suited for capturing global facial structures by analyzing spatial frequency and orientation, providing a holistic representation of facial features. On the other hand, LBP excel in capturing fine-grained textures and details within localized facial regions, enhancing the discriminative power of the authentication system. The proposed GW-LBP features offer a promising avenue for enhancing the accuracy, interpretability, and security of FA system. By bridging the gap between global structure and local texture features [2,3], the methodology as shown in Fig. 1, aims to set new benchmarks in biometric security, paving the way for reliable and scalable solutions in diverse application domains.

In this context, the study evaluates the performance of the GW-LBP methodology across multiple modern classification models, including Convolutional Neural Network (CNN), ResNet50 and Vision Transformer (ViT). Performance metrics such as accuracy, precision, recall, F1-score, and Equal Error Rate (EER) are analyzed to assess the effectiveness of proposed methodology in real-world scenarios. The findings contribute to advancing the field of facial authentication by providing insights into the comparative advantages of hybrid feature extraction techniques over deep learning-based classifiers.

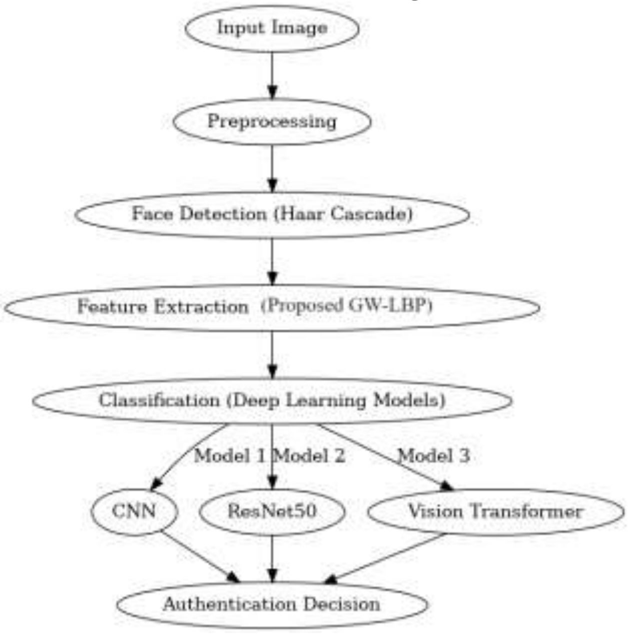

Figure 1: *Flowchart of the proposed FA system*

2. Material and Methods

A comprehensive block diagram for the FA system is presented in Fig. 2, detailing both the training and testing phases [2]. In the training phase, the input facial image undergoes preprocessing followed by face detection. Upon successful detection, the face is forwarded to the feature extraction module to derive distinct features and to check liveness detection. These extracted features are then processed using a classifier and stored in the database, thereby completing the enrollment process of the user's facial image. In the testing phase, the captured facial image is compared one-to-one against the stored reference template in the database. The user is authenticated or rejected based on the similarity score between the query image and the reference template.

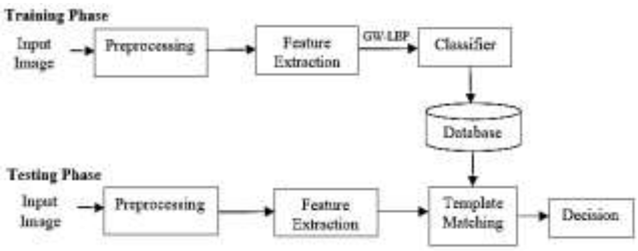

Figure 2. *Block diagram of FA system*

2.1 Data Acquisition

The ORL Database [2] is a publicly available dataset that includes face images from 40 subjects, with ten images per subject. The images in this collection showcase a diverse range of facial variations, captured under different lighting conditions and facial expressions, allowing for comprehensive evaluation. This dataset presents a challenge due to the diverse range of facial expressions and lighting conditions. The size of each picture is 92x112 pixels, with each pixel containing 256 levels of grayscale to to reduce complexity.

2.2 Pre-processing: Face Detection

Preprocessing step is performed on each facial image to enhance the quality and consistency of the data. The Histogram equalization [4] is utilized to enhance contrast to improve feature extraction, followed by the normalization of pixel values to a range [0, 1], to facilitate model training. Then, the Haar Cascade algorithm [4] is applied to detect and crop the face region from each image. This ensures that the feature extraction process focuses on the facial area.

