Computer Science • 26(3) 2025 https://doi.org/10.7494/csci.2025.26.3.6243

Atul Kumar Uttam
Rohit Agrawal
Anand Singh Jalal

ATTENTION-BASED MULTIPLE-REPRESENTATION METHOD FOR FINGERPRINT-PRESENTATION-ATTACK DETECTION

Abstract *Fingerprint biometrics are one of the most common authentication mechanisms; however, such systems are often compromised by presentation attacks that are made by presentation-attack instruments. Most fingerprint-presentation-attack-detection approaches show poor performance due to the large variations in the presentation-attack instruments and the limited feature representation of the input fingerprint. Therefore, this article proposes a hybrid model of shallow and deep features with multiple representations of input fingerprints. To obtain these shallow and deep features, we first enhanced the texture of the input fingerprint through a novel median adaptive local binary pattern filter and an existing binarized statistical image feature. After this, the input fingerprint image and two textured enhanced images are concatenated along with the channel dimension for multiple representations. Finally, an extended ResNeXt architecture with channel and spatial attention (EResNeXt) was used for relevant feature extraction and presentation attack detection. EResNeXt was evaluated in the LivDet-2015 and LivDet-2017 data sets, and ACEs (average classification errors) were obtained at 0.94 and 0.49, respectively.*

Keywords presentation attack detection, ResNeXt, channel attention, spatial attention, median adaptive local binary pattern

Citation Computer Science 26(3) 2025: 31–53

1. Introduction

To recognize a human being, different biometric traits like one's face, fingerprint, voice, etc. are widely used. Among these biometrics, fingerprint-based biometric systems are cost-effective and widely used; however, fingerprint biometrics systems are vulnerable to various forms of attacks [28]. In recent years, people without extensive knowledge of biometric systems have successfully performed spoofing attacks (presentation attacks) with fake fingerprints that have been created with different spoofing materials [33] [34]. These spoofing materials are also known as presentation-attack instruments (PAIs) [10]. Such types of fingerprint-presentation attacks make the existing fingerprint biometric system unreliable. To address these attacks on fingerprint biometric systems, it is recommended to incorporate a fingerprint-presentation-attack-detection (FPAD) module before the recognition system. FPAD-based solutions are categorized into two categories; software-oriented (SW), and hardware-oriented (HW) methods. SW-oriented methods are based on static features (such as image quality, ridge structure [45], perspiration [20], and texture analysis [1]) or dynamic features (such as finger pressure and the skin blanching of fingerprints) for presentation-attack detection. Moreover, these SW-oriented approaches rely on either feature extraction that is crafted manually or automatic feature extraction using deep-learning techniques. Although HW-oriented methods use thermal imaging [16], optical coherence tomography technology [22], infrared and short waves [35], and pulse oximeters are used to detect artificial fingerprints [29] based on liveliness properties. These methods are more complex to implement and integrate with the existing system and incur additional costs, which limit their usefulness and applicability. In contrast, anti-spoofing techniques that use software can distinguish between genuine and counterfeit fingerprints by utilizing conventional image-processing methods. However, SW-oriented FPAD methods face several issues, such as feature selection, generalization, and vulnerability to noise. Among these issues, generalizations due to novel PAIs pose a significant challenge. Since these techniques have poor generalizations due to the limited representations of input fingerprints, a new hybrid PAD method has been proposed in this article.

The major contributions of the proposed fingerprint-presentation-attack-detection method are as follows:

- The proposed approach first improves the input image by using two distinct texture descriptors: a novel median adaptive local binary pattern (MALBP) descriptor, and the binarized statistical image features (BSIF) descriptor [8]. This enhancement process aims to improve the texture representation within the image, thus avoiding conventional feature extraction from fingerprint data.
- We introduce a novel approach for representing an input fingerprint through multiple representations by converting the original image and its enhanced versions into a three-channel image. This step significantly aids in extracting a wide range of distinctive features from the input fingerprint.

- We propose the integration of the residual block of ResNeXt [38] with channel and spatial attention (also known as convolution block attention module [CBAM] [36]) for vital feature extraction. The channel-attention component prioritizes significant channels, while the spatial-attention component focuses on essential regions.
- These refined features are then directed into a global average pooling layer, thus facilitating their alignments with specific class categories. Following this, the average features that correspond to each category are input into soft-max layers, producing likelihood estimates that underpin the final classification decision.
- Our experiment revealed that combining multiple representations of input fingerprint enhanced the generalization performance. This improvement was particularly notable when these concatenated images along the channel dimension were processed through ResNeXt with CBAM.

The rest of the article is organized as follows: in Section 2, the related works are discussed; in Section 3, the detailed proposed method is presented; in Section 4, the results and analysis of the proposed work are given; and finally, in Section 5, the conclusion is discussed, along with potential future work.

2. Related work

In this research article, we focused on software-oriented approaches for FPAD. This section presents a brief overview of the SW-oriented FPAD approaches.

2.1. Handcrafted feature-extraction-based methods

The majority of the techniques that are covered in this section use a two-step process for FPAD. First, the fingerprint is enhanced for texture-related features; later, these features are utilized as input to a machine-learning-based model to categorize them. For instance, Nikam et al. [24] suggested a texture analysis for an FPAD challenge by utilizing local binary pattern (LBP) and wavelet analysis. Ghiani et al. [7] combined statistical analysis and textural analysis approaches for feature extraction. In order to decrease the dimensionality of the feature space and increase the discriminative capability of the retrieved features, they were binarized. The support-vector classifier was then fed features that were binarized. For feature extraction, Jia et al. [12] proposed a method that incorporated multi-scale analysis and texture analysis approaches. These properties were encrypted using the local ternary pattern (LTP) operator. Galbally et al. [6] proposed a method that used wavelet transform to break down a fingerprint image and LBP to extract the texture information. A method for extracting an image's histograms of invariant gradient (HoIG) characteristics was developed by Gottschlich et al. [9]; these features were used to train an SVM classifier that distinguished between real and fake fingerprints. Principal-component analysis (PCA) and multi-scale local phase quantization (LPQ) characteristics were employed by Yuan et al. [42]. Jiang et al. [13] employed co-occurrence matrix (CM) characteristics to detect phony fingerprints; additional methods included using features from

HoG [41], the difference matrix with neighborhood gray-tone (NGTDM) [3], the Weber local binary descriptor (WLBD) [37], and an ensemble of adaptive LBP (LABP) and uniform LBP (ULBP) with an SVM classifier [31] for FPAD. These methods involved extracting an image's texture or structural elements as features; after this, a machine-learning classifier was taught to differentiate between real and false fingerprints. A summary of all these methods is provided in Table 1. Most of the handcrafted feature-extraction methods that were discussed above have limitations in terms of their abilities to represent complex and diverse patterns in an input fingerprint; this could lead to a lack of generalization abilities across new PAIs that have not been used in the training process.

