Journal of Theoretical and Applied Information Technology
31st March 2026. Vol.104. No.6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ADAPTIVE FEATURE EXTRACTION FRAMEWORK FOR ROBUST FACE ANTI-SPOOFING WITH CROSS-DOMAIN GENERALIZATION

KARTHIKA S[1] , G PADMAVATHI[2]

[1]Assistant Professor, Department of Information Technology,

Avinashilingam Institute for Home Science and Higher Education for Women, India

[2]Professor, Department of Computer Science,

Avinashilingam Institute for Home Science and Higher Education for Women, India

E-mail: [1]karthika_it@avinuty.ac.in, [2]padmavathi_cs@avinuty.ac.in

ABSTRACT

Face spoofing detection is crucial for the security of facial recognition systems; however, many existing methods struggle to generalize across varying acquisition conditions such as changes in lighting, camera angles, and multiple types of spoofing attempts. To address this challenge, this work proposes a novel feature extraction framework that combines three advanced components: Adaptive Kernel Generator (AKG), Discrete Style Assembly (DSA), and Adaptive Style Transfer (AST). AKG dynamically adjusts feature extraction based on instance-specific characteristics, enhancing sensitivity to subtle variations in spoofing attacks. DSA categorizes input samples into distinct style categories, enabling synthesis of style-specific features that are resilient to different presentation attack instruments (PAIs) and environmental conditions. AST further refines feature representations by adaptively transferring stylistic information from reference images, ensuring consistency and accuracy across diverse scenarios. The framework is built upon a modified ResNet-18 backbone optimized for single-channel face inputs, serving as the initial feature extractor before enhancement by the AKG, DSA, and AST modules. The model is evaluated on Replay-Attack, SiW-Mv2, and OULU-NPU datasets, each offering unique variations in spoofing scenarios. Experimental results demonstrate that the proposed model outperforms a baseline ResNet-18 model, achieving up to 94% accuracy in cross-dataset testing, highlighting its effectiveness and improved generalization performance in real-world face spoofing detection. Unlike conventional face anti-spoofing methods that rely on fixed feature extractors or classifier-level adaptation, this work introduces a style-aware, feature-level adaptation strategy that improves cross-domain generalization without requiring target-domain data.

Keywords: *Face Spoofing, Domain Generalization, Adaptive Kernel Generator, Discrete Style Assembly, Adaptive Style Transfer*

1. INTRODUCTION

Face recognition is now a critical aspect of biometric security, widely employed in applications such as smartphone unlocking, facial payment systems and access control [1]. Despite their convenience and non-intrusive nature, face biometric authentication systems are increasingly vulnerable to spoofing attacks, also referred to as presentation attacks (PAs). These attacks include print photo attacks, video replay attacks, 3D mask attacks, and sophisticated makeup techniques designed to deceive the system [2]. These vulnerabilities highlight the urgent need for robust Presentation Attack Detection (PAD), commonly referred to as Face Anti-Spoofing (FAS), to ensure the trustworthiness of face recognition systems. Accordingly, numerous techniques have been developed by the researchers and can be generally categorized into traditional machine learning methods and advanced deep learning methods [2]. These techniques aim to effectively distinguish between genuine and spoofed facial inputs, thereby enhancing the security of biometric systems.

Traditional FAS methods rely on handcrafted features, including texture-based, motion-based, frequency-based, and image quality-

based approaches. Texture features such as Local Binary Patterns (LBP) ([3], [4], [5]), Histogram of Oriented Gradients (HoG) ([6], [7], [8]), and Speeded-Up Robust Features (SURF) [9] have demonstrated limited success in detecting spoofing attacks. [10] proposed novel motion-based countermeasure using optical flow, rigid and non-rigid face motions. A face liveness detection method has been developed using high frequency descriptors and Fourier spectra in [11].

Although these methods have shown some success, they often struggle with poor generalization across different datasets due to the variability in capture devices and environmental conditions. More advanced techniques have emerged, leveraging Convolutional Neural Networks (CNNs) ([12],[13]) and Recurrent Neural Networks (RNNs) [14]. These methods leverage rich semantic information and robust feature representations, showing improved performance and generalization abilities.

However, even the most promising deep learning-based approaches face significant challenges when it comes to cross-dataset generalization [15]. Models trained solely on a single dataset often demonstrate poor performance when applied to new, unseen datasets. This performance degradation arises because models fail to capture variations in feature distributions caused by different presentation attack instruments and recording environments [16]. This dataset bias limits their effectiveness in real-world applications. The primary factor contributing to this challenge is the variability in facial texture patterns, chromatic variations, and the diversity of attack modalities, all of which result in substantial differences in feature distributions across various domains.

Based on these observations, this work identifies cross-domain robustness as the central concern of face anti-spoofing. Since variations in devices, illumination, and attack types mainly affect feature representations, this study focuses on adaptive feature extraction rather than classifier-level fine-tuning. By improving the stability and adaptability of extracted features, the proposed approach aims to achieve more reliable generalization across unseen spoofing scenarios.

Some researchers address these challenges using domain adaptation methods, which aim to improve performance in unseen scenarios by aligning feature distributions or learning domain-invariant representations. One limitation of domain adaptation techniques in face anti-spoofing is their reliance on labeled data from the source domain, which may not always be easily accessible or fully representative of the diverse variations found in the target domain. Researchers have tackled this issue using domain generalization techniques in face anti-spoofing, which involve training models across diverse source domains without requiring labeled data from specific target domains. This approach aims to improve model robustness and generalization to new scenarios, minimizing the dependency on domain-specific labeled data. Some of the prominent domain generalization techniques include Adversarial domain generalization [17], where models are trained across diverse source domains without specific target domain labels. Other approaches focus on aligning feature distributions and class-level distributions across domains to improve model robustness and generalization to new scenarios.

