Proceedings of the International Conference on Industrial Engineering and Operations Management

Publisher: IEOM Society International, USA DOI: 10.46254/NA11.20260434
Published: June 9, 2026

Towards Explainable Fingerprint Presentation Attack Detector using Deep Learning

Syeda Zartashia Shafique
Dept. of Software Engineering
University of Engineering and Technology
Taxila, Pakistan
zartashia.shafique@riphah.edu.pk

Ali Javed
Dept. of Software Engineering
University of Engineering and Technology
Taxila, Pakistan
ali.javed@uettaxila.edu.pk

Mohammed Abouheaf
Robotics Engineering
Bowling Green State University
Bowling Green, OH, MI, USA
mabouhe@bgsu.edu

Muteb Aljasem
Assistant Professor, School of Engineering
Bowling Green State University
Bowling Green, OH, MI, USA
aljasem@bgsu.edu

Sonia Chacko
Assistant Professor
Robotics Engineering
Bowling Green State University
Bowling Green, OH, MI, USA
soniamc@bgsu.edu

Abstract

Fingerprint presentation attack detection (FPAD) has been a pivotal part of biometric security systems that aim to distinguish genuine fingerprints from spoofed ones. Existing FPADs have achieved reasonable performance; however, they exhibit limited interpretability and generalization, especially when encountering unseen attacks. To address these limitations, we propose an explainable fingerprint presentation attack detector framework that incorporates a MobileNet-inspired feature extractor and a customized classification head along with explainable analysis. Efficient

feature extraction is ensured using the MobileNet feature extractor, thereby reducing computational cost. The classification head refines the feature representation, thus improving the accuracy and explainability. A feature extractor built upon batch normalization, multiple convolutional layers, and ReLU, followed by fully connected layers, including dense layers and dropout, for final classification. Incorporating explainability into proposed approach offers transparency in decision-making and enables insight into the most responsible features for detecting spoofing attempts. Extensive experimentations on the standard and diverse LivDet 2011 benchmark dataset show that our model achieves high accuracy while providing intuitive explanations of its predictions.

Keywords
Biometrics Liveness Detection, Deep learning, Fingerprint Anti-spoofing, Presentation Attack Detection, Trustworthy AI.

1. Introduction
Fingerprint Presentation Attack Detection (FPADs) is a pivotal component of biometric authentication systems due to its convenience, reliability, and widespread applications in areas such as banking, mobile devices, and border control (Muhammad et al. 2024). FPAD remains vulnerable to unseen Presentation Attacks (PAs), where an attack is planted using impersonated replicas of real fingerprints, typically made of materials such as silicone, latex, or 3D replicas, posing significant challenges to system security (Agarwal et al. 2022). Researchers have been developing robust and generalizable Presentation Attack Detection (PAD) methods to enhance system reliability. These methods first rely on handcrafted features, which show some effectiveness but often fail to capture fine textural differences between spoof and live fingerprints due to the diversity of spoofing techniques, environmental variations, and limited generalizability across datasets, thus making it difficult to detect PAs (Shukla et al. 2019).

Deep learning (DL) models have proven to be a game-changer for PAD, as the data-driven feature extraction showed promising results in detecting known PAs (Heusch et al. 2020). PAD is approached by Convolutional Neural Network (CNNs) as a binary classification problem, where the output of CNN models is only the PA score, and the classes are "live" vs "Spoof". Despite their performance, they remain vulnerable to unseen attacks (Yambay et al. 2017). One of the drawbacks they have is a lack of transparency in their decision-making, further limiting their trustworthiness and adaptability in complex spoofing scenarios where transparency and generalizability are critical (Buhrmester et al. 2021).

Explainable AI (XAI) surfaced as a transformative approach that helps mitigating limitation of opaqueness and explains internal decision-making processes. Incorporating XAIs into the concerned framework not only contributes to the model prediction's interpretability but also contributes to its robustness against spoofing attacks. Developing an explainable fingerprint presentation attack detector that improves both the model's effectiveness and explanation capabilities is the focus of this research. By leveraging XAI (Uchiyama et al. 2023), this research bridges the gap between transparency and performance, thereby paving the way for more trustworthy and reliable fingerprint liveness detection systems. The primary contributions of this paper are outlined as follows:

- We introduce a two-stage image enhancement pipeline that integrates filter-based edge-preserving sharpening with HSV-domain brightness normalization to effectively mitigate sensor variability, enabling substantial improvements in low-quality fingerprint images by improving the visibility of discriminative ridge structures before feature extraction.
- We leverage a transfer learning approach and introduce an explainable FPAD framework combining a lightweight MobileNet-based feature extractor with a customized classification module, which achieves model parameter reduction while preserving high classification accuracy and enabling real-time inference speeds suitable for deployment in resource-constrained biometric systems.
- We establish a comprehensive and explainable framework that integrates both model-specific and model-agnostic explainability techniques into the FPAD pipeline to deliver multi-level decision transparency and visual explanations, thereby contributing to user trust and transparency.
- Rigorous experimentation was conducted on the standard and diverse LivDet2011 dataset containing various types of sensors and spoofing materials, showing the framework's effectiveness in challenging fingerprint presentation attack scenarios.