2.3 Feature Extraction

Feature extraction [8] plays a pivotal role in FA by converting raw facial images into meaningful representations. It captures distinctive features such as texture, edges and patterns, which are crucial for distinguishing between different facial templates or individuals. Effective feature extraction techniques, such as GW and LBP, enhance the robustness and accuracy of FA systems. By leveraging these techniques, face authentication systems can reliably identify and verify individuals, even under varying conditions and potential spoofing attacks.

2.3.1 Global Feature: GW

- Global features [2] are the structural configuration of face organs and facial contour as a whole. Gabor filters are applied to extract texture features at 5 scales and 8 orientations [6]. This property makes them highly effective for analyzing facial textures under varying conditions.
- Each GW [6] is parameterized by a specific frequency and orientation, allowing it to respond to specific spatial patterns within the image. This makes Gabor features highly discriminative for detecting fine details in facial textures. It is defined as equation 1:

$$G(x,y;\lambda,\theta,\psi,\sigma,\gamma)=exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)cos\left(2\Pi\frac{x'}{\lambda}+\psi\right) \quad (1)$$

where $x' = x\cos\theta + y\sin\theta$ and $y' = -x\sin\theta + y\cos\theta$. Parameters λ denotes wavelength, θ denotes orientation, ψ denotes phase offset, σ denotes standard deviation and γ denotes aspect ratio.

- Gabor features are less sensitive to changes in lighting conditions, making them suitable for real-world applications where illumination variations are common.
- In the context of facial liveness detection, Gabor wavelets [6] are used to extract features that represent the underlying texture of the facial skin. Genuine facial images exhibit natural skin textures with complex spatial patterns, while spoofed images (such as photos or masks) often lack such intricate details. By analyzing the Gabor responses across different scales and orientations, a liveness detection system effectively differentiates between live and fake facial inputs.

2.3.2 Local Feature: LBP

- Local features [2] occur frequently and are determined by the orientation and position of the facial images. LBP [7] is a widely used technique for texture analysis, known for its simplicity and computational efficiency.
- LBP captures the local texture information by comparing each pixel with its neighbouring pixels.
- LBP encodes the texture information by creating a binary pattern for each pixel, based on the intensity differences with its neighboring pixels. This binary pattern is then converted into a decimal value to form the LBP code. The LBP code for a pixel (x,y) is denoted as equation 2:

$$LBP(x,y) = \sum_{p=0}^{P-1} s(i_p - i_c) \cdot 2^p \quad (2)$$

where s(x) is 1 if $x \geq 0$ and 0 otherwise, i_c is the center pixel intensity, and i_p are the intensities of P surrounding pixels.

- The LBP operator is extended to achieve rotation invariance, making it robust to changes in orientation.
- It is computationally efficient, making it suitable for real-time applications. It requires minimal processing power, which is advantageous for liveness detection systems operating on embedded devices.
- For facial liveness detection, LBP features are used to capture the micro-textures [7] of the facial skin. Genuine facial images exhibit fine-grained textures, while spoofed images often show smoother surfaces due to printing or display artifacts. By extracting LBP features, the system distinguishes between the natural micro-

textures of a live face and the artificial textures of a spoofed face.

2.3.3 Proposed GW-LBP feature

The GW [6] and LBP [7] features are concatenated at feature level fusion to create a combined GW-LBP feature vector for each facial image. For this work, Gabor filters are applied to extract texture features at 5 scales and 8 orientations and total feature vector length of $5 \times 8 = 40$ is used. For LBP extraction, taking the radius of 1, number of neighbors as 8 gives total length of feature vectors $2^8 = 256$ is used. Therefore, the combined feature vector length of $40 + 256 = 296$ i.e., GW-LBP fusion captures both the global texture (Gabor) and local texture (LBP) information. Moreover, by leveraging the strengths of both methods, a more comprehensive proposed GW-LBP feature set as shown in Fig. 3 is obtained, leading to improved discrimination between live and fake faces as well as robust FA system.