Additionally, representation learning methods intends to extract features that are consistent across domains, while meta-learning and transfer learning techniques leverage knowledge from related tasks or domains to enhance model adaptation to unseen datasets. Uncertainty modeling is also utilized to handle difficult-to-transfer samples and improve overall generalization performance across different domains. However, these existing domain generalization methods in face anti-spoofing struggle to effectively integrate diverse feature representations across multiple domains. This limitation reduces their robustness on unseen datasets, particularly due to inadequate modeling of class-level information and multimodal structures.

This research gap is addressed by introducing a novel approach that integrates Adaptive Kernel Generator (AKG), Discrete Style Assembly (DSA), and Adaptive Style Transfer (AST) for robust and context-aware feature extraction. Firstly, the Adaptive Kernel Generator (AKG) dynamically adjusts kernel parameters based on input data characteristics, facilitating more precise feature extraction that captures domain-specific nuances effectively. Secondly, the Discrete Style Assembly (DSA) generates diverse feature representations, accommodating variations introduced by different presentation attack instruments (PAIs). This allows the proposed model to be trained to generalize well across a wide range of spoofing set-ups, from print attacks to video

replays. Furthermore, Adaptive Style Transfer (AST) complements these techniques by adapting feature extraction methods to varying image qualities and environmental conditions. By integrating these adaptive mechanisms, the proposed framework enhances generalization beyond the training domains, improving robustness and reliability in real-world face anti-spoofing scenarios.

The proposed framework focuses not only on extracting features that are consistent across different domains but also on training a robust classifier capable of generalizing effectively across multiple domains. This dual approach guarantees that the model can perform robustly under diverse lighting scenarios and camera angles. The model's capability to manage variations in facial expressions and occlusions is also enhanced, significantly boosting the security and reliability of face biometric authentication systems. Extensive experiments conducted on multiple benchmark datasets showcase the advantage of this proposed method over current state-of-the-art techniques, underscoring its potential for effective application in practical environments.

The key contributions of the proposed framework include:

- Integration of Adaptive Kernel Generator (AKG), Discrete Style Assembly (DSA), and Adaptive Style Transfer (AST) improves the model's capability to capture subtle variations and style-specific features in face spoofing detection.
- Both feature-level differences and local-level variations are extracted for distinguishing between real and spoofed faces
- Cross-dataset validation is performed using the datasets Replay-Attack, SiW-Mv2, and OULU-NPU.

The remaining sections of this paper are organized as follows. The Related Work section discusses existing literature on face anti-spoofing and domain generalization techniques. The Proposed Framework section presents the adaptive feature extraction architecture. The Methodology section describes the data preprocessing, feature extraction, classification process, and evaluation procedures. The Evaluation and Results section reports and analyses the experimental findings. Finally, the Conclusion section highlights the key contributions of this study and outlines potential directions for future research.

2. RELATED WORK

2.1 Machine Learning And Deep Learning Based Anti-Spoofing Methods

Machine learning methods primarily focus on designing features from texture and using natural attributes of facial images and videos. For instance, optical flow values from face movements are used in [18] and classified them with Support Vector Machines (SVM). Similarly, color texture analysis for distinguishing real faces from spoofed ones is employed in [4].

To overcome the limitations of handcrafted feature–based methods, deep learning methods focus on extracting complex and distinctive features through advanced neural network architectures. Deep learning is applied in face anti-spoofing in different ways. These methods can be categorized into two types: deep learning models that extract features directly from face images, and approaches that first extract handcrafted features and then employ deep learning models for higher-level semantic representation. A deep Convolutional Neural Network in [19] learned highly discriminative features, greatly improving face anti-spoofing performance with data pre-processing.

A deep neural network proposed by the work [20] combined LSTM units with CNNs, where LSTM was used to model long-term dependencies and CNN was employed for extracting local features. [21] proposed a Deep Partial CNN that extracted deep partial features from convolutional layers and applied PCA for dimensionality reduction, achieving strong discrimination between real and fake faces. However, these methods often rely on dataset-specific feature patterns, which can reduce robustness when evaluated across unseen datasets.

A 3D CNN is introduced by [13] effectively learns spatiotemporal features from video frames for face anti-spoofing, leading to substantial enhancement in performance. However, the issue of complex backgrounds remains inadequately addressed. The work [22] presented the Spatio-Temporal Anti-Spoofing Network (STASN), which enhanced face anti-spoofing by using realistic data synthesis and capturing both global and local features.

A multi-cue integration framework for face anti-spoofing is proposed by [23] that

combines image quality and motion cues using a hierarchical neural network and achieved perfect discrimination on benchmark datasets. A new FAS approach is developed in [24] that uses Residual Spatial Gradient Blocks (RSGB) and Spatio-Temporal Propagation Modules (STPM) to capture detailed spatial and temporal features from multiple frames and attained cutting-edge performance. However, its reliance on depth information may limit its effectiveness in scenarios lacking reliable depth data.

An end-to-end FAS method with Anti-interference Feature Distillation (AFD), Global Spatial Attention Learning (GSAL), and pyramid binary mask supervision is proposed by [25] and accomplished better performance on both intra-dataset and cross-dataset tests. Despite their effectiveness, such models often involve increased computational complexity and require extensive training resources.

NAS-FA, a neural architecture search method for face anti-spoofing is developed by [26] which achieved state-of-the-art results across nine datasets and introduced a new cross-dataset testing protocol. Another research work [27] introduced DRL-FAS, a framework that uses CNN for global feature extraction and RNN for sequential analysis of local sub-patches, combined with deep reinforcement learning for optimization. A face anti-spoofing algorithm is introduced in [28] utilizing multi-scale perceptual image quality assessment features with SVM. Nevertheless, many of these approaches still experience performance degradation under cross-dataset evaluation due to domain shift.