- We provide qualitative explainability analysis showing that the proposed model predominantly focuses on fingerprint anatomical characteristics, including ridge continuity, pore distributions, and texture irregularities, thereby supporting the trustworthiness of the learned decision process.

In contrast to the existing methods that only aim to achieve maximum classification accuracy or use explainability methods for post-processing analysis, the proposed approach integrates lightweight deep feature extraction, fingerprint image enhancement, and multi-level explainability analysis in a single pipeline. This integration allows for high detection performance while also offering an easy interpretation of decisions, addressing two key challenges in modern fingerprint liveness detection systems: computational efficiency and model trustworthiness. The rest of the paper is organized in a way that Section 2 provides a review of related work. Section 3 provides a description of the proposed methodology. Section 4 discusses experiments and results, and finally, Section 5 concludes our study with future directions.

2. Literature Review

The rapid advancement in artificial intelligence, particularly in DL, has transformed the FPADs to a greater extent. Earlier research work focused on handcrafted features, including texture histograms, local binary patterns, and ridge orientation, expressive yet underperformed when encountered with different sensor or material conditions. With the emergence of CNNs, feature extraction becomes more adaptive and data-driven, leading models to learn minute visual cues effectively in PADs. The key challenge CNNs face to date is generalizability, and it can be improved with attentive fine-tuning of the model (Poojary et al. 2019). Deep pixel-wise supervision improves local texture learning for PA (Yu et al. 2021). MobileNetV2 is fine-tuned and integrates inverted residuals and linear bottlenecks, which help in optimizing representational efficiency (Sandler et al. 2018). Additionally, lightweight CNNs were proposed, specialized in low computational, memory, and energy consumption. In PADs, they result in achieving performance with low computational resources (Park et al. 2019). This is further supported by real-time inference in embedded systems with MobileNet and EfficientNet (Howard et al. 2017; Tan et al. 2019).

The other limitation CNNs face is a lack of interpretability, which leads to the failure in validation, regulatory compliance, and understanding of the model's decision-making process (Carta et al. 2025). To mitigate this limitation, researchers proposed XAIs that provide human-understandable insights. The most popular one includes saliency maps and heatmaps generated via gradient-based attribution methods, Grad-CAM, and Integrated Gradients. In FPADs, Grad-CAM visually presents regions of interest in fingerprint images, such as ridge anomalies or pore structures (Liu et al. 2024). Saliency maps help in improving generalization by using perceptual importance maps to direct the model's focus toward discriminative regions (Adam et al. 2025). Other approaches include LIME, which aids in local approximations and the integration of physical domain knowledge, such as sweat pore dynamics (Uttam et al. 2024). Another concerned point is that the lack of geometrical representativeness in training data may lead to generalization problems (Pereira et al. 2020). XAIs bridge this gap by explaining the decision boundaries and representativeness of learned features.

Recently, researchers have been investigating adversarial methods to improve detector resilience, as well as multimodal sensing. Techniques like GANs and Optical Coherence Tomography aid in increasing robustness (Spinoulas et al. 2021). XAI is now increasingly viewed as a core component to ensure trustworthiness, as it validates whether models attend to genuine fingerprint patterns rather than spurious background noise. Regardless of the widespread attention around XAIs, researchers treat them as a diagnostic tool rather than as an integral part of architectural innovation, as experiments are often performed in isolation, without in-model integration. This study closes this gap by developing an efficient model tailored for PAD and coupling it with multiple XAI analyses to visualize and understand its decision logic.

3. Methodology

This section provides a detailed overview of the method in two stages; the first stage covers all technical aspects, while the latter stage discusses how XAI analysis aids in interpreting the decision process of our proposed method (Figure 1).