Table 1
Summary of handcrafted features for fingerprint-presentation-attack detection

Reference	Feature Descriptor	Data Set	Performance (ACE %)
Nikam et al. [24]	Texture and wavelet method using LBP	Custom data set	2.59
Ghiani et al. [7]	BSIF descriptor	LivDet 2013	Equal error rate 7.22
Jia et al. [12]	Local ternary pattern with multi-scale block (MBLTP)	LivDet 2011	9.77
Galbally et al. [6]	Image quality-based	LivDet 2009	8.23
Gottschlich et al. [9]	HIG-based texture descriptor	LivDet 2013	12.2
Jia et al. [11]	MSLBP	LivDet 2011	7.47
Yuan et al. [42]	LPQ and PCA	LivDet 2011	8.62
Jiang et al. [13]	Co-occurrence Matrix	LivDet 2009, LivDet 2011	6.8, 10.98
Agrawal et al. [3]	Neighborhood gray-tone difference matrix	LivDet 2011	3.25
Yuan et al. [41]	Gamma-corrected HoG features with SVM	LivDet 2013	5.27
Sharma et al. [31]	Ensemble LABP and ULBP with SVM	LivDet 2009, LivDet 2011, LivDet 2013	4.23, 3.83, 3.57
Xia et al. [37]	WLBD	LivDet 2011, LivDet 2013, LivDet 2015	5.96, 1.89, 9.67

2.2. CNN-based methods

Several researchers have proposed CNN-based methods for fingerprint-presentation-attack detection, thus eliminating the need for manual feature engineering. Nogueira et al. [25] was the first to apply CNNs for this purpose; they also compared the performance with LBP-based methods. Their results showed that CNNs outperformed

LBP; later, they also used transfer learning with VGG16 and AlexNet for liveness detection. Marasco et al. [21] applied CNN-based models such as CaffeNet, GoogLeNet, and Siamese networks. Chugh et al. [5] extracted minutiae-centered patches from fingerprints and used MobileNet-v1 for FPAD. Zhang et al. [47] presented a lightweight CNN called SlimResCNN (which was based on the ResNet model). Yuan et al. [43] proposed image-scale equalization to normalize input-fingerprint images to a consistent scale for improving the performance of CNN; this helped reduce the impact of image variability due to the differences in finger placements and pressures. The authors of [46] proposed a compact CNN architecture called FLDNet, which consisted of ten dense convolutional blocks with decreasing feature map sizes, followed by an average pooling and a softmax-activation function for classification. Liu et al. [19] proposed a global network and a local network; the global network processed the entire fingerprint image in order to extract global features, while the local network processed small image patches in order to extract local features. The concatenated features of the two sub-networks were then passed to a fully connected layer for classification. Agarwal et al. [2] conducted a relative study between handcrafted and deep features for PAD. The authors used transfer learning for deep features, while LBP and HoG were used for handcrafted feature-based methods. The findings demonstrated that the deep-learning-oriented FPAD methods generally outperformed the handcrafted feature-based methods in terms of accuracy and resilience. An overview of these CNN-based approaches is provided in Table 2, along with the data set.

Table 2
Summary of CNN-based methods for fingerprint-presentation-attack detection

Reference	Method	Pipeline	Data set	Performance (ACE %)
Nogueira et al. [25]	CNN, PCA with SVM	CNN with PCA and SVM	LivDet 2009, LivDet 2011, LivDet 2013	3.93, 6.45, 3.55
Marasco et al. [21]	CNN	CNN	LivDet 2013	3.4
Chugh et al. [5]	MobileNet-v1	Patch extraction (centered around minutiae) with CNN	LivDet 2011, LivDet 2013, LivDet 2015	1.67, 0.25, 0.97
Zhang et al. [47]	SlimResCNN	Ten-layer CNN with nine residual blocks	LivDet 2017	4.75
Yuan et al. [43]	Image-scale equalization	Five-layer CNN with image-scale-equalization layer for varying sizes of fingerprints	LivDet 2011, LivDet 2013	6.45, 3.7

Table 2 cont.

Reference	Method	Pipeline	Data set	Performance (ACE %)
Zhang et al. [46]	FLDNet	Five-layer residual – dense block is introduced	LivDet 2013, LivDet 2015	0.7, 1.76
Liu et al. [19]	Mobilenet-V3 base CNN	Global local-based rethinking module with score-level fusion	LivDet 2017	2.28

Deep-learning-based methods offer significant performance quality, but they need to be carefully fine-tuned for the FPAD task. These methods are also data-hungry, and deeper models face the problem of saturation. CNN models are computationally expensive and require pre-processing steps in order to achieve significant performance. Studies have shown that the performances of these models drop considerably in unknown PAI scenarios. In general, handcrafted feature-extraction approaches are more interpretable than deep-feature-extraction techniques. Handcrafted features often capture the distinctive qualities of the data, which can be used to understand underlying patterns and insights. In contrast, deep-feature-extraction algorithms develop sophisticated representations in a data-driven manner to read features and comprehend underlying patterns. Li et al. [17] examined the effect of the background in deep-learning-based FPAD methods and found that removing the background will significantly improve the deep-learning-based FPAD method's performance. Kaur et al. [15] utilized local texture features such as scale, translation, and the rotation invariant for fingerprint and iris PAD. Tang et al. [32] applied adversarial networks, style transfers, and diffusion models to generate fingerprint images that enhanced the detection of live fingerprints that were classified as fake.

2.3. Handcrafted-CNN-based hybrid methods

A few hybrid approaches that combine handcrafted features and deep features have been suggested for FPAD by different authors in recent years. Hybrid approaches can leverage the strengths of both methods and potentially overcome some of their limitations. One approach that was used by [4] and [44] was to use handcrafted features as input to a deep neural network, which could learn more-complex representations based on the handcrafted features. These approaches have been shown to increase the performance of FPAD – especially when the training data is limited. Another approach by [30] used a pre-trained deep neural network and LBP with BSIF for extracting deep and shallow features from an input-fingerprint image. The output of these shallow- and deep-feature-based methods was then an ensemble for FPAD. One major limitation of such methods is that these methods failed to learn more-complex and -diverse features that are robust against different types of PAIs; hence, we have proposed a novel hybrid approach for FPAD in this research article. In our FPAD method, multiple representations of fingerprints have been used as input to the extended ResNeXt architecture.