2.2 Domain Generalization Method Face Anti-Spoofing Methods

Traditional end-to-end deep learning methods for face anti-spoofing often struggle with new conditions like different lighting, facial features, and camera qualities, as well as emerging attack methods like advanced masks. As a result, these methods may not be reliable for high-security applications, leading researchers to focus increasingly on improving how well these models generalize to varied scenarios. Hence, developing models with enhanced adaptability to diverse conditions and emerging attack types is crucial. To address these challenges, strategies such as domain adaptation and domain generalization have been explored. [29] presented SSDG, which enhances face anti-spoofing by making real faces compact and fake faces dispersed. Using asymmetric learning and triplet loss, SSDG achieves superior results on several datasets. A dual-branch framework is designed to extract camera-invariant features, where one branch performed high-frequency domain decomposition, while the other integrated high- and low-frequency information. This approach achieved improved generalization across various camera models, outperforming existing methods in both intra- and cross-dataset evaluations. However, these methods primarily focus on global feature alignment and provide limited consideration of class-level consistency.

A Shuffled Style Assembly Network (SSAN) is proposed in [30], which separates content and style features and uses adversarial and contrastive learning to enhance domain generalization in face anti-spoofing. [31] developed Dual Reweighting Domain Generalization (DRDG), which iteratively reweights samples and features to emphasize domain-invariant information, resulting in improved generalization performance.

A FAS method using negative data augmentation which is proposed by [32] achieving strong generalization without real-world attack samples. It outperforms existing techniques by using synthesized samples and specialized loss functions. A Quality-Invariant Domain Generalization (QIDG) is introduced by [33] which aligned liveness features into a quality-invariant space using a teacher-student model. By employing dual adversarial learning and quality feature assembly, QIDG improved generalization across domains and outperformed existing methods on five public datasets. Such approaches often rely on complex training strategies or synthetic data generation, which may limit their scalability in real-world settings.

A method SA-FAS is proposed by [34], a that learns domain-specific features while retaining decision boundaries consistent across domains by aligning live-to-spoof transitions, achieved strong cross-domain performance. [35] presented a framework that enhances face anti-spoofing by creating a generalized feature space for live faces and a domain-agnostic classifier using low-rank decomposition. By employing a Common Specific Decomposition (CSD-S) layer, the method demonstrated improved performance across four major datasets. Despite improved performance,

transferring difficult samples across domains remains a challenging problem.

Previous research by [36] has shown that feature correlations, represented by covariance matrices, capture domain-specific styles in images. As described by [37], the whitening transformation is used to remove these correlations by standardizing each feature to have unit variance. This technique has been demonstrated to effectively eliminate domain-specific styles in various applications, including image translation [38], style transfer [39], and domain adaptation [40]. More recently, [41] employed whitening loss to improve generalization in semantic segmentation. Despite its proven effectiveness, instance whitening has not been extensively explored in the context of domain generalization for face anti-spoofing (DG FAS).

Recent developments in adaptive intelligent systems emphasize the necessity of robust feature learning mechanisms capable of operating under heterogeneous and dynamically shifting environments. The proposed Adaptive Feature Extraction Framework directly addresses cross-domain generalization challenges in face anti-spoofing by modeling variations introduced by diverse presentation attack instruments, illumination changes, and acquisition devices. Similar to machine learning strategies applied for uncovering complex cybercrime patterns across distributed datasets [42], the present framework prioritizes resilient pattern extraction under domain variability. The integration of Adaptive Kernel Generator reflects principles of bio-inspired intelligent optimization, where dynamic adaptation enhances contextual decision-making and system robustness [43], [44]. The Discrete Style Assembly and Adaptive Style Transfer modules further manage feature distribution shifts by preserving discriminative content while harmonizing stylistic inconsistencies across domains. Secure biometric authentication infrastructures require such adaptive mechanisms, complementing advanced encryption-driven secure data sharing systems [45]. From a broader AI perspective, domain-resilient learning architectures contribute to sustainable and intelligent digital ecosystems that support scalable, trustworthy decision frameworks across sectors [46], [47]. Through dynamic kernel modulation, style-aware feature synthesis, and instance-level normalization, this research advances robust and generalizable biometric security modeling.

The current literature highlights the importance of aligning features across different domains for improved generalization; however, such approaches often neglect the alignment at the class level and the challenges associated with transferring difficult samples. The proposed framework seeks to fill these gaps by incorporating adaptive techniques, including instance whitening, to improve feature generalization and mitigate the influence of domain-specific characteristics and challenging samples on model efficacy.

The literature indicates that many face anti-spoofing methods struggle to generalize across domains due to variations in acquisition settings and attack types. Although deep learning and domain generalization techniques have improved performance, dataset bias and weak class-level consistency remain challenges. Accordingly, this work focuses on learning adaptive, domain-robust feature representations for reliable face anti-spoofing in unseen scenarios.

It is hypothesized that learning adaptive and style-aware feature representations at the feature extraction stage can significantly improve cross-domain generalization in face anti-spoofing systems, compared to conventional approaches that rely on fixed feature extractors or classifier-level adaptation.

3. METHODOLOGY

The detailed steps of the proposed methodology are clearly illustrated in Figure 1, providing a visual representation of the overall workflow and its key components. The proposed approach aims to reliably differentiate real faces from spoofed faces by learning strong and adaptable feature representations. First, input face images are preprocessed through resizing, augmentation, and histogram equalization. These processed images are then fed into a modified ResNet-50 model enhanced with the Adaptive Feature Extraction Framework (AFEF). AFEF consists of three modules: Adaptive Kernel Generator (AKG) Discrete Style Assembly (DSA), and Adaptive Style Transfer (AST)—which work together to highlight spoof-related cues while suppressing irrelevant variations. Finally, the trained model predicts whether an unseen face is real or spoofed.