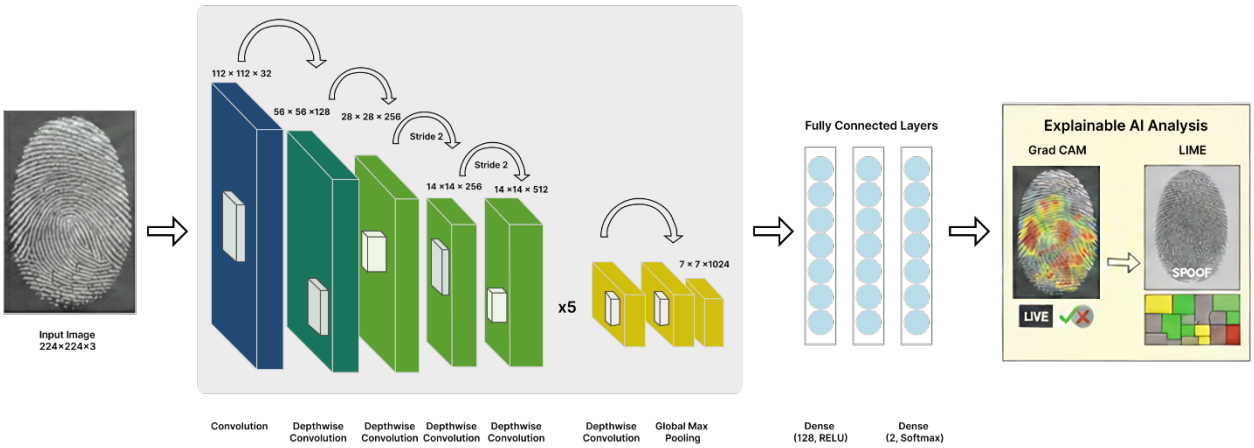


Figure 1. Overview of the Proposed Approach's Workflow.

3.1. Data Preprocessing

In this study, the benchmark used is the LivDet 2011 cooperative dataset, which has a large amount of variation in quality for the various sensors, potentially impairing model performance. To address this, we used a two-stage pre-processing pipeline to isolate discriminative ridge patterns and to reduce sensor-induced noise. The first stage enables preserving overall brightness, ensuring the integrity of critical fingerprint minutiae via a customized edge-sharpening kernel. Next stage, selective illumination normalization is applied in the HSV colour space to enhance the visibility in low light regions by increasing the value of the V channel by 25%. This will help the model to concentrate on biologically meaningful features and not the artifacts of the sensors.

3.2. Feature Extractor

A feature extractor is proposed to ensure accurate classification and effective feature discrimination, illustrated in Fig. 1, that leverages MobileNet's computational efficiency and the robust representational capacity. The core of the feature extractor belongs to depthwise separable convolutions represented in Eq. 1, which aids MobileNet to extract discriminative features, resulting in reducing parameters without sacrificing representational power, making it suitable for real-time applications. Depth-wise convolutions used for feature extraction are defined as:

$$Y = \sum_{k=1}^{M} (X_k * W_k).P_k \tag{1}$$

Here, X_k represents the input feature map's k^{th} channel. Meanwhile, W_k is the depthwise convolution filter applied to the k^{th} channel. Multiplying them successfully captures spatial features within each channel independently. Feature extraction goes from low-level, i.e., basic ridge edges and textures, to high-level features, including pore density and surface irregularities. Depthwise convolutions help in the model's effective learning while conserving spatial details, resulting in feature maps having discriminative characteristics. These maps are then fed to the customized classification head for decision-making.

3.3. Classification Head

To ensure our model provides outputs classified correctly, we incorporated the customized classification head illustrated in Fig. 1. The head provides classification based on the fed feature maps, which contain batch normalization (BN), dense, dropout, and softmax layers. BN layers balance feature scaling, dense layer (128 units) learns complex relationships among the extracted features with ReLU, and dropout is also incorporated to eliminate overfitting, resulting in better classification. Softmax, represented in Eq. 2 at the final stage, makes the classification decision by converting learned representations into probability scores. We can define Softmax as:

$$y = \frac{e^{I_c}}{\sum_{j=1}^{K} e^{I_j}} \tag{2}$$

Where e^{I_c} exponentiates the input value I_c, ensuring positive values, $\sum_{j=1}^{K} e^{I_j}$ represents the sum of all exponentiated input values, and y is the final output that presents the probability of the input belonging to the class c. The proposed work pipeline is illustrated in Fig. 1.

3.4. Optimizer

To investigate the optimizer that best fits the concerned model, we take AdaMax, the variant of the most popular Adam, as it combines the benefits of momentum optimization and adaptive learning rates, making it robust and stable for PAD. While the model is being trained, AdaMax guides the feature extractor to learn discriminative features rather than overfitting to noise. The Adamax update rule can be defined as:

First moment (mean) update:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (3)$$

Exponential moving infinity norm:

$$u_t = \max(\beta_2 u_{t-1}, | g_t |) \quad (4)$$

Bias correction for:

$$m_t: \widehat{m}_t = \frac{m_t}{1-\beta_1^t} \quad (5)$$

Parameter update:

$$\theta_{t+1} = \theta_t - \frac{\alpha}{u_t} \widehat{m}_t \quad (6)$$

In Eqs. (3-6), α represents the learning rate, $g_t = \nabla_\theta J(\theta_t)$ is the gradient at step t , while β_1, β_2 are Exponential decay rates and u_t approximate the infinity norm. In Eq. (3), m_t smooths out gradient noise, while Eq. 4 ensures stable updates forbidding multiplication with a small number, especially when gradients vary widely in magnitude. Unlike Adam, which scales updates by a running average of squared gradients, AdaMax scales them by the largest recent gradient magnitude, making it more robust to outliers.