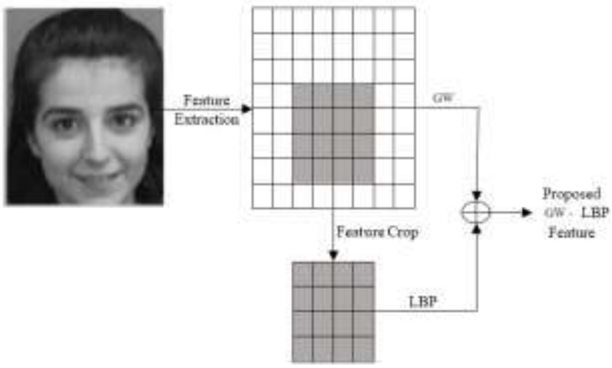

Figure 3. *Proposed GW-LBP feature*

2.4 Classification

The final step involves classification [8] of extracted features to authenticate the user. The proposed approach assesses the efficacy of different state-of-the-art deep learning models such as CNN, ResNet50 and ViT in categorizing the retrieved features, as mentioned below:

2.4.1 CNN Model

A CNN model [2, 9] tailored for facial authentication using GW-LBP features would typically involve multiple convolutional layers followed by pooling layers and fully connected layers. Explanation of CNN Layers:

- **Convolutional Layers (Conv1, Conv2, Conv3)**: These layers apply convolutional filters to the input, learning spatial hierarchies of features. Each convolutional layer is followed by an activation function (ReLU) to introduce non-linearity.
- **Max Pooling Layers (MaxPool1, MaxPool2, MaxPool3)**: These layers downsample the spatial dimensions (height and width) of the feature maps, reducing the computational load and controlling overfitting.
- **Flatten Layer**: Converts the 2D feature maps into a 1D feature vector that can be fed into the fully connected layers.
- **Fully Connected (Dense) Layers**: These layers learn to combine the features extracted by the convolutional layers to make the final classification. They typically include activation functions such as ReLU for intermediate layers and softmax for the output layer.

2.4.2 ResNet50 Model

ResNet50 [10] is a deep network with 50 residual connections. These connections help in training deeper networks by mitigating the vanishing gradient problem and improving feature extraction depth. Deeper networks improve the accuracy and robustness of FA system. Each set of residual blocks contains multiple bottleneck blocks, which are the building blocks of the ResNet architecture. Explanation of the ResNet50 architecture:

- **Input Image**: The input to the model, usually an RGB image of fixed size.
- **Conv1: 7x7, 64, stride 2**: The first convolutional layer with a 7x7 filter, 64 output channels, and a stride of 2.
- **Max Pooling: 3x3, stride 2**: Max pooling layer with a 3x3 filter and a stride of 2 to reduce the spatial dimensions.
- **Conv2_x: 3x3, 64**: The first set of residual blocks (Conv2_x) containing 3 bottleneck blocks, each with a 3x3 filter and 64 output channels.
- **Conv3_x: 3x3, 128**: The second set of residual blocks (Conv3_x) containing 4 bottleneck blocks, each with a 3x3 filter and 128 output channels.
- **Conv4_x: 3x3, 256**: The third set of residual blocks (Conv4_x) containing 6 bottleneck blocks, each with a 3x3 filter and 256 output channels.
- **Conv5_x: 3x3, 512**: The fourth set of residual blocks (Conv5_x) containing 3 bottleneck blocks, each with a 3x3 filter and 512 output channels.
- **Average Pooling: 7x7**: Average pooling layer with a 7x7 filter to reduce the spatial dimensions to 1x1.

- **Fully Connected: 1000**: Fully connected layer with 1000 output neurons, corresponding to the number of classes.
- **Softmax**: Softmax layer to output the probability distribution over the 1000 classes.

2.4.3 Vision Transformer (ViT) Model

ViT [11] is a transformer-based model that uses self-attention mechanisms to classify facial features by capturing global dependencies between input features. Explanation of different layers of ViT model:

- **Input Layer**: Accepts combined GW-LBP features.
- **Patch Embedding**: Converts each P×P patch to a D-dimensional embedding.
- **Positional Embedding**: Adds positional information to each patch embedding.
- **Class Token**: Learnable token prepended to the patch embeddings.
- **Transformer Encoder**: 12 stacked encoder layers, each with multi-head self-attention and FFN.
- **Multi-Head Self-Attention**: Attention mechanism with multiple heads.
- **Layer Normalization**: Normalization before and after attention/ Feed-Forward Network.
- **FFN**: Two fully connected layers with ReLU activation.
- **Output Layer:** Fully connected layer for classification.
- **Softmax Activation:** Softmax activation for classification.