3.1 Source Domains

Three public FAS datasets are used to evaluate the efficiency of the proposed framework: (i) Replay-Attack dataset [3], which includes video recordings from 50 subjects captured under various lighting conditions. This dataset includes 1300 videos in total; however, only 1200 (training, development, and test sets) are used for anti-spoofing evaluation, excluding the 100-video enrollment set. (ii) SiW-Mv2 dataset [48], which features videos captured from multiple angles and under different lighting conditions (iii) OULU-NPU dataset [49], which provides high-quality spoofing videos from numerous participants. Together, these datasets include diverse capture devices, types of attacks including print attacks and video replays, varied illumination conditions, backgrounds, and diverse racial diversity. The combined characteristics of these datasets introduce significant domain shifts, challenging face anti-spoofing methods to generalize effectively across different real-world scenarios. Further specifications, including device types, attack scenarios, and training-testing splits, are given in Table 1.

3.2 Preprocessing

The pre-processing steps are critical and pivotal for the proposed feature extraction framework.

3.2.1 Face detection, resizing and cropping

Face detection, achieved using a robust Haar cascades algorithm localizes facial regions within video frames, concentrating subsequent analyses on pertinent areas and enhancing computational efficiency. Resizing standardizes facial images to consistent dimensions, (W_{resize}, H_{resize}), ensuring uniformity across datasets and optimizing model performance by reducing variability in input sizes. Furthermore, cropping extracts the detected faces from frames, removing unnecessary background noise and concentrating the analysis on the most informative parts of the image. These steps collectively prepare input data in a structured manner, enhancing the framework's ability to extract robust features essential for accurate face recognition and spoofing detection.

3.2.2 Data augmentation

Applying appropriate data augmentation techniques is crucial to effectively prepare the datasets for enhancing the performance and robustness of CNNs in image classification tasks. In this study, a range of transformations are applied to the training data to make the model more adaptable and resilient. These transformations include Random Horizontal Flip, which flips images horizontally, Random Rotation that rotates images up to 10 degrees, and Random Resized Crop that adjusts the size of the images between 80% and 100% of the original size. These geometric transformations help the model recognize features regardless of their orientation or size. Additionally, color-based transformations like color jittering are used to adjust brightness, contrast, saturation, and hue of the images. This helps the model learn to identify objects under different lighting conditions and color variations, simulating real-world scenarios. The augmented images are converted into PyTorch tensors and used for network training. This augmentation strategy helps reduce overfitting while improving the model's ability to generalize to unseen data.

3.2.3 Histogram equalization

Histogram equalization is employed to enhance the contrast of images in the dataset, which is particularly beneficial for improving the performance of the model. This technique works by redistributing the pixel value intensities in a facial image to span the entire available range, thereby equalizing the histogram of the grayscale values. This process is mathematically defined by transforming the pixel intensity values I in the image using the cumulative distribution function (CDF) of the histogram. In this implementation, histogram equalization is used to standardize the lighting conditions across different images, ensuring that the features are more uniformly distributed. This helps in mitigating the effects of poor lighting conditions or shadows that might otherwise obscure important details in the images. By enhancing the global contrast of the images, this pre-processing step enables the ResNet-18 to better differentiate between subtle features, leading to more accurate feature extraction and classification.

Mathematically, if H(i) denotes the histogram of intensity level i, and P(i) represents the corresponding probability, then the cumulative distribution function is computed as in Equation 1:

$$C(i) = \sum_{j=0}^{i} P(j) \quad (1)$$

The new equalized pixel value I′ for a pixel with original intensity I is then computed as shown in Equation 2:

$$I' = \text{round}\left((CDF(I) - CDF_{min}) X \frac{L-1}{MXN - CDF_{min}}\right) \quad (2)$$

where 'L' denotes the total intensity levels, and M x N represents the overall pixel count

in the image. This transformation ensures that the intensities are spread out more evenly across the entire range, enhancing the visual quality and feature detectability.

3.2.4 Normalization

Normalization is applied to scale the pixel values of facial images to a consistent range, ensuring uniform input distribution for the model. This step stabilizes and accelerates the training process by preventing the network from becoming biased toward features with larger numerical values. In this work, the input images are normalized by centering pixel intensities around zero and scaling them to unit variance. Mathematically, each pixel intensity I is transformed to a normalized value I' as shown in the Equation (3).

$$I' = \frac{I-\mu}{\sigma} \quad (3)$$

where μ and σ denote the mean and standard deviation of the pixel intensities, respectively. This operation standardizes illumination differences across images and improves the model's ability to extract consistent spoof-related cues, ultimately contributing to more stable optimization and robust learning during training.

3.3 Feature Extraction

This research introduces a novel Adaptive Feature Extraction Framework (AFEF) to enhance Face Anti-Spoofing (FAS) through robust, context-aware feature extraction. The proposed framework consists of three modules: Adaptive Kernel Generator (AKG), Discrete Style Assembly (DSA), and Adaptive Style Transfer (AST). Figure 2 illustrates the proposed feature extraction framework and its constituent modules.

3.3.1 Adaptive Kernel Generator (AKG)

Static filters may not be sufficient to learn the diverse array of variations found in different domains. The Adaptive Kernel Generator (AKG) improves the model's generalization ability by creating instance-specific filters that adjust to the distinctive characteristics of each domain, ensuring more robust performance on unseen data. AKG dynamically adjusts kernel parameters according to the input data's characteristics, optimizing feature extraction to capture subtle domain-specific nuances. Additionally, by generating dynamic, instance-adaptive filters, AKG minimizes the risk of overfitting. The AKG framework comprises two primary components: (i) a static convolution branch that utilizes fixed parameters and processes half of the input channels through a static filter (ii) a dynamic kernel branch that generates instance-adaptive filters, adjusting parameters based on each input instance.