3.5 Explainability Analysis

To ensure transparency of the decision-making process, we run the explainability analysis over the model output. This helps in interpreting the model's decision-making and eliminates the opaque nature of CNNs. XAIs are particularly categorized into model-specific and model-agnostic approaches. Model-specific techniques use the internal architecture of the model to draw insights, making it restricted to that model; however, model-agnostic techniques are significantly more flexible as they can draw insights from any model regardless of its underlying complexity. In this work, we employ multiple XAIs, Gard-CAM, and LIME, one from each category, as illustrated in Fig. 1.

In the proposed framework, the feature extractor determines cues from the input image that are influentially important; in our case, ridge details and pore density. Coupled with a classifier that determines how much each extracted feature contributes to the final decision. Integrated XAI analysis aids in visualizing this interaction by highlighting specific regions of inputs that strongly influenced the classifier's decision. In our method, we selected XAI methods that not only handle fine-grained texture analysis but also keep it computationally efficient. Grad-CAM, a model-specific method, takes architecture into consideration and provides the visual heat maps without requiring the computational overhead. Alongside LIME, a model-agnostic technique, provides insights about the data on which the model is trained without caring about the architecture. Consequently, ensures that the model's decisions are not arbitrary, but grounded in biologically and visually meaningful fingerprint regions.

3.6 Implementation Details

The proposed model is trained using Mobilenet weights on the LivDet 2011 (David et al. 2012) dataset split into training, testing, and validation sets at a ratio of 65:25:10, ensuring balanced representation. Before feeding images into the model, input samples are resized to 224 × 224 pixels and normalized within the [0, 1] range. To achieve further performance gains, we integrated an image enhancement, which deliberately refines edge clarity and boosts local contrast, further optimizing g the visibility of fine structural details. Consequently, these amplified textural cues allow our method to effectively isolate the subtle anomalies critical for fingerprint spoof detection. The ImageDataGenerator utility from Keras was used to handle image augmentation and normalization. We integrated Adamax with a learning rate of 0.0001, leveraging adaptive gradient scaling. The model was trained using a categorical cross-entropy loss function, as the task has two output classes over 15 epochs. The loss simplifies to:

$$L = -\frac{1}{N} \sum_{i=1}^{N} [\, y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (7)$$

Where y_i is the predicted probability that sample i belongs to class c (output from softmax layer), $c = 1$ if sample i belongs to class c, and 0 otherwise (one-hot encoded true label), N is the total number of samples, and C is the number of classes.

Table 1. Livdet2011 Dataset Summary

Sensor	**Training / Testing (Real/Fake)**	**Image Size**	**Resolution (in dpi)**	**Spoofing Materials**
Biometrika	(1000/1000) (1000/1000)	312×372	500	EcoFlex, Latex, Silgum, Gelatin, WoodGlue
Digital Persona	(1004/1000) (1000/1000)	355×391	500	Silicone, Latex, Play-Doh, WoodGlue, Gelatin
Italdata	(1000/1000) (1000/1000)	640×480	500	Latex, EcoFlex, Gelatin, WoodGlue, Silgum
Sagem	(1008/1008) (1000/1036)	352×384	500	Latex, Gelatin, WoodGlue, Play-Doh

4. Dataset

Experiments were conducted on the LivDet2011 (David et al. 2012) dataset, visualized in Figure 2, which contains live and spoof fingerprint images collected from multiple sensors and diverse spoofing materials. Thus, introducing significant variability during training and testing. This diversity partially simulates real-world deployment scenarios. The details are shared in Table 1.

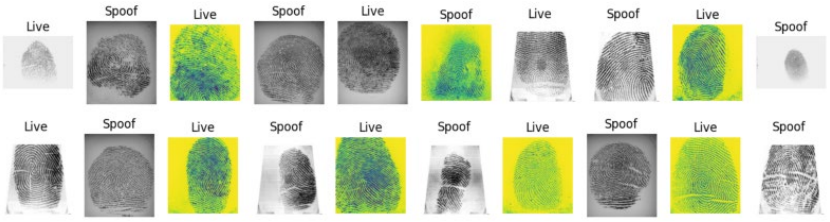


Figure 2. Dataset Images from Multiple Sensors of LivDet 2011.

5. Results and Discussion

This section presents both the quantitative and qualitative results obtained from the proposed model on the LivDet2011 dataset. Moreover, details of the dataset are also discussed.

5.1 Performance Evaluation Metrics

The proposed method's performance is evaluated on standard ISO/IEC 30107-3 metrics, including IM Accuracy, Precision, Recall, F1-Score, APCER, BPCER, and ACER. The metrics are defined as: **(a) Accuracy** measures the proportion of correctly classified input images among all test dataset, **(b) Precision** indicates the measure of correctly classified input samples are actually have that label, **(c) Recall** provides us with the measure of model's ability to correctly classify inputs among all samples, **(d) F1-Score** provides a balanced measure of classification performance by combining precision and recall, **(e) APCER** rate of spoof fingerprint which are misclassified as live, **(f) BPCER** is the rate of bona fide fingerprint samples that are misclassified as spoof, **(g) ACER** computes the average of APCER and BPCER to provide a balanced evaluation of spoof detection. A distinct performance aspect is provided by each of the metrics, thus assuring both detection capability and practical robustness of the proposed method.