These three models are trained on ORL dataset using proposed feature vector GW-LBP, to classify the facial templates for user authentication, improving overall accuracy and security of FA system.

2.5 Performance Metric

Any deep learning research study must evaluate model performance. In this different classification-focused study, we heavily use the following metrics [12-18] to determine the performance:

- **Accuracy**: Indicates the overall correctness of face authentication predictions [2] across all classes (faces).
- **Precision**: It measures the proportion of correctly identified authentic faces among all faces predicted as authentic.
- **Recall**: Indicates the proportion of correctly identified authentic faces among all actual authentic faces.
- **F1-score**: It is a harmonic mean of precision and recall. It provides a balanced measure for model performance.
- **Equal Error Rate (EER)**: It is a threshold where False Acceptance Rate (FAR) and False Rejection Rate (FRR) are equal, indicating the point of optimal balance in authentication performance.

3. Results and Discussions

The testing of the different classifiers is done on ORL dataset, which ensure the system's effectiveness in real-world scenarios. The classification models are trained using 80% of the dataset, with the remaining 20% used for testing. Then the trained classifiers are ready to the test and classify each image as real or fake and FA. The experimental approach is carried out using the Jupiter Lab component of the Anaconda development environment, which is built on the Python programming language.

3.1 Analysis of Different Deep Learning Classifier using Standard Performance Metrics

Standard evaluation metrics such as accuracy, precision, recall, F1-score and EER are employed to assess the performance of each classifier models, as shown in table 1. From computational perspective, the GW-LBP methodology demonstrates high accuracy, precision, recall, and F1-score across all models, indicating robust facial authentication capabilities. In the context of FA, lower EER indicates better performance because it represents the threshold where the system balances between accepting impostors and rejecting legitimate users. As per table 1, lower EER values achieved by all classifier model for proposed feature indicates better system performance in balancing between accepting legitimate users and rejecting impostors, highlighting the effectiveness of GW-LBP in authentication tasks. Consistent performance across ResNet50, ViT and CNN models underscores the generalizability and reliability of the GW-LBP methodology in diverse deep learning architectures. Moreover, the experimental results demonstrate that the proposed approach using novel facial features with liveness detection and deep learning models achieves high accuracy and robustness than individual feature based approaches. The ViT model, in particular, shows superior performance than CNN and ResNet50 due to its ability to capture and classify global context through self-attention mechanisms.

3.2 Analysis of different deep learning classifier models using Receiver Operating Characteristic (ROC) curve

Table 1. *Performance of proposed approach with various features and classifiers based FA system*

Feature Extraction	Classification Model	Accuracy (in %)	Precision (in %)	Recall (in %)	F1-score (in %)	EER (in %)
GW Feature	ResNet50	91.5	92.0	91.0	91.5	5.0
	ViT	90.8	91.2	90.5	90.9	5.5
	CNN	89.7	90.0	89.5	89.8	6.2
LBP Feature	ResNet50	89.2	89.5	89.0	89.2	6.8
	ViT	88.5	88.8	88.2	88.5	7.0
	CNN	87.8	88.0	87.5	87.8	7.5
Proposed GW-LBP Feature	**ResNet50**	**97.2**	**96.4**	**96.0**	**96.2**	**3.2**
	ViT	**97.9**	**97.1**	**95.7**	**96.4**	**2.5**
	CNN	**96.5**	**96.3**	**94.8**	**96.0**	**3.4**

In order to shed additional light on the findings of table1, Fig. 4-6 offers a comparative assessment of the proposed approach.