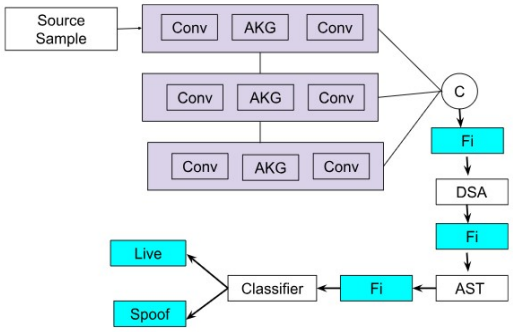

Figure 2: Proposed Adaptive Feature Extraction Framework (AFEF)

Let X_i be the feature map of the i^{th} sample, with D representing the number of channels, A denoting the height, and B indicating the width, resulting in dimensions D×A×B. The input feature map is segmented into two parts along the channel dimension as shown in Equation 4:

$$X_i = [\hat{X}_i, \tilde{X}_i] \quad (4)$$

where $\hat{X}_i \in R^{C/2XHXW}$ and $\tilde{X}_i \in R^{C/2XHXW}$

The static convolution branch processes $\tilde{X}_i$ using a static kernel $f\theta_s$ with fixed parameters is shown in Equation 5:

$$\tilde{Z}_i = f\theta_s(\tilde{X}_i) \quad (5)$$

The dynamic kernel branch generates instance-adaptive kernels W_i from $\hat{X}_i$. This is achieved through global average pooling followed by a learnable convolution block f_{θ_1}, as shown in Equation 6.

$$W_i = f\theta_1(avgpool(\hat{X}_i)) \quad (6)$$

In this equation, W_i represents the instance-specific convolutional kernel, f_{θ_1} denotes the parameterized convolution block, and $\hat{X}_i$ is the input feature map from the previous layer.

The instance-adaptive kernels Wi are then used in a dynamic convolution operation to extract features from $\hat{X}$I as shown in Equation 7.

$$\hat{Z}_i = \text{conv}(\hat{X}_i, W_i) \quad (7)$$

Here, $\hat{Z}_i$ denotes the dynamically extracted feature map, and $\text{conv}(\cdot)$ represents the convolution operation parameterized by the adaptive kernels.

The output feature maps from the static and dynamic branches, denoted as $\tilde{Z}_i$ and $\hat{Z}_i$, are concatenated along the channel dimension to form the combined representation Z_i, as defined in Equation (8):

$$Z_i = [\tilde{Z}_i, \hat{Z}_i] \quad (8)$$

Finally, the concatenated feature map Z_i is passed through an additional convolution block f_{θ_2} to produce the enhanced feature representation F_i, as given in Equation (9):

$$F_i = f\theta_2(Z_i) \quad (9)$$

In these formulations, f_{θ_2} denotes the subsequent learnable convolutional function, and F_i represents the enhanced feature map incorporating both static and dynamic information.

Equations (4)-(8) collectively describe the adaptive kernel generation process, which allows the model to adjust its feature extraction mechanism based on the input characteristics and improve robustness across domains.

3.3.2 Discrete Style Assembly (DSA)

DSA generates diverse feature representations to manage the variability introduced by different Presentation Attack Instruments (PAIs), helping the model to generalize effectively and reducing the risk of overfitting to specific attack types. Inspired by farthest point sampling (FPS) techniques used in point cloud down-sampling, DSA ensures that the selected basis styles from various source samples represent the broadest possible style diversity. During each epoch of model training, DSA dynamically updates the basis styles to reflect changes in the style space. Specifically, DSA employs FPS to iteratively select a set of L basis styles from all available samples for each category, such as real faces and spoof samples. These basis styles are chosen to maximize dissimilarity from the remaining styles, ensuring comprehensive coverage of the style space. The mean (μ) and variance (σ^2) of the basis styles for each category are then computed and stored in memory banks to capture their statistical characteristics.

To generate style-diversified samples, DSA combines basis styles using weights sampled from a Dirichlet distribution $B(\alpha_1, \ldots, \alpha_L)$. These weights control how the basis styles are combined, which helps in generating novel styles within each category. The newly generated styles are applied to augment the original features. For each content feature of an instance, the appropriate set of basis styles (either real or spoof) is selected based on the instance's category label. This ensures that the style reassembly process maintains the integrity of the instance's classification, thereby avoiding label changes that could compromise the FAS task. The style augmentation process is mathematically represented as in Equations 10 and 11:

$$\mu^c_{aug} = W^c \cdot \mu^c_{base}, \ \sigma^c_{aug} = W^c \cdot \sigma^c_{base} \quad (10)$$

$$F_{aug} = \sigma^c_{aug}\left(\frac{F_{org} - \mu(F_{org})}{\sigma(F_{org})}\right) + \mu^c_{aug} \quad (11)$$

Where $\mathbf{W_c}$ are the weights sampled from the Dirichlet distribution, μ_{basec} is the mean value and σ_{basec} is the standard deviation of the basis styles for the respective category c. Algorithm 1 illustrates the proposed Adaptive Feature Extraction Framework.

3.3.3 Adaptive Style Transfer (AST)

AST adapts feature extraction methods to varying image qualities and environmental conditions using adaptive style transfer techniques. The goal is to refine feature extraction to align with the distinctive characteristics of input images, improving the model's robustness against different spoofing tactics and reliability in diverse operational environments. The style transfer mechanism in AST is based on the Adaptive Instance Normalization (AdaIN) technique. The core idea is to adjust the mean and variance of feature maps to match the target style statistics, ensuring the robustness of the extracted features in image quality and conditions. The input feature map F of an image is first normalized using instance normalization (IN) to remove style-specific information. This is done as in the following Equation 12:

$$\hat{F} = \frac{F-\mu(F)}{\sigma(F)} \quad (12)$$

Where F be the feature map, μ(F) denotes the mean value and σ(F) the standard deviation of F. To adaptively transfer the style, AST computes the target mean μtgt and variance σ_{tgt} from a style image or a set of style images. The normalized feature map $\hat{F}$ is then adjusted to match these target statistics as shown in Equation 13.

$$F_{AST} - \sigma_{tgt}.\hat{F} + \mu_{tgt} \quad (13)$$

Here, σ_{tgt} and μ_{tgt} are dynamically computed according to the characteristics of the input data to ensure adaptive feature extraction.