3. Proposed framework

The framework of the suggested method is represented in Figure 1, which is based on a hybrid approach that has the advantage of shallow and deep features. The proposed framework consists of two components: the texture enhancement of an input fingerprint, and feature extraction with classification. In the first component (i.e., the texture enhancement), multiple representations of a fingerprint were used as input to the ResNeXt-based CNN. The network can learn more-diverse and -complex features that are resilient to various PAI types by employing multiple representations. Additionally, we have proposed a median adaptive local binary pattern (MALBP) to improve the input fingerprint's texture. The first component produces two textured enhanced images: the first image was obtained by applying the novel MALBP method to the input image, while the second image was obtained by enhancing the texture of the input image by using the state-of-the-art binarized statistical image filter (BSIF) method. Furthermore, these three images (i.e. the input images, the MALBP-enhanced image, and the BSIF-enhanced image) were concatenated along the channel dimension. The second component was utilized to extract the salient features and classification. We extended the state-of-the-art ResNeXt [38] by combining the attention module in the convolution block (CBAM) [36] for extracting the salient features and classification. The concatenated image was further fed to the extended ResNeXt network in order to classify live or fake fingerprints. The detailed workings of these components are provided in the subsections below.

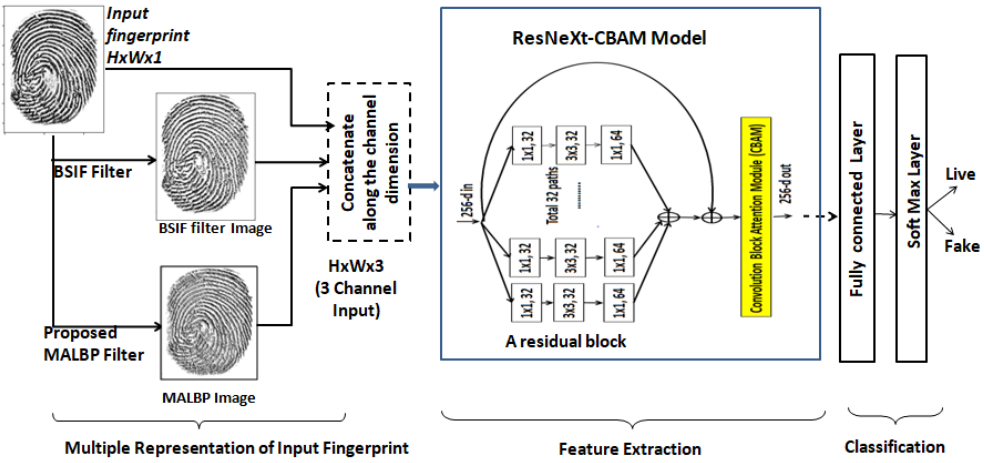

Figure 1. Block diagram of proposed EResNeXt-CBAM for FPAD

3.1. Texture enhancement

In this component, two textured enhanced images are extracted using the proposed MALBP and BSIF.

3.1.1. Median Adaptive Local Binary Pattern (MALBP)

LBP [26] is the most widely used and a simple-yet-effective filter in computer-vision applications for texture enhancement; it works by contrasting the pixel values of the local neighborhood that surrounds each pixel with its core pixel. However, the patterns that are extracted by LBP are affected by noise and illumination variations, thus resulting in a loss of discriminating power.

Figure 2. Original fake fingerprint (002_6_0_Fake) LivDet 2015 image (left) and MALBP enhanced image (right)

The proposed MALBP is an extension of LBP; this addresses these limitations. In MALBP, we computed the binary pattern of a pixel by comparing the pixel values of the 3×3 neighborhood around it with a robust median value instead of the central pixel value (as in [26]) or mean value [30]. This makes the proposed MALBP less sensitive to noise and illumination variations. Comparing the neighboring pixel values with the median or mean value (LABP) [30] has the advantage of reducing the impacts of outliers in the neighborhood as compared to the original LBP. Equations (1), (2), and (3) are used to compute the MALBP code:

$$\mathrm{MALBP}_{N,R} = \sum_{k=0}^{N-1} f(p_c - p_k)2^k \tag{1}$$

$$MT = \mathrm{median}(p_k, p_c) \tag{2}$$

$$f(d) = \begin{cases} 1, & \text{if } d > MT \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

where p_c and p_k are the intensity values of the core pixel, and the k^{th} pixel in radii R and N is the number of pixels in a given neighborhood. To enhance the texture features of the input fingerprint, we applied the MALBP texture descriptor by computing

the MALBP code from Equations (1), (2), and (3). Then, we applied a sharpening filter with a 3 × 3 kernel in order to enhance the contrast of the image. Finally, the fingerprint was normalized to set the pixel values to within a range of (0, 255). The overall effect of this process is given in Figure 2.

3.1.2. Binarized Statistical Image Features (BSIF)

BSIF [14] is a texture descriptor that encodes texture information (local) in an image into a binary string.

The BSIF filter is robust to noise and illumination changes, which can be problematic for other methods. The BSIF filter is also computationally efficient, making it suitable for real-time FPAD as shown by previous studies [7, 8, 30] and [18]. BSIF linear filters are trained on natural image patches, allowing them to adapt to a wide range of textures while reducing over-fitting. BSIF textured enhanced image I of size h,w is obtained by using Equation (4):

$$R_i = \sum_{h,w} W_i(h,w)I(h,w) = W_i^T I_i \tag{4}$$

$$B_i = \begin{cases} 1, & \text{if } R_i > 0 \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

where W_i is the linear filter of the same size. A binary image B of the same size as R is computed by threshold R_i at 0 using Equation (5). The sample input and output of the BSIF filter are shown in Figure 3.

Figure 3. Original fake fingerprint (002_6_0_Fake) image (left) and BSIF enhanced image (right)

3.2. Feature extraction and classification

The processed multi-channel enhanced image from component one is fed to the extended ResNeXt with the convolution block attention module (CBAM) for salient feature extraction and classification. The specifics of these are given in the subsections below.

3.2.1. Extended ResNeXt with CBAM

In the proposed framework, we utilized ResNeXt [38] as a feature extractor as opposed to the other deep-learning-based methods (such as DenseNet [4,30], etc.). The advantage of using the ResNeXt is that it can handle large amounts of data and high-level features.

Also, ResNeXt uses a split-transform-merge strategy, thus enabling it to acquire more-complex features and better capture variations in the data from multiple paths ($C = 32$). ResNeXt also utilizes the grouped convolution that captures local features in the input data. However, the topology of ResNeXt [38] is not designed for the FPAD task, so it has been modified (as shown in Table 3). In the first layer, a smaller filter size of 5×5 was used (originally 7×7), resulting in a smaller local receptive field; this means that extended ResNeXt can capture more local details in a fingerprint image.