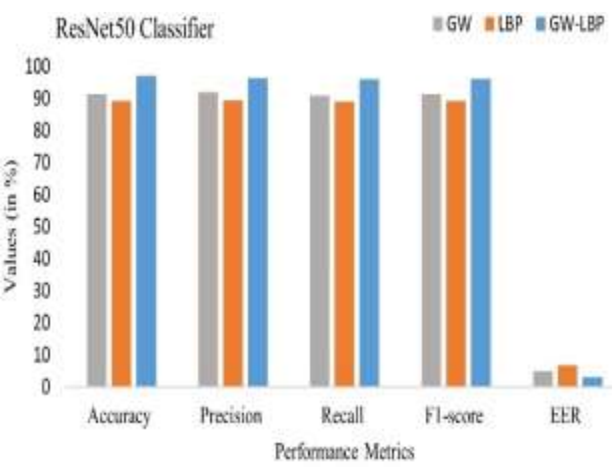

Figure 4: *Comparison of Different Feature Extraction Technique based FA System using ResNet50 Classifier*

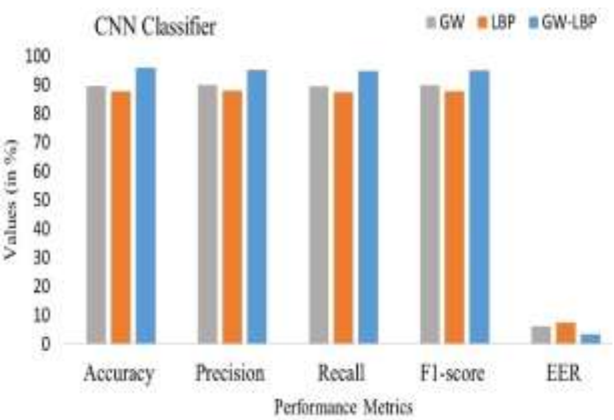

Figure 5: *Comparison of Different Feature Extraction Technique based FA System using CNN Classifier*

ROC curve [15] is a fundamental tool for evaluating the performance of a binary classification system. It is widely used in various fields, including medical diagnostics, machine learning, and biometric systems, to assess the accuracy and discriminative power of a classifier. Its ability to provide a threshold independent assessment, visualize trade-offs and handle imbalanced datasets makes it an indispensable metric in classification tasks. The True Positive Rate (TPR) also known as sensitivity or recall measures the proportion of actual positives correctly identified by the classifier. The False Positive Rate (FPR) measures the proportion of actual negatives incorrectly identified as positives. The ROC curve plots TPR against FPR, providing insights into both types of errors and allows for easy comparison of multiple classifiers. The classifier with the highest area under the ROC curve (AUC) [15] is considered to have better performance. As shown in Fig. 7, ViT outperforms CNN and ResNet50 classifiers.

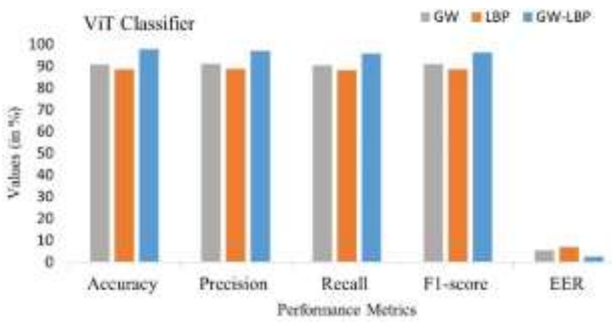

Figure 6: *Comparison of Different Feature Extraction Technique based FA System using ViT Classifier*

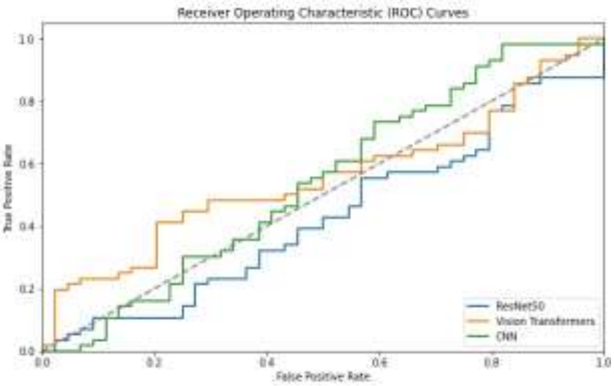

Figure 7: *ROC Curves for Different Classifiers*

ViT offers significant advantages over traditional CNNs and ResNet50, particularly in terms of global context understanding through self-attention mechanisms, scalability with large datasets, parameter efficiency, transfer learning capabilities, robustness to adversarial attacks and flexibility with image resolutions. These advantages make ViT a powerful tool for image classification tasks in modern deep learning applications.