The target mean μ_{tgt} and variance σ_{tgt} are not fixed but are dynamically adjusted based on the input features. This adaptation is crucial for handling varying image qualities and environmental conditions. First, the input feature map F is processed using global average pooling to get a concise representation of the style information as in the Equation 14.

$$f_{gap} = \frac{1}{HXW}\sum_{h=1}^{H}\sum_{w=1}^{W} F_{h.w} \quad (14)$$

where H – Height and W – Weight of the feature map F.

The pooled feature f_{gap} is subsequently processed through two dense layers, each using a Rectified Layer Unit activation function applied between them to calculate the target statistics. The weights in the final layer are normalized using a softmax function to ensure stability during adaptation, as defined in Equations 15 and 16.

$$\mu_{tgt} = FC_2\left(ReLU\left(FC_1(f_{gap})\right)\right) \quad (15)$$

$$\sigma_{tgt} = \sqrt{FC_2\left(ReLU\left(FC_1(f_{gap})\right)\right)} \quad (16)$$

In these equations, FC_1 and FC_2 denote fully connected layers, μ_{tgt} represents the target mean statistics, and σ_{tgt} denotes the target standard deviation derived from the pooled feature vector.

The loss function of the AST module is designed to ensure that the style-transferred features preserve discriminative information while adapting effectively to different domain styles. The content loss component maintains the semantic integrity of the original feature representation, whereas the style loss enforces alignment between the source and target style statistics. The total loss function combines both components with appropriate weighting, as expressed in Equations 17–19.

$$L_{content} = E[\|F - F_{AST}\|^2] \quad (17)$$

$$L_{style} = E\left[\left\|\mu(F_{AST}) - \sigma_{tgt}\right\|^2 + \left\|\sigma(F_{AST}) - \sigma_{tgt}\right\|^2\right] \quad (18)$$

$$L_{AST} = \lambda_c L_{content} + \lambda_s L_{style} \quad (19)$$

Here, $L_{content}$ and L_{style} represent the content and style losses, respectively, while λ_c and λ_s are the weighting factors that balance content preservation and style adaptation.

3.4 Classification

In this phase, the discriminative feature representations produced by the AFEF module are mapped to binary class scores for face anti-spoofing. The extracted features are first flattened and passed through a batch normalization layer to stabilize learning. A dropout layer is then applied to reduce overfitting by randomly deactivating a portion of neurons during training.

Next, the refined features are fed into a fully connected layer that generates class logits, which are converted into real and spoof probabilities using the sigmoid activation function. The model is trained using Binary Cross-Entropy Loss, and the Adam optimizer is employed to update network parameters efficiently. The training pipeline follows a standard forward and backward propagation procedure to minimize classification loss.

In addition to the primary classification objective, the AST module includes a style-guided learning objective combining content loss and style loss. The total AST loss is expressed as a weighted sum in Equation 20 as follows:

$$L_{AST} = \lambda_C L_{Content} + \lambda_C L_{style} \quad (20)$$

where λ_c and λ_s control the trade-off between preserving feature content and enforcing style consistency.

Model performance is finally evaluated on validation and test sets using accuracy and related metrics, ensuring reliable separation between real and spoof faces across varying conditions.

3.5 Performance Evaluation

In the evaluation, the key metrics such as Half Total Error Rate (HTER) and Accuracy (Acc) are used to assess the model's effectiveness in both intra-dataset and cross-dataset testing. HTER is calculated as the mean of the False Rejection Rate (FRR) and the False Acceptance Rate (FAR), gives a balanced estimation of the model's error rate. FRR represents the proportion of genuine faces incorrectly classified as spoofed, while FAR denotes the proportion of spoofed faces misclassified as genuine. Accuracy measures the proportion of correctly classified samples, providing a straightforward indicator of classification correctness.

4. EVALUATION AND RESULTS

This study evaluates the effectiveness of the proposed Adaptive Feature Extraction Framework (AFEF) for face anti-spoofing under four experimental settings: (i) Single Source Domain Experiment, which assesses the model's ability to detect face spoofing within a single dataset; (ii) Cross Domain Training and Testing, where the model is trained on one dataset and evaluated on another to assess generalization capability; (iii) Dual-Domain Training and Testing, which involves creating an intermediate dataset by combining two different datasets for training and testing the model; and (iv) Multi-Domain Training and Testing, which examines robustness by training on multiple datasets and evaluating performance on unseen domains. Model performance is measured using Accuracy and Half Total Error Rate (HTER), and results are compared against a baseline ResNet-18 model. The experimental outcomes are summarized in Tables 2–5.

4.1 Single-Source Domain Evaluation

This experiment evaluates model performance when both training and testing are conducted on the same dataset. Table 2 presents the results across the three benchmark datasets. The proposed AFEF improves performance across all single-source settings. The most notable gain occurs on the Replay dataset, where accuracy increases from 72.14% to 99.88% and HTER decreases significantly. Similar improvements are seen on SiW-Mv2 and OULU, confirming that AFEF enhances spoof-discriminative feature learning when training and testing occur within the same domain.

4.2 Cross Domain Evaluation

Table 3 shows a comparison of performance between a baseline ResNet-18 and the proposed feature extraction framework across various datasets, where the models are trained and tested on different datasets. The results show that the proposed AFEF consistently outperforms the baseline ResNet-18 across all single-source settings. A particularly large improvement is observed on the Replay dataset, where the framework significantly boosts detection accuracy and reduces error rate, indicating stronger discrimination of spoof patterns. The gains achieved on SiW-Mv2 and OULU further confirm that AFEF enhances feature quality and improves in-domain spoof detection performance across diverse acquisition conditions.

4.3 Dual Domain Evaluation

Table 4 presents the results where training and testing are performed on a combined dataset formed from two domains. The proposed AFEF consistently improves performance across all dataset combinations, achieving higher accuracy and lower HTER than the baseline ResNet-18. This shows that learning from blended domain distributions strengthens the model's ability to capture diverse spoof cues. Overall, dual-domain training enhances robustness and improves the model's ability to generalize across varying attack conditions and acquisition environments.