5.2 Performance Evaluation

This experiment aims to evaluate the proposed model's classification performance across the LivDet 2011 dataset, a suitable match for this analysis due to its diverse spoofing materials and multiple sensor types, which is split into a ratio of 65:25:10.

The experimental results achieve a test accuracy of approximately 99% with a batch size of 32, revealing robust classification performance of the proposed model. This enabled the model to learn discriminative features of the fingerprint, thus the model identifies live fingerprints reliably. Further substantiated by the high accuracy of

classification observed across the test data set, as indicated by the confusion matrix in Figure 3(a). The fine-tuned classification head's ability of capturing textural irregularities and ridge distortions characteristics is proved by the ACER of 0.19% and 99.81% accuracy. Table 2 shows a breakdown of the other performance metrics for a more detailed analysis.

Our proposed approach shows almost smooth convergence from the training and validation curve, as shown in Figure 3(b). The accuracy trends and losses observed across epochs are consistent, which indicates that the optimizer and regularization parameters employed were appropriate for maintaining a consistent learning behavior. Adamax significantly helps in stabilizing learning during fine-tuning of MobileNet layers. By reducing sensitivity to an outlier gradient that is often caused by different texture variations of the spoof image, in our results, Adamax guarantees a more stable gradient descent, thus producing consistent accuracy gains.

Table 2. Results on LivDet 2011

Accuracy	**Precision**	**Recall**	**F1 Score**	**APCER**	**BPCER**	**ACER**	**IMA**
99.9	99.87	99.82	99.84	0.20	0.18	0.19	**99.81**

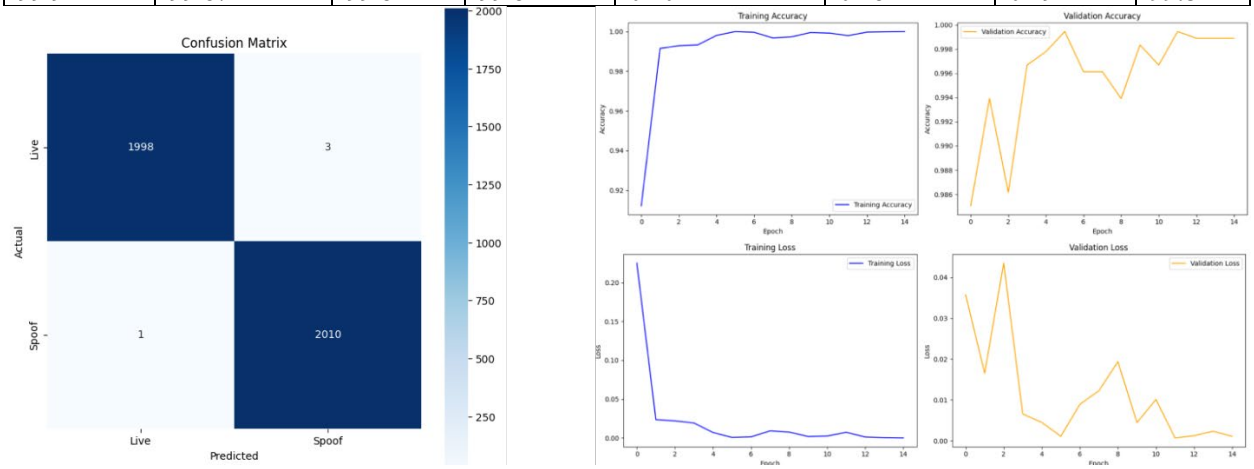


Figure 3. (a) Classification Accuracy via Confusion Matrix on LivDet 2011,
(b) Graphical Representation of Training and Validation Trajectories for both Accuracy and Loss, Indicating Smooth Convergence.

5.3 Cross-Sensor Evaluation

The experiment's objective is to examine the robustness and generalization ability of the proposed framework. To this end, a cross-sensor evaluation was carried out on LivDet 2011. In this experiment, the data from one sensor was fully excluded from the training process and only used for testing; the other three sensors were used for the training process, thus successfully replicating more challenging and realistic deployment environments. This is adopted to validate the model's transferability to an unseen sensor domain and acquisition scenarios. The most accurate was the Sagem sensor with an accuracy of 70%. A significant drop in accuracy can be seen in Table 3, ranging from 53% to 70%, showing that the model is not only capturing fingerprint liveness features but also features dependent on the sensor. When tested on unseen sensor images, there are variations in image resolution, image sensing technology, acquisition parameters, and the physical characteristics of the spoof material, resulting in a large domain shift that negatively affects the classification. However, our model scores competitive accuracy across all sensor domains, and Sagem scores 70% shows that there is some transferability across sensor domains.