3.3 Comparison of Proposed Approach with Different State-of-the-Art Techniques

The field of FA methods has been the subject of investigation from a variety of perspectives. Recently, [12] proposed an interest point-based FA system that makes use of an adaptive neuro fuzzy inference system (ANFIS) classifier. Similarly, [13] advocated the use of Principal Component Analysis (PCA) and CNN for the purpose of FA utilizing eigenfaces. A unique method for feature extraction is proposed in the work [14], which is referred to as Relative Gradient Magnitude Strength (RGMS). RGMS is able to extract features for every two surrounding regions, which is a very essential step in the process of extracting more discriminative local patterns for the human face. The classification was carried out with the assistance of a Deep Neural Network (DNN) classifier. A Hybrid Robust Point Set Matching Convolutional Neural Network (HPPSM-CNN), was proposed in [15] with the intention of improving facial recognition in situations where there are no constraints. [2] proposed a DCT-ULBP feature set, which is an integration of global (Discrete Cosine Transform, DCT) and local (Uniform Local Binary Pattern, ULBP) feature extraction approaches. The purpose of this feature set is to extract discrete and resilient characteristics from the facial picture by utilizing CNN classifier. The work in [16] presented a CNN-based face recognition method that makes use of sequence and hidden temporal information. In another work [17], an integrated technique for variation-resistant facial recognition that is based on CNN architecture was proposed. A comparative analysis of the performance of the aforementioned FA approaches with regard to feature extraction, classification and the use of the ORL dataset is presented in Table 2.

The comparison highlights the efficacy of the proposed approach in outperforming existing state-of-the-art approaches with respect to all classifiers. Therefore, it can be inferred that the proposed texture based features represents a significant adva-

Table 2: *Comparison of proposed approach with state-of-the-art FA techniques on ORL Dataset*

Ref.	Feature Used	Classifier	Accuracy
[12], 2019	Variance, maximum intensity & mean	ANFIS	96 %
[13], 2019	PCA	CNN	93 %
[14], 2020	RGMS	DNN	96.5 %
[15], 2022	RPSM-CNN	HRPSM-CNN	97 %
[2], 2023	DCT-ULBP	CNN	94.58 %
[16], 2023	Sequence & hidden temporal features	CNN	96.79 %
[17], 2023	CNN	CNN	93.75 %
[18], 2024	Locality preserving projections	Autoencoder	96.57 %
Ours	**GW-LBP**	**ResNet50**	**97.2 %**
		ViT	**97.9 %**
		CNN	**96.5 %**

ncement in FA system. By combining the strengths of global and local feature extraction with dynamic liveliness cues, the proposed methodology ensures high accuracy, security, and reliability. This holistic approach with deep learning based classifier addresses vulnerabilities in FA system and paves the way for robust biometric authentication solution.

4. Conclusions

FA is essential for secure user verification in the digital age, providing both convenience and security. This study addresses challenges such as low illumination and spoofing attacks through advanced facial feature extraction and liveness detection integrated with deep learning classifiers. The proposed texture based GW-LBP feature, combines global (GW) and local (LBP) features, enhances robustness and is crucial for distinguishing between real and fake users. The evaluation on the ORL dataset, the approach achieved accuracies of 96.5% with CNN, 97.2% with ResNet50, and 97.9% with ViT classifiers, respectively. Thus, it can be inferred that the proposed approach showed significant improvements in every performance metrics. Consequently, this approach not only advances the current state-of-the-art but also offers a robust solution that can effectively counteract spoofing attempts and perform well under varying environmental conditions, thus advancing the field of biometric security and FA for various practical applications.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] Jha, K., Jain, A., & Srivastava, S. (2024). A Secure Biometric-Based User Authentication Scheme for Cyber-Physical Systems in Healthcare. *Int. J. Exp. Res. Rev*, 39, 154-169. https://doi.org/10.52756/ijerr.2024.v39spl.012

[2] Jha, K., Srivastava, S., & Jain, A. (2023, March). Integrating Global and Local Features for Efficient Face Identification Using Deep CNN Classifier. In 2023 International Conference on Device *Intelligence, Computing and Communication Technologies,(DICCT)* (pp. 532-536). IEEE. https://doi.org/10.1109/DICCT56244.2023.10110170

[3] Huang, J., Yuen, P. C., Lai, J. H., & Li, C. H. (2004). Face recognition using local and global features. *EURASIP Journal on Advances in Signal Processing*, 1-12. https://doi.org/10.1155/S1110865704312187

[4] Meddeb, H., Abdellaoui, Z., & Houaidi, F. (2023). Development of surveillance robot based on face recognition using Raspberry-PI and IOT. *Microprocessors and Microsystems*, 96, 104728.