4.4 Multi-Domain Cross Evaluation

This experiment evaluates the model by training on multiple datasets and testing on a third, unseen dataset. This setup tests cross-domain generalization by exposing the model to varied spoofing conditions during training and evaluating its ability to adapt to a completely new environment. As shown in Table 5, the proposed AFEF consistently improves performance over the baseline ResNet-18, demonstrating stronger robustness and better transferability across diverse attack types and acquisition settings.

In this setting, the model is trained on combinations of two datasets and tested on a third,

unseen dataset. This setup simulates real-world deployment where training and deployment domains differ. As shown in Table 5, the proposed AFEF consistently outperforms the baseline ResNet-18 across all multi-domain configurations, demonstrating stronger generalization and lower error rates. These outcomes confirm the framework's ability to transfer spoof-discriminative features across diverse environmental and presentation attack variations.

Across all four evaluation settings, single-source, cross-domain, dual-domain, and multi-domain, the proposed AFEF consistently improves performance over the baseline ResNet-18. The framework achieves higher accuracy and lower HTER in most cases, particularly in challenging cross-dataset and multi-domain scenarios. These results demonstrate that AFEF effectively enhances feature robustness and generalization, making it well-suited for real-world face anti-spoofing systems where attack conditions and acquisition environments vary significantly.

4.5 Comparative Evaluation

Most research on face spoofing attack detection focuses on a few common datasets like CASIA-MFSD (C), Idiap Replay-Attack (I), MSU-MFSD (M), and OULU-NPU (O). However, there is little exploration of how face spoofing detection systems generalize across different dataset combinations.

Most existing face anti-spoofing (FAS) studies evaluate their models on a limited set of benchmark datasets such as CASIA-MFSD, Idiap Replay-Attack, MSU-MFSD, and OULU-NPU. However, there is limited investigation into how Face Anti-spoofing systems generalize across combinations of diverse modern datasets. In particular, no prior work has jointly evaluated Replay-Attack, SiW-Mv2, and OULU-NPU, three datasets that collectively cover a wide range of presentation attack instruments, capture devices, and environmental conditions.

Replay-Attack includes varied lighting and attack mediums, SiW-Mv2 features high-resolution and diverse pose variations, and OULU-NPU introduces multiple devices and illumination changes. Together, they provide a comprehensive and challenging test bed for cross-domain robustness in real-world scenarios. The inclusion of all three datasets makes this evaluation setup distinct from existing studies and enables a deeper understanding of generalization performance.

Due to this unique dataset configuration, a direct comparison with prior state-of-the-art works is not possible. Therefore, a relative performance comparison is conducted using limited-source domain settings, following protocols commonly used in domain generalization and leave-one-out (LOO) validation studies. Accuracy and Half Total Error Rate (HTER) metrics are used for consistency with prior literature. The comparative results against representative state-of-the-art methods under limited-source evaluation settings are summarized in Table 6.

4.6 Discussion

The results demonstrate that the proposed AFEF enhances spoof detection performance across all evaluation settings by producing more discriminative and domain-robust features than a baseline ResNet-18 model. This improvement is consistent with recent advances in domain generalization for face anti-spoofing and extends prior work by evaluating performance on a broader and more challenging combination of datasets, including Replay-Attack, SiW-Mv2, and OULU-NPU. The observed gains in both cross-domain and multi-domain settings highlight the framework's suitability for real-world biometric systems operating under diverse and unseen attack conditions. Nevertheless, performance fluctuations under extreme domain shifts indicate sensitivity to highly heterogeneous acquisition environments, suggesting scope for further improvement.

4.6.1 Limitations of the proposed work

Although the proposed Adaptive Feature Extraction Framework improves robustness under cross-domain settings, it has certain limitations. The experimental results show that performance can fluctuate when the model is exposed to extreme domain shifts, suggesting that highly diverse acquisition conditions are still difficult to handle. In addition, the evaluation was conducted on three public benchmark datasets, which may not adequately capture emerging spoofing techniques or low-quality attack samples commonly encountered in unconstrained scenarios. The use of multiple adaptive modules also increases training complexity, which could limit scalability in practical deployments. Addressing these issues will require further analysis and optimization to improve both robustness and computational efficiency.

4.7 Difference From Prior Work

Most existing face anti-spoofing methods rely on fixed feature extractors, classifier-level optimization, or data augmentation to improve generalization. In contrast, the proposed Adaptive Feature Extraction Framework (AFEF) focuses on learning adaptive and style-aware features to better handle domain variability. Unlike prior works that address domain shift mainly through adversarial learning or sample reweighting, this approach explicitly models domain and style variations during feature extraction, leading to improved robustness in cross-domain and multi-domain evaluations.

The framework demonstrates improved cross-dataset generalization and stable performance across multiple benchmark datasets. However, the use of multiple adaptive modules increases training complexity, and performance degradation may still occur under extreme domain shifts, suggesting directions for future refinement. These results confirm that the primary objective of improving cross-domain generalization in face anti-spoofing has been successfully achieved in comparison to existing approaches reported in the literature.

5. CONCLUSION

The study shows that adaptive feature learning is an effective strategy in improving the robustness of face anti-spoofing systems when tested under heterogeneous domains and unseen attack conditions. The experimental results confirm that cross-domain generalization continues to be a challenging issue, particularly in the presence of varying acquisition environments, and show that the proposed Adaptive Feature Extraction Framework (AFEF) is capable of alleviating this problem. By learning more stable and discriminative representations than conventional fixed feature extractors, the framework consistently performs better in cross-dataset and multi-domain evaluations. The experimental results support the stated hypothesis by demonstrating that adaptive, feature-level modeling leads to more robust performance across cross-dataset and multi-domain evaluation settings.