Table 3. Cross Sensor Evaluation on LivDet 2011

Testing Sensor	**Training Sensors**	**Accuracy (%)**
Biometrika	Digital Persona + Italdata + Sagem	53
Italdata	Biometrika + Digital Persona + Sagem	60

Digital Persona	Biometrika + Italdata + Sagem	55
Sagem	Biometrika + Italdata + Digital Persona	70

This experiment provides some insight into the limitations of FPADs. Achieving high performance on a benchmark dataset does not guarantee strong cross-sensor generalization. The difference in performance between sensors suggests that some sensors have more common characteristics than others. The relatively high accuracy achieved by the Sagem sensor shows that learned representations have partial transferability between some sensor configurations, while the accuracy results of the Biometrika and Digital Persona sensors show higher sensitivity to sensor-centric variations. Despite competitive results, robustness against unseen sensor domains remains an open challenge.

5.4 Evaluation of Explainability Analysis

To ensure the model's decision is explainable, XAI techniques were applied. One is model-specific, i.e., Grad-CAM, while the other belongs to model-agnostic, i.e., LIME. The results reveal which regions most influenced the model's decision, providing valuable insights into the internal feature representations learned during training.

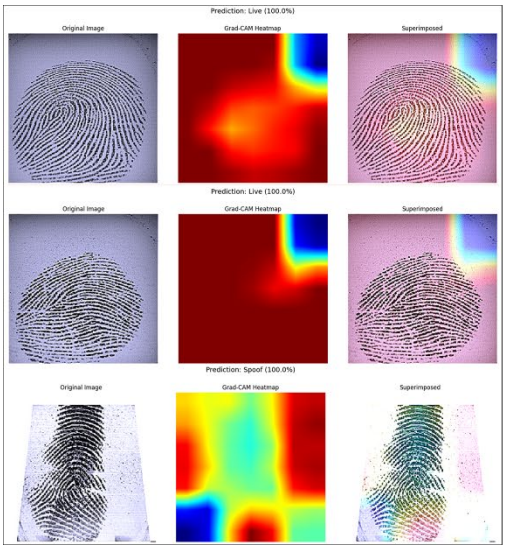
5.4.1 Analysis via Grad-CAM

We employ Grad CAM analysis to make our model transparent in nature. We can define Grad-CAM as follows:

$$L^{c}_{Grad\text{-}CAM} = ReLU\left(\sum_{k} \alpha^{c}_{k} A^{k}\right) \quad (8)$$

Where A^{k}: activation map of the k^{th} convolutional channel, α^{c}_{k}: Importance weight for class c and channel k, y^{c}: Output score for class c before the softmax layer, Z: number of spatial locations (width × height), and $\frac{\partial y^{c}}{\partial A^{k}_{ij}}$: Gradient of the score for class c with respect to the feature map A^{k} at position (i, j).

In this work, the Grad-CAM was adapted to produce visual heatmaps. Grad-CAM works by calculating the gradient of the target class score with respect to the final convolutional layer's feature maps. The resultant weighted mixture highlights the spatial areas that are essential in the model's decision-making. Furthermore, ReLU ensures that only features that have a positive impact on the prediction of a class are visualized. The resulting heatmaps presented in Figure 4(a) uncover the fingerprint regions where the proposed model mostly focuses, i.e., ridge continuity, pore density, and edge curvature, when identifying live samples and making decisions. On the flip side, spoof images prompt high activation in regions with surface inconsistencies, glossy patches, and artificial ridge deformation. Observed findings validate that the classification head results in guiding model attention towards genuine biometric traits instead of irrelevant background noise by enhancing feature localization.



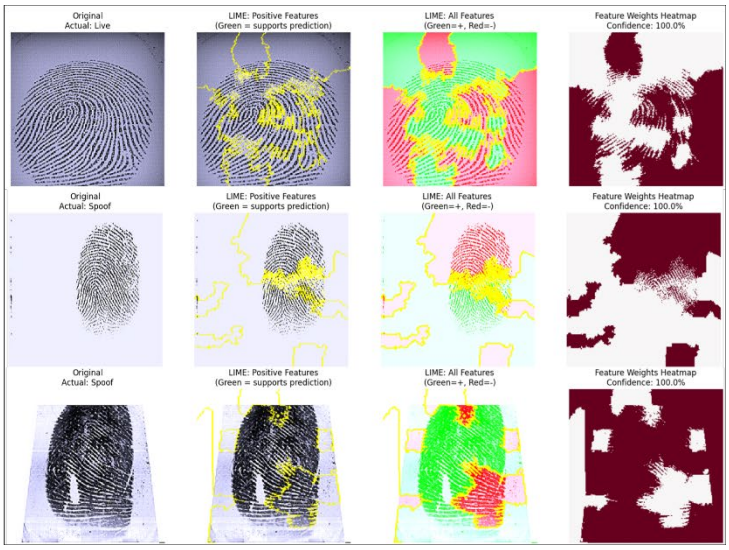


Figure 4. (a) Grad-CAM illustrating regions responsible for liveness classification. Red and yellow regions in Grad-CAM indicate highly influential fingerprint structures. (b) In Lime, green superpixels represent areas supporting the model prediction.