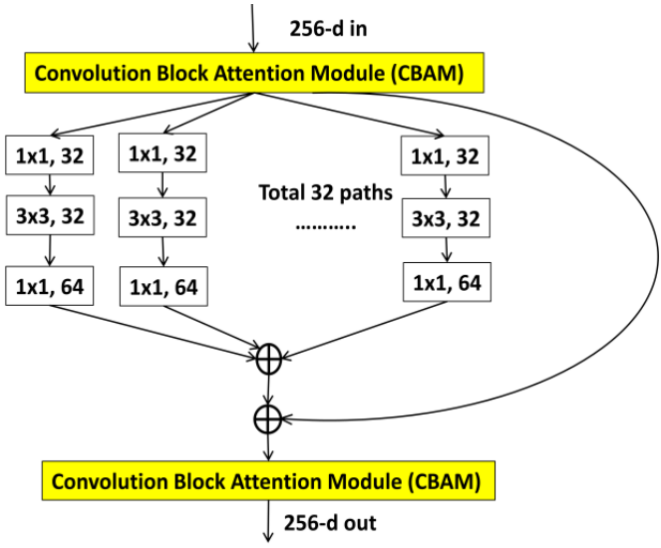

Figure 4. Residual block of proposed EResNeXt-CBAM

Furthermore, the sizes of the successive layers have been reduced to achieve faster computation and better feature representation (as shown in Table 3). This improves network performance by decreasing over-fitting and increasing generalization. Furthermore, ResNeXt has been extended by integrating the CBAM module in each

residual block. The CBAM module consists of channel and spatial attention (CA and SA) as two sub-modules that are applied sequentially. As presented in Figure 4, the CBAM module has been applied after each residual block to refine the feature maps and capture more of the fine-grained details of the fingerprint. CBAM discriminates important features from others by adding a very small number of additional parameters. The integration of CBAM to ResNeXt-50 significantly enhances its generalization performance over the new PAIs that were not seen at training time, as it focuses on only salient features. The proposed EResNeXt was inspired by ResNext-50. Considering the convolution block attention module as a single layer, the proposed model has 66 layers (Table 3. When presenting Conv 1 as the input layer, Conv 2 has 12 layers (including the max pool layer), Conv 3 has 16 layers, Conv 4 has 24 layers, Conv 5 has 12 layers, and Conv 6 has the output layer).

Global average pooling is applied to the output of the last residual block. After this, the feature vector is routed across two completely linked layers (FC), including 512 and 2 nodes, respectively. Using a soft-max activation function, the final layer produces the probability distribution across the two classes (false or real fingerprints). The complete topology of EResNeXt-CBAM is provided in Table 3. EResNeXt was inspired by ResNeXt-50, which is a CNN model that utilizes cardinality and skip connections.

Table 3
Layered architecture of proposed model

Stage	Output	Proposed EResNeXt-CBAM (32×4d)	
conv1	112 × 112	5 × 5, 64, stride 2	
conv2	56 × 56	3 × 3 max pool, stride 2	
		1 × 1, 32	×3
		3 × 3, 32, C = 32	
		1 × 1, 64	
		CBAM	
conv3	28 × 28	1 × 1, 64	×4
		3 × 3, 64, C = 32	
		1 × 1, 128	
		CBAM	
conv4	14 x 14	1 × 1, 128	×6
		3 × 3, 128, C = 32	
		1 × 1, 256	
		CBAM	
conv5	7 × 7	1 × 1, 256	×3
		3 × 3, 256, C = 32	
		1 × 1, 512	
		CBAM	
conv6	1 × 1	global average pool	
		2-d fc, softmax	
#parameters		12·10^6	

Let X be the input feature map of the output of the model that can be calculated by Equation (6):

$$z = x + \sum_{i=1}^{c} f_i(x) \tag{6}$$

where f_i represents the transformation function, which is a convolution operation followed by non-linearity.

3.2.2. Channel attention

Channel attention (CA) focuses on the most relevant channel features of a fingerprint image while ignoring noise or irrelevant details. CA has been described as a mechanism for selectively weighting CNN's feature maps [36]. Given an input feature map X that has a shape of channel (C), height (H), and width (W), the channel-attention map($\mathrm{C_{am}}$) takes the shape of channel (C), 1, 1 and is calculated from Equation (7) using averaged pooled features $\mathrm{F_{avg\_pool}}(X)$ and max pooled features $\mathrm{F_{max\_pool}}(X)$:

$$\mathrm{Cam} = \sigma\left(M\left(\mathrm{F_{avg\_pool}}(X)\right) + M\left(\mathrm{F_{max\_pool}}(X)\right)\right) \tag{7}$$

where M is a multi-layer perceptron (MLP) with one hidden layer (described as $M(z) = w_1(w_0(z))$), and w_0 and w_1 are the weight matrixes of sizes ($C/r \times C$) and ($C \times C/r$), respectively. Using average pooling and max pooling procedures, CA integrates spatial information from a feature map to produce two unique spatial context descriptors ($F_{avg\_pool}(X)$, and $F_{max\_pool}(X)$). After this, these descriptors are sent across a shared network that consists of a single multi-layer perceptron (MLP) with a hidden layer. The sigmoid activation keeps the attention weights between 0 and 1 and can be interpreted as a probability distribution over the channels. Furthermore, the output of the channel attention is fed to the spatial attention.

3.2.3. Spatial attention block

The approach employs spatial attention to target particular areas of a fingerprint picture that are prone to presentation attack artifacts, including irregular patterns or textures. This limited method lowers the possibility of false positives from unrelated areas of the picture, which can assist in increasing detection accuracy. Furthermore, spatial attention makes the model more robust to changes in fingerprint image appearance, such as resolution, illumination, or angle alterations. The spatial attention map (Sam) is calculated from Equation (8) using averaged pooled features $\mathrm{F_{avg\_pool}}(X)$ and max pooled features $\mathrm{F_{max\_pool}}(X)$ concatenated along with convolution, using a typical convolution layer of 7×7. Given a channel refined input X the spatial attention map Sam is calculated by Equation (8):

$$\mathrm{Sam} = \sigma\left(\mathrm{Cov}_{7\times7}\left(\left[\mathrm{F_{avg\_pool}}(X); \mathrm{F_{max\_pool}}(X)\right]\right)\right) \tag{8}$$

3.2.4. Channel and spatial attention (CBAM)

Given an input feature map X with a size of $C \times H \times W$, Equations (7) and (8) compute the Cam and Sam of sizes $C \times 1 \times 1$ and $1 \times H \times W$, respectively. The Cam and Sam are placed sequentially (as was suggested in [36]). The CBAM is used to calculate the entire attention module, which is calculated using Equations (9) and (10):

$$F_1 = \mathrm{Cam}(X) \otimes X \tag{9}$$

$$F_{11} = \mathrm{Sam}(F_1) \otimes F_1 \tag{10}$$

4. Results and analysis

We performed comprehensive tests on the benchmark LivDet-2015 [23] and LivDet-2017 [39] data sets to evaluate the proposed model (EResNeXt-CBAM). The data sets that were used in this study (mentioned in Sections 4.1. and 4.2) gave the usual assessment metrics for the EResNeXt-CBAM FPAD method's performance evaluation. Furthermore, the performance of the EResNeXt-CBAM in unknown PAIs is reported in Section 4.3.