[5] Li, C., Huang, Y., Huang, W., & Qin, F. (2021). Learning features from covariance matrix of gabor wavelet for face recognition under adverse conditions. *Pattern Recognition*, 119, 108085. https://doi.org/10.1016/j.patcog.2021.108085

[6] Hassan, M. M., Hussein, H. I., Eesa, A. S., & Mstafa, R. J. (2021). Face recognition based on gabor feature extraction followed by fastica and lda. *Computers, Materials and Continua*, 68(2), 1637-1659. https://doi.org/10.32604/cmc.2021.016467

[7] Freitas Pereira, T. D., Komulainen, J., Anjos, A., De Martino, J. M., Hadid, A., Pietikäinen, M., & Marcel, S. (2014). Face liveness detection using dynamic texture. *EURASIP Journal on Image and video Processing*, 2014, 1-15. https://doi.org/10.1186/1687-5281-2014-2

[8] Jha, K., Jain, A., & Srivastava, S. (2024). Analysis of Human Voice for Speaker Recognition: Concepts and Advancement. *Journal of Electrical Systems*, 20(1), 582-599. https://doi.org/10.52783/jes.806

[9] Bacak, A., Şenel, M., & Günay, O. (2023). Convolutional neural network (CNN) prediction on meningioma, glioma with Tensorflow. *International Journal of Computational and Experimental Science and Engineering*, 9(2), 197-204. https://doi.org/10.22399/ijcesen.1306025

[10] Kırelli, Y., & Aydın, G. (2023). Classification of Histopathological Images in Automatic Detection of Breast Cancer with Deep Learning Approach. *International Journal of Computational and Experimental Science and Engineering*, 9(4), 359-367. https://doi.org/10.22399/ijcesen.1332504

[11] Yu, Z., Cai, R., Cui, Y., Liu, X., Hu, Y., & Kot, A. C. (2024). Rethinking vision transformer and masked autoencoder in multimodal face anti-spoofing. *International Journal of Computer Vision*, 1-22. https://doi.org/10.1007/s11263-024-02055-1

[12] Rejeesh, M. R. (2019). Interest point based face recognition using adaptive neuro fuzzy inference system. *Multimedia Tools and Applications*, 78(16), 22691-22710. https://doi.org/10.1007/s11042-019-7577-5

[13] Zafaruddin, G. M., & Fadewar, H. S. (2019). Face recognition using eigenfaces. *In Computing, Communication and Signal Processing: Proceedings of ICCASP* 2018 (pp. 855-864). Springer Singapore.

[14] ElBedwehy, M. N., Behery, G. M., & Elbarougy, R. (2020). Face recognition based on relative gradient magnitude strength. *Arabian Journal for Science and Engineering*, 45(12), 9925-9937. https://doi.org/10.1007/s13369-020-04538-y

[15] Tamilselvi, M., & Karthikeyan, S. (2022). An ingenious face recognition system based on

HRPSM_CNN under unrestrained environmental condition. *Alexandria Engineering Journal*, 61(6), 4307-4321. https://doi.org/10.1016/j.aej.2021.09.043

[16] Qi, X., Wu, C., Shi, Y., Qi, H., Duan, K., & Wang, X. (2023). A convolutional neural network face recognition method based on BILSTM and attention mechanism. *Computational Intelligence and Neuroscience*, 2023(1), 2501022. https://doi.org/10.1155/2023/2501022

[17] Benradi, H., Chater, A., & Lasfar, A. (2023). A combined method based on CNN architecture for variation-resistant facial recognition. *International Journal of Electrical and Computer Engineering Systems*, 14(9), 993-1001.

[18] Ran, R., Feng, J., Li, Z., Wang, J., & Fang, B. (2024). Locality preserving projections with autoencoder. *Expert Systems with Applications*, 242, 122750. https://doi.org/10.1016/j.eswa.2023.122750