5.4.2 Analysis via Local Interpretable Model Agnostic Explanations (LIME)

The LIME framework was used to approximate the model's local decision by generating perturbed samples of the input, observing the model's responses, and learning a locally faithful interpretable surrogate model. For each test image, LIME highlighted super pixels that either supported or opposed the model's prediction. In other words, LIME approximates the behavior of a complex model, defined as:

$$\xi(x) = \text{argmin g} \in \text{G } L(f, g, \pi_x) + \Omega(g) \quad (9)$$

Where G represents the family of interpretable models, such as linear models or decision trees, $L(f, g, \pi_x)$ is the loss function that measures how well the interpretable model g approximates the original model f near the input x, π_x is a locality kernel that defines the neighborhood around the input sample x, and $\Omega(g)$ is a regularization term that enforces simplicity in the model g to ensure human interpretability.

The resulting LIME visualizations in Fig. 4(b) provided complementary evidence by segmenting the input image into super pixels and evaluating their influence on the model's output. Green segments on the images represented decision-supportive regions, while red segments opposed the specific decision. For live samples, supportive regions were mainly concentrated along continuous ridge structures, whereas for spoofed images, LIME highlighted uneven or broken ridge patterns and material boundaries.

LIME explains that genuine ridge patterns are evidence of live fingerprints, whereas irregular surface patches are of spoofed ones. Consistency can be seen from coinciding results of LIME and Grad-CAM analysis, indicating that both model-specific and model-agnostic explanation techniques highlight similar discriminative fingerprint regions, which indicates that the classifier primarily relies on biologically meaningful ridge structures rather than background artifacts. Thus provide a tradeoff between performance, optimization, and interpretability. Classification head aids in refining feature discrimination, while incorporated XAI certifies that the model's focus regions align with insightful biometric traits. Our approach includes both the functional and ethical dimensions of AI-driven PAD.

5.4 Comparative Analysis

To ensure belief in accuracy, a comparison against recent state-of-the-art PAD methods is made in Table 4. All of them are trained over the standard LivDet 2011 benchmark dataset. Our proposed model, with an IM Accuracy of 99.81%, outperforms all the methods, while Jian et al. (Jain et al. 2021) come second with the accuracy reaching around 96%. (Ali et al. 2025) performed the lowest of all with an IM Accuracy of 92%. The comparison ensures the proposed model not only outperforms the SOTA method but also boosts trust and practical reliability in real-world scenarios via integrating XAI analysis.

Table 4. Comparative Analysis With Existing Methods

Methods by Other Authors, Year	**Integrated Matching Accuracy (%)**
Jian et al., 2021	96.2
Contreras et al., 2022	93.17
Mehboob et al., 2023	95.61
Li et al., 2024	94.84
Ali et al., 2025	92.00
(Proposed Method)	**99.81**

6. Conclusion

This paper has proposed a novel explainable CNN deep learning framework for FPAD that enables effectiveness and explainability in fingerprint PAD. Experiments on the benchmark LivDet 2011 dataset indicate that the proposed method is effective in detecting presentation attacks. It also confirms that the structural modifications and the chosen optimization strategy are the factors that influence the model's performance. However, model robustness and generalization remain an open challenge. Further, explainable AI analysis validated the model classified based on important fingerprint features rather than spurious cues. The observed findings confirm that careful architecture refinements, when integrated with XAI, result in a reliable, explainable biometric system. Potential future directions involve training with diverse and larger datasets, integration of a hybrid interpretable framework with an attention

module, and investigating domain-aware learning strategies to improve generalization. Additionally, exploring multi-model biometric modalities within a single explainable mixture of experts framework could be a great contribution.

References

Agarwal, D., and Bansal, A., "Fingerprint liveness detection through fusion of pores perspiration and texture features," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 7, pp. 4089-4098, 2022.

Ali, M. S., Akram, A., Rashid, J., Jaffar, M. A., Shah, D., Ali, S., and Tahir, M., "Fake fingerprint classification using hybrid features learning with gradient boosting," *Applied Computational Intelligence and Soft Computing*, vol. 2025, no. 1, p. 8442143, 2025.

Buhrmester, V., Münch, D., and Arens, M., "Analysis of explainers of black box deep neural networks for computer vision: A survey," *Machine Learning and Knowledge Extraction*, vol. 3, no. 4, pp. 966-989, 2021.

Carta, S., Casula, R., Orrù, G., Micheletto, M., and Marcialis, G. L., "Interpretability of fingerprint presentation attack detection systems: a look at the “representativeness” of samples against never-seen-before attacks," *Machine Vision and Applications*, vol. 36, no. 2, p. 44, 2025.