4.1. Experimental setup

The proposed EResNeXt-CBAM model was trained on an Nvidia GeForce GTX TITAN X GPU with 128 GB RAM across 50 epochs. We chose to use the Adam optimizer with a batch size of 32 and an initial learning rate of 0.0001. Table 3 provides the specific architectural configuration.

4.2. Data sets

Table 4 provides information about the LivDet 2015 [23] and LivDet 2017 [39] data sets. The training set in the LivDet 2015 data set contained about 4500 real fingerprints and 4473 false fingerprints that were composed of Ecoflex, Latex, WoodGlue, BodyDouble, Playdoh, and Gelatin PAIs. The test data set had some unknown PAIs (LiquidEcoflex, RTV, OOMOO, and Gelatin) that were not involved in the training set. In the training set, the LivDet 2017 data set contained around 3000 live-fingerprint images and 5100 false-fingerprint images from three scanners. The test set included 5100 live-fingerprint images and more than 6000 false-fingerprint images. The PAIs that were used in the training (WoodGlue, Ecoflex, BodyDouble) and the test set (Gelatin, Latex, LiquidEcoflex) in the LivDet 2017 data set were different, making this data set ideal for evaluating the FPAD model for unknown PAIs or cross-material performance.

Table 4
Data set description

Data set	Scanner	Scanner abbreviation	Training set		Test set	
			#Live Images	#Fake Images	#Live Images	#Fake Images
LivDet 2015 [23]	Green Bit	GB_15	1000	1000	1000	1500
	Biometrika	Bio_15	1000	1000	1000	1500
	Digital Persona	DP_15	1000	1000	1000	1500
	Cross Match	CM_15	1510	1473	1000	1448
LivDet 2017 [39]	Green Bit	GB_17	1000	1200	1700	2040
	Orcanthus	Or_17	1000	1200	1700	2040
	Digital persona	DP_17	999	1199	1700	2028

4.3. Performance-evaluation metrics

To provide an equitable comparison with the most advanced techniques, we employed the average classification error (ACE) – a well-recognized assessment statistic that is included in the ISO standard [10] for bio-metrics. A smaller ACE value is desirable. Equation (11) was used to calculate the average rates of the false positives and false negatives for the fingerprints:

$$\text{ACE} = \frac{\text{sum of rate of misclassified live and fake fingerprints}}{2} \tag{11}$$

4.4. Performance evaluation

We evaluated the EResNeXt model using the data sets that are listed in Table 4 and compared the outcomes with the state-of-the-art findings to evaluate the model's effectiveness generalization.

4.4.1. Model tuning

We examined the performance of the suggested technique (both with and without the texture-enhancement stage) to determine the impact of this step. Table 5 illustrates the effect, with the texture-enhancement method reducing the model's overall ACE score in the LivDet 2015 data set by more than two-fold. This clearly showed that the texture-improvement phase that was presented in this study made use of several representations of fingerprint images, allowing the network to learn more-complex and -diverse features. We also assessed the performance of the baseline CNN ResNeXt without the CBAM layer. Table 5 strongly illustrates that EResNeXt with CBAM improved its performance – even with fewer parameters (as shown in Table 3).

Table 5
Comparisons of proposed model with and without texture-enhancement step that we proposed on LivDet 2015 [23] data set

Data set	Without texture enhancement	With texture enhancement	Without texture enhancement	With texture enhancement
	EResNeXt-CBAM		ResNeXt	
	ACE [%]	ACE [%]	ACE [%]	ACE [%]
GB_15	4.794	1.41	15.2	9.23
Bio_15	2.4624	0.72	9.11	6.65
DP_15	5.55	1.5	11.23	7.06
CM_15	1.13	0.15	8.56	5.38
Average	3.48	0.945	11.02	7.08

4.4.2. Cross-material performance discussion

The cross-material scenario was a practical setting where the generalization performance of the model was evaluated using unknown PAIs. It is important to note that the sensor did not change when the model was being tested and trained. The performance of the model was assessed using the same process as in earlier studies [4, 19, 27, 40]. In this scenario, the proposed model was evaluated using fingerprint photos that were not used to train the model. In the LivDet 2015 data set's Green Bit Sensor, for example, the training data set was comprised of fingerprints created from Ecoflex, Latex, Gelatin, and Wood Glue that were not tested (Liquid-eco-flex and RTV). Tables 6 and 7 provide a full overview of the training and testing materials that were used in the LivDet 2015 and LivDet 2017 data sets. When compared to the findings of the FAPD methods [4,27], and [40], EResNeXt-CBAM outperformed most of the methods in the LivDet 2015 data set. In all of the data sets from LivDet 2017, EResNeXt-CBAM outperformed the FPAD methods [4, 19], and [27].

Table 6
Comparative study in cross material scenario on LivDet 2015 data set [23]

Data set	Training material	Test material	EResNeXt-CBAM (ours)	Anusha et al. [4]	Yuan et al. [40]	Rai et al. [27]
			ACE	ACE	ACE	ACE
GB_15	Ecoflex, Wood-Glue, Gelatin, Latex	Liquid Ecoflex, RTV	1.4	2.82	1.03	1.32
Bio_15			0.7	2.57	0.8	3.23
DP_15			1.5	2.63	2	3.65
CM_15	Body Double, Ecoflex, Playdoh	OOMOO, Gelatine	0.15	0.37	0.31	1.61

Table 7
Comparative study on LivDet 2017 data set

Data set	Training material	Test material	EResNeXt CBAM (Our)	Anusha et al. [4]	Liu et al. [34]	Rai et al. [27]
			ACE	ACE	ACE	ACE
GB_17	Wood Glue, Ecoflex, and Body Double	Latex, Gelatine, Liquid Ecoflex	0.5	0.68	1.92	4.06
Or_17			0.28	0.03	1.67	6.19
DP_17			0.7	0.71	3.25	4.6

EResNeXt-CBAM performed better than the majority of the state-of-the-art techniques in the LivDet 2015 and LivDet 2017 data sets; this is demonstrated by the bar graphs (Figures 5 and 6). Our approach also performed better in terms of ACE on the majority of the data sets (GB_15, Bio_15, DP_15, CM_15) when compared to the other techniques (Anusha et al. [4], Yuan et al. [40], and Rai et al. [27]).

The LivDet 2017 data sets showed similar tendencies, with the suggested technique routinely achieving lower ACE values (thus, suggesting higher performance). In the cross-material scenarios, the proposed method outperformed the existing PAD methods [4, 27], and [40]. Figures 7 and 8 depict the ROC curves that indicated the output quality of the proposed FPAD approach. The suggested technique had a greater true positive rate (owing to the utilization of multiple representations of the input fingerprint) as well as greater channel and spatial attention in the ResNeXt base architecture.

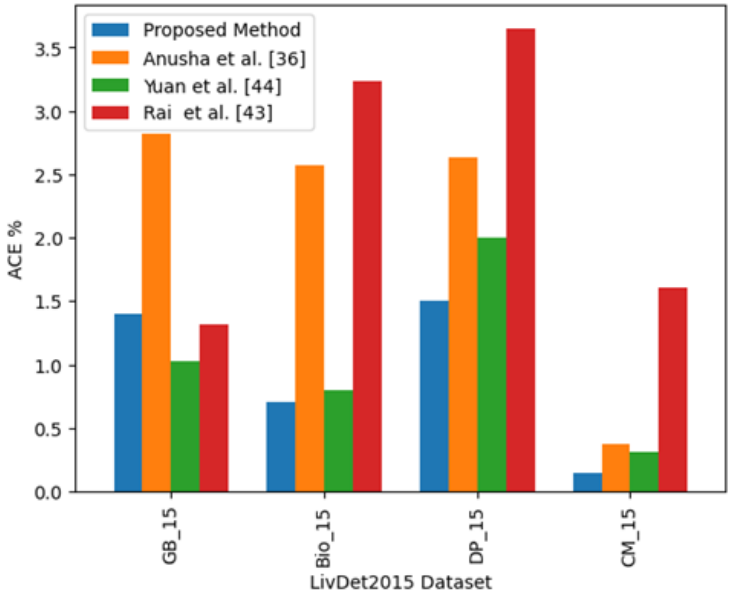

Figure 5. Bar graph comparison of LivDet 2015 data sets using suggested EResNeXt-CBAM technique with other FPAD methods

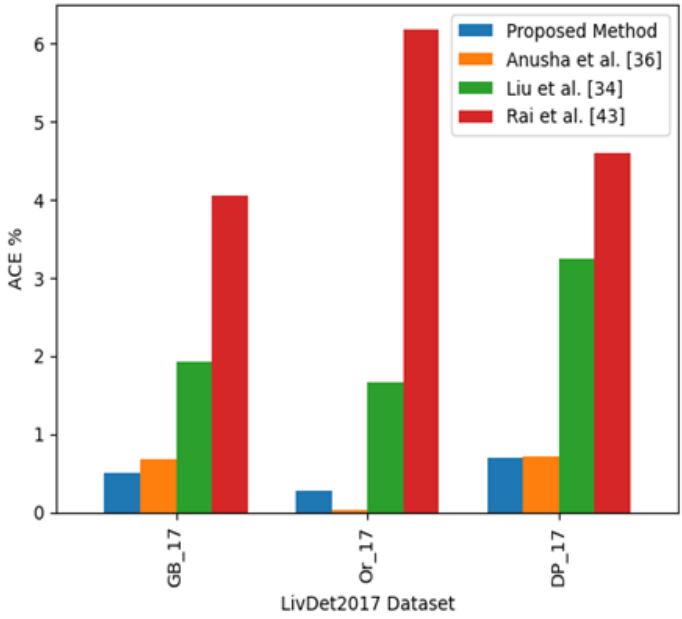

Figure 6. Bar graph comparison of LivDet (2017) data sets using suggested EResNeXt-CBAM technique with other FPAD methods

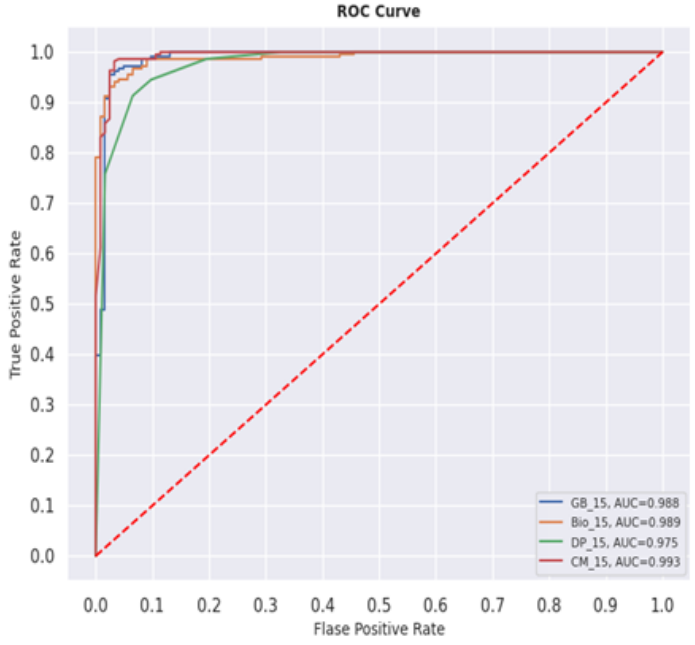

Figure 7. ROC curve analysis on LivDet2015 data sets using suggested EResNeXt-CBAM technique

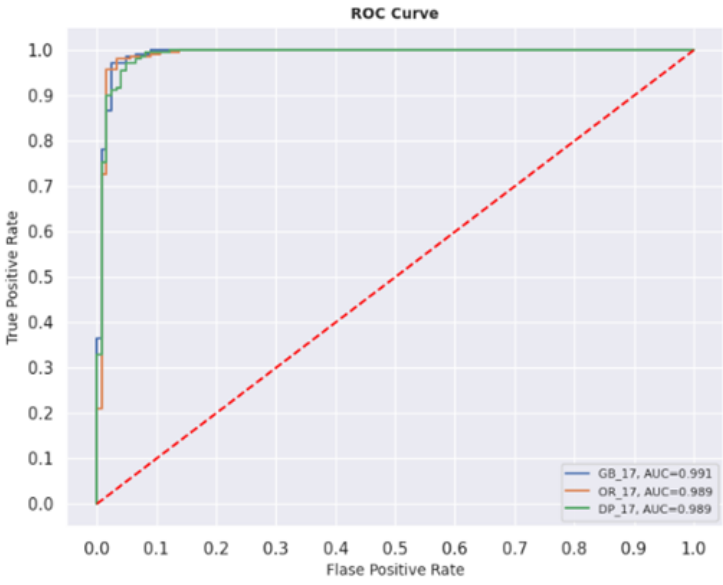

Figure 8. ROC curve analysis on LivDet2017 data sets using suggested EResNeXt-CBAM technique

5. Conclusion

To enhance the generalization performance in cross-material scenarios, the proposed EResNeXt-CBAM model first enhances the texture of an input fingerprint with the proposed novel MALBP and BSIF filters.

Compared to LBP, the proposed MALBP texture descriptor is more resilient to variations in illumination and noise, as it generates a binary pattern of a pixel by contrasting the pixel values of its neighboring area with a stable median value instead of using the value of the core pixel or mean value. These multiple representations of an input image are then concatenated along the channel dimension for extracting high-dimensional spatial features by passing them to the EResNeXt-CBAM model. ResNeXt was used as base CNN, as it utilizes skip connections to avoid vanishing gradient issues.