Contreras, R. C., Nonato, L. G., Boaventura, M., Boaventura, I. A. G., Dos Santos, F. L., Zanin, R. B., and Viana, M. S., "A new multi-filter framework for texture image representation improvement using set of pattern descriptors to fingerprint liveness detection," *IEEE Access*, vol. 10, pp. 117681-117706, 2022.

Heusch, G., George, A., Geissbühler, D., Mostaani, Z., and Marcel, S., "Deep models and shortwave infrared information to detect face presentation attacks," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 2, no. 4, pp. 399-409, 2020.

Howard, Andrew G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and Adam, H., "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.

Jian, W., Zhou, Y., and Liu, H., "Densely connected convolutional network optimized by genetic algorithm for fingerprint liveness detection," *IEEE Access*, vol. 9, pp. 2229-2243, 2020.

Li, J., Wang, Y., and Erhu Zhang., "Striver: an image descriptor for fingerprint liveness detection," *Signal, Image and Video Processing*, vol. 18, no. 11, pp. 8229-8239, 2024.

Liu, G., Zhang, J., Chan, A. B., and Hsiao, J. H., "Human attention guided explainable artificial intelligence for computer vision models," *Neural Networks*, vol. 177, p. 106392, 2024.

Mehboob, R., Dawood, H., and Dawood, H., "An encoded histogram of ridge bifurcations and contours for fingerprint presentation attack detection," *Pattern Recognition*, vol. 143, p. 109782, 2023.

Mohammad, N., Khatoon, R., Nilima, S. I., Akter, J., Kamruzzaman, M., and Sozib, H. M., "Ensuring security and privacy in the internet of things: challenges and solutions," *Journal of Computer and Communications*, vol. 12, no. 8, pp. 257-277, 2024.

Park, E., Cui, X., Nguyen, T. H. B., and Kim, H., "Presentation attack detection using a tiny fully convolutional network," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 11, pp. 3016-3025, 2019.

Pereira, J. A., Sequeira, A. F., Pernes, D., and Cardoso, J. S., "A robust fingerprint presentation attack detection method against unseen attacks through adversarial learning," *In 2020 International Conference of the Biometrics Special Interest Group (BIOSIG)*, IEEE, pp. 1-5, 2020.

Poojary, R., and Pai, A., "Comparative study of model optimization techniques in fine-tuned CNN models," *In 2019 International Conference on Electrical and Computing Technologies and Applications (ICECTA)*, IEEE, pp. 1-4, 2019.

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., and Chen, L. J., "Mobilenetv2: Inverted residuals and linear bottlenecks," *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4510-4520, 2018.

Shukla, A. K., Janmaijaya, M., Abraham, A., and Muhuri, P. K., "Engineering applications of artificial intelligence: A bibliometric analysis of 30 years (1988–2018)," *Engineering Applications of Artificial Intelligence*, vol. 85, pp. 517-532, 2019.

Spinoulas, L., Mirzaalian, H., Hussein, M. E., and AbdAlmageed, W., "Multi-modal fingerprint presentation attack detection: Evaluation on a new dataset," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 3, no. 3, pp. 347-364, 2021.

Tan, M., and Le, Q., "Efficientnet: Rethinking model scaling for convolutional neural networks," *In International Conference on Machine Learning*, PMLR, pp. 6105-6114, 2019.

Uchiyama, T., Sogi, N., Niinuma, K., and Fukui, K., "Visually explaining 3D-CNN predictions for video classification with an adaptive occlusion sensitivity analysis," *In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 1513-1522, 2023.

Uttam, A. K., Agrawal, R., and Jalal, A. S., "Handcrafted Feature-Based Fingerprint Presentation Attack Detection," *In 2024 International Conference on Intelligent Systems for Cybersecurity (ISCS)*, IEEE, pp. 1-4, 2024.

Webster, S., and Czajka, A., "Saliency-Guided Training for Fingerprint Presentation Attack Detection," *arXiv preprint arXiv:2505.02176*, 2025.

Yambay, D., Becker, B., Kohli, N., Yadav, D., Czajka, A., Bowyer, K. W., and Schuckers, S., "LivDet iris 2017—Iris liveness detection competition 2017," *In 2017 IEEE International Joint Conference on Biometrics (IJCB)*, IEEE, pp. 733-741, 2017.

Yambay, D., Ghiani, L., Denti, P., Marcialis, G. L., Roli, F., and Schuckers, S., "LivDet 2011—Fingerprint liveness detection competition 2011," *In 2012 5th IAPR International Conference on Biometrics (ICB)*, IEEE, pp. 208-215, 2012.

Yu, Z., Li, X., Shi, J., Xia, Z., and Zhao, G., "Revisiting pixel-wise supervision for face anti-spoofing," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 3, no. 3, pp. 285-295, 2021.