Furthermore, we extended ResNeXt with attention modules for identifying the most significant discriminative features of FPAD. When tested on the benchmark LivDet 2015 and LivDet 2017 data sets, the proposed model outperformed the previous state-of-the-art methods in terms of generalization performance. Future work may include extending EResNeXt-CBAM for cross-sensor and cross-data set performance to establish a data set-independent model.

References

[1] Agarwal D., Bansal A.: Fingerprint liveness detection through fusion of pores perspiration and texture features, *Journal of King Saud University – Computer and Information Sciences*, vol. 34(7), pp. 4089–4098, 2022. doi: 10.1016/j.jksuci.2020.10.003.

[2] Agarwal S., Rattani A., Ravindranath Chowdary C.: A comparative study on handcrafted features v/s deep features for open-set fingerprint liveness detection, *Pattern Recognition Letters*, vol. 147, pp. 34–40, 2021. doi: 10.1016/j.patrec.2021.03.032.

[3] Agrawal R., Jalal A.S., Arya K.V.: Fake fingerprint liveness detection based on micro and macro features, *International Journal of Biometrics*, vol. 11(2), pp. 177–206, 2019. doi: 10.1504/ijbm.2019.099065.

[4] Anusha B.V.S., Banerjee S., Chaudhuri S.: DeFraudNet: End2End fingerprint spoof detection using patch level attention. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 2695–2704, 2020. doi: 10.1109/wacv45572.2020.9093397.

[5] Chugh T., Cao K., Jain A.K.: Fingerprint spoof buster: Use of minutiae-centered patches, *IEEE Transactions on Information Forensics and Security*, vol. 13(9), pp. 2190–2202, 2018. doi: 10.1109/tifs.2018.2812193.

[6] Galbally J., Marcel S., Fierrez J.: Image Quality Assessment for Fake Biometric Detection: Application to Iris, Fingerprint, and Face Recognition, *IEEE Transactions on Image Processing*, vol. 23, pp. 710–724, 2014. doi: 10.1109/tip.2013.2292332.

[7] Ghiani L., Hadid A., Marcialis G.L., Roli F.: Fingerprint liveness detection using binarized statistical image features. In: *2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, IEEE, 2013. doi: 10.1109/btas.2013.6712708.

[8] Ghiani L., Hadid A., Marcialis G.L., Roli F.: Fingerprint liveness detection using local texture features, *IET Biometrics*, vol. 6(3), pp. 224–231, 2017. doi: 10.1049/iet-bmt.2016.0007.

[9] Gottschlich C., Marasco E., Yang A.Y., Cukic B.: Fingerprint liveness detection based on histograms of invariant gradients. In: *IEEE International Joint Conference on Biometrics*, IEEE, 2014. doi: 10.1109/btas.2014.6996224.

[10] ISO/IEC 30107-3:2023: *Information Technology – Biometric Presentation Attack Detection – Part 3: Testing and Reporting*, 2023.

[11] Jia X., Yang X., Cao K., Zang Y., Zhang N., Dai R., Zhu X., Tian J.: Multi-scale local binary pattern with filters for spoof fingerprint detection, *Information Sciences*, vol. 268, pp. 91–102, 2014. doi: 10.1016/j.ins.2013.06.041.

[12] Jia X., Yang X., Zang Y., Zhang N., Dai R., Tian J., Zhao J.: Multi-scale block local ternary patterns for fingerprints vitality detection. In: *2013 International Conference on Biometrics (ICB)*, IEEE, 2013. doi: 10.1109/icb.2013.6612964.

[13] Jiang Y., Liu X.: Spoof fingerprint detection based on co-occurrence matrix, *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 8(8), pp. 373–384, 2015. doi: 10.14257/ijsip.2015.8.8.38.

[14] Kannala J., Rahtu E.: BSIF: Binarized statistical image features. In: *Proceedings of the 21st international conference on pattern recognition (ICPR2012)*, pp. 1363–1366, IEEE, 2012. https://ieeexplore.ieee.org/document/6460393.

[15] Kaur B.: Fingerprint and Iris liveness detection using invariant feature-set, *Multimedia Tools and Applications*, vol. 83, pp. 60833–60859, 2024. doi: 10.1007/s11042-023-17854-w.

[16] Kowalski M.: A study on presentation attack detection in thermal infrared, *Sensors*, vol. 20(14), 3988, 2020. doi: 10.3390/s20143988.

[17] Li H., Ramachandra R.: Does Capture Background Influence the Accuracy of the Deep Learning Based Fingerphoto Presentation Attack Detection Techniques? In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 1034–1042, 2024. doi: 10.1109/wacvw60836.2024.00113.

[18] Li Q., Chan P.P.K.: Fingerprint liveness detection based on binarized statistical image feature with sampling from Gaussian distribution. In: *2014 International Conference on Wavelet Analysis and Pattern Recognition*, pp. 13–17, IEEE, 2014. doi: 10.1109/icwapr.2014.6961283.

[19] Liu H., Zhang W., Liu F., Wu H., Shen L.: Fingerprint presentation attack detector using global-local model, *IEEE Transactions on Cybernetics*, vol. 52(11), pp. 12315–12328, 2022. doi: 10.1109/tcyb.2021.3081764.

[20] Marasco E., Sansone C.: Combining perspiration-and morphology-based static features for fingerprint liveness detection, *Pattern Recognition Letters*, vol. 33(9), pp. 1148–1156, 2012. doi: 10.1016/j.patrec.2012.01.009.

[21] Marasco E., Wild P., Cukic B.: Robust and interoperable fingerprint spoof detection via convolutional neural networks. In: *2016 IEEE Symposium on Technologies for Homeland Security (HST)*, IEEE, 2016. doi: 10.1109/ths.2016.7568925.

[22] Moolla Y., Darlow L., Sharma A., Singh A., Merwe van der J.: Optical coherence tomography for fingerprint presentation attack detection, *Handbook of Biometric Anti-Spoofing: Presentation Attack Detection*, pp. 49–70, 2019. doi: 10.1007/978-3-319-92627-8_3.

[23] Mura V., Ghiani L., Marcialis G.L., Roli F., Yambay D.A., Schuckers S.A.: LivDet 2015 fingerprint liveness detection competition 2015. In: *2015 IEEE 7th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, 2015. doi: 10.1109/BTAS.2015.7358776.

[24] Nikam S.B., Agarwal S.: Texture and wavelet-based spoof fingerprint detection for fingerprint biometric systems. In: *2008 first international conference on emerging trends in engineering and technology*, pp. 675–680, IEEE, 2008. doi: 10.1109/icetet.2008.134.

[25] Nogueira R.F., de Alencar Lotufo R., Machado R.C.: Evaluating software-based fingerprint liveness detection using convolutional networks and local binary patterns. In: *2014 IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications (BIOMS) Proceedings*, pp. 22–29, IEEE, 2014. doi: 10.1109/bioms.2014.6951531.

[26] Ojala T., Pietikainen M., Maenpaa T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24(7), pp. 971–987, 2002. doi: 10.1109/tpami.2002.1017623.

[27] Rai A., Dey S., Patidar P., Rai P.: MoSFPAD: An end-to-end ensemble of MobileNet and Support Vector Classifier for fingerprint presentation attack detection, *Computer and Security*, vol. 148, 104069, 2023. doi: 10.1016/j.cose.2024.104069.

[28] Ratha N.K., Connell J.H., Bolle R.M.: Enhancing security and privacy in biometrics-based authentication systems, *IBM Systems Journal*, vol. 40(3), pp. 614–634, 2001. doi: 10.1147/sj.403.0614.

[29] Reddy P.V., Kumar A., Rahman S.M.K., Mundra T.S.: A New Antispoofing Approach for Biometric Devices, *IEEE Transactions on Biomedical Circuits and Systems*, vol. 2, pp. 328–337, 2008. doi: 10.1109/tbcas.2008.2003432.

[30] Sharma D., Selwal A.: HyFiPAD: a hybrid approach for fingerprint presentation attack detection using local and adaptive image features, *The Visual Computer*, vol. 38, pp. 2999–3025, 2022.

[31] Sharma D., Selwal A.: An intelligent approach for fingerprint presentation attack detection using ensemble learning with improved local image features, *Multimedia Tools and Applications*, vol. 81, pp. 22129–22161, 2022.

[32] Tang W., Figueroa D., Liu D., Johnsson K., Sopasakis A.: Enhancing Fingerprint Image Synthesis with GANs, Diffusion Models, and Style Transfer Techniques, *arXiv preprint arXiv:240313916*, 2024. doi: 10.48550/arXiv.2403.13916.

[33] The Times of India: Man 26 learns 'cloning fingerprints' online 'hacks' nearly 500 bank accounts, https://timesofindia.indiatimes.com/city/bareilly/man-26-learns-cloning-fingerprints-online-hacks-nearly-500-accounts/-with-bankmitrass-help/articleshow/81158623.cms, 2021. Accessed: 22.07.2024.

[34] The Times of India: How crooks are exploiting gaps in Aadhaar system in Delhi, https://timesofindia.indiatimes.com/city/delhi/how-crooks-are-exploiting-gaps-in-aadhaar-system-in-delhi/articleshow/98744284.cms, 2023. Accessed: 22.07.2024.

[35] Tolosana R., Gomez-Barrero M., Kolberg J., Morales A., Busch C., Ortega-Garcia J.: Towards fingerprint presentation attack detection based on convolutional neural networks and short wave infrared imaging. In: *2018 International Conference of the Biometrics Special Interest Group (BIOSIG)*, IEEE, 2018. doi: 10.23919/biosig.2018.8553413.

[36] Woo S., Park J., Lee J.Y., Kweon I.S.: CBAM: Convolutional Block Attention Module. In: V. Ferrari, M. Hebert, C. Sminchisescu, Y. Weiss (eds.), *Computer Vision – ECCV 2018. 15th European Conference, Munich, Germany, September 8–14, 2018, Proceedings, Part VII*, pp. 3–19, Springer International Publishing, Cham, 2018. doi: 10.1007/978-3-030-01234-2_1.

[37] Xia Z., Yuan C., Lv R., Sun X., Xiong N.N., Shi Y.Q.: A novel weber local binary descriptor for fingerprint liveness detection, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50(4), pp. 1526–1536, 2018. doi: 10.1109/tsmc.2018.2874281.

[38] Xie S., Girshick R., Dollár P., Tu Z., He K.: Aggregated Residual Transformations for Deep Neural Networks. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5987–5995, 2017. doi: 10.1109/cvpr.2017.634.

[39] Yambay D., Schuckers S., Denning S., Sandmann C., Bachurinski A., Hogan J.: Livdet 2017 – fingerprint systems liveness detection competition. In: *2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pp. 1–9, IEEE, 2018. doi: 10.1109/btas.2018.8698578.

[40] Yuan C., Jiao S., Sun X., Wu Q.M.J.: MFFFLD: A multimodal-feature-fusion-based fingerprint liveness detection, *IEEE Transactions on Cognitive and Developmental Systems*, vol. 14(2), pp. 648–661, 2021. doi: 10.1109/tcds.2021.3062624.

[41] Yuan C., Sun X.: Fingerprint liveness detection using histogram of oriented gradient based texture feature, *Journal of Internet Technology*, vol. 19(5), pp. 1499–1507, 2018.

[42] Yuan C., Sun X., Lv R.: Fingerprint liveness detection based on multi-scale LPQ and PCA, *China Communications*, vol. 13(7), pp. 60–65, 2016. doi: 10.1109/cc.2016.7559076.

[43] Yuan C., Xia Z., Jiang L., Cao Y., Wu Q.M.J., Sun X.: Fingerprint liveness detection using an improved CNN with image scale equalization, *IEEE Access*, vol. 7, pp. 26953–26966, 2019. doi: 10.1109/access.2019.2901235.

[44] Yuan C., Xia Z., Sun X., Wu Q.M.J.: Deep residual network with adaptive learning framework for fingerprint liveness detection, *IEEE Transactions on Cognitive and Developmental Systems*, vol. 12(3), pp. 461–473, 2019. doi: 10.1109/tcds.2019.2920364.

[45] Yuan C., Yu P., Xia Z., Sun X., Wu Q.M.J.: FLD-SRC: fingerprint liveness detection for AFIS based on spatial ridges continuity, *IEEE Journal of Selected Topics in Signal Processing*, vol. 16(4), pp. 817–827, 2022. doi: 10.1109/jstsp.2022.3174655.

[46] Zhang Y., Pan S., Zhan X., Li Z., Gao M., Gao C.: FLDNet: Light dense CNN for fingerprint liveness detection, *IEEE Access*, vol. 8, pp. 84141–84152, 2020. doi: 10.1109/access.2020.2990909.

[47] Zhang Y., Shi D., Zhan X., Cao D., Zhu K., Li Z.: Slim-ResCNN: A deep residual convolutional neural network for fingerprint liveness detection, *IEEE Access*, vol. 7, pp. 91476–91487, 2019. doi: 10.1109/access.2019.2927357.

Affiliations

Atul Kumar Uttam
GLA University, Department of Computer Engineering and Applications, Mathura 281406, India, atulkumaruttam@gmail.com

Rohit Agrawal
GLA University, Department of Computer Engineering and Applications, Mathura 281406, India, rohit.agrawal@gla.ac.in

Anand Singh Jalal
School of Computer Science & Information Technology, Devi Ahilya Vishwavidyalaya (DAVV), Indore, MP, India, anandsinghjalal@gmail.com

Received: 16.04.2024
Revised: 12.01.2025
Accepted: 1.07.2025