At the same time, the analysis reveals certain limitations. The experiments show that performance can still vary under extreme domain shifts, suggesting that highly diverse acquisition conditions remain difficult to handle. Moreover, the evaluation was carried out on only three benchmark datasets, which may not adequately capture newer spoofing techniques or poor-quality attack samples often encountered outside controlled settings. A more detailed investigation is therefore needed to better understand these cases. Future extensions may consider alternative learning strategies, including adversarial objectives or different network designs, to improve robustness. Evaluating the framework on deepfake attacks, 3D mask spoofing, and data collected from real deployment environments would further clarify its practical reliability in security-critical applications.

REFERENCES:

[1] X. Yang, J. Liu, Z. Li, and X. Liu, “DADM: Dual Alignment of Domain and Modality for Face Anti-Spoofing,” *arXiv Preprint*, arXiv:2503.00429, 2025.

[2] S. Zhou, C. Xie, Z. Li, Y. Zhang, and K.-Y. Zhang, “Test-Time Domain Generalization for Face Anti-Spoofing,” *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Seattle, USA, June 2024.

[3] I. Chingovska, A. Anjos, and S. Marcel, “On the Effectiveness of Local Binary Patterns in Face Anti-Spoofing,” *Proceedings of the BIOSIG 2012 – International Conference of the Biometrics Special Interest Group*, Darmstadt, Germany, September 2012, pp. 1–7.

[4] Y. Moon, I. Ryoo, and S. Kim, “Face Anti-Spoofing Method Using Color Texture Segmentation on FPGA,” *Hindawi Journal*, 2021.

[5] F. Sthevanie and K. N. Ramadhani, “Spoofing Detection on Facial Images Recognition Using LBP and GLCM Combination,” *IOP Conference Series: Journal of Physics: Conference Series*, Vol. 971, 2018, Article No. 012014.

[6] D. Karmakar, P. Mukherjee, and M. Datta, “Spoofed Facial Presentation Attack Detection by Multivariate Gradient Descriptor in Micro-Expression Region,” *Pattern Recognition and Image Analysis*, Vol. 31, 2021, pp. 285–294.

[7] A. Maurya and S. Tarar, “Spoofed Video Detection Using Histogram of Oriented Gradients,” *Proceedings of the 3rd International Symposium on Computer Vision and the Internet (VisionNet ’16)*, ACM, India, 2016, pp. 1–7.

[8] Y. A. Rehman, L. M. Po, Z. Liu, W. Zou, and W. Ou, "Perturbing Convolutional Feature Maps with Histogram of Oriented Gradients for Face Liveness Detection," *Proceedings of the CISIS–ICEUTE 2019 Conference*, Springer, Italy, June 2019.

[9] Z. Boulkenafet, J. Komulainen, and A. Hadid, "Face Anti-Spoofing Using Speeded-Up Robust Features and Fisher Vector Encoding," *IEEE Signal Processing Letters*, Vol. 24, No. 2, 2017, pp. 141–145.

[10] L. Lei, Z. Xia, J. Wu, L. Yang, and H. Han, "Face Presentation Attack Detection Based on Optical Flow and Texture Analysis," *Journal of King Saud University – Computer and Information Sciences*, Vol. 34, No. 4, 2022, pp. 1455–1467.

[11] W. Liu, "Face Liveness Detection Using Analysis of Fourier Spectra Based on Hair," *Proceedings of the International Conference on Wavelet Analysis and Pattern Recognition*, IEEE, Chengdu, China, July 2014, pp. 75–80.

[12] S. Maphisa and D. Coulter, "Face Anti-Spoofing Based on Convolutional Neural Networks," *Proceedings of the International Conference on Intelligent Data Science Technologies and Applications (IDSTA)*, IEEE, South Africa, 2022, pp. 68–74.

[13] J. Gan, S. Li, Y. Zhai, and C. Liu, "3D Convolutional Neural Network Based on Face Anti-Spoofing," *Proceedings of the 2nd International Conference on Multimedia and Image Processing (ICMIP)*, IEEE, Hong Kong, China, March 2017.

[14] U. Muhammad, T. Holmberg, W. C. Melo, and A. Hadid, "Face Anti-Spoofing via Sample Learning Based Recurrent Neural Network," *Proceedings of the British Machine Vision Conference (BMVC)*, BMVA Press, Cardiff, UK, September 2019.

[15] C.-S. Fahn, C.-P. Lee, and M.-L. Wu, "A Cross-Dataset Evaluation of Anti-Face-Spoofing Methods Using Random Forests and Convolutional Neural Networks," *Proceedings of the Artificial Intelligence and Cloud Computing Conference (AICCC '19)*, ACM, New York, USA, 2019, pp. 89–96.

[16] Z. Yu, Y. Qin, X. Li, C. Zhao, Z. Lei, and G. Zhao, "Deep Learning for Face Anti-Spoofing: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 45, No. 5, 2023, pp. 5609–5631.

[17] B. Zhang, B. Tondi, and M. Barni, "Adversarial Examples for Replay Attacks Against CNN-Based Face Recognition with Anti-Spoofing Capability," *Computer Vision and Image Understanding*, Vols. 197–198, 2020, Article No. 102988.

[18] M. Smiatacz, "Liveness Measurements Using Optical Flow for Biometric Person Authentication," *Metrology and Measurement Systems*, Vol. 19, No. 2, May 2012, pp. 257–268.

[19] N. A.-H. Taha, T. Hasan, and M. A. Younus, "Face Spoofing Detection Using Deep CNN," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, Vol. 12, No. 13, 2021, pp. 4363–4373.

[20] Z. Xu, S. Li, and W. Deng, "Learning Temporal Features Using LSTM-CNN Architecture for Face Anti-Spoofing," *Proceedings of the 3rd IAPR Asian Conference on Pattern Recognition (ACPR)*, IEEE, Kuala Lumpur, Malaysia, November 2015.

[21] L. Li, X. Feng, Z. Boulkenafet, Z. Xia, M. Li, and A. Hadid, "An Original Face Anti-Spoofing Approach Using Partial Convolutional Neural Network," *Proceedings of the IEEE 6th International Conference on Image Processing Theory, Tools, and Applications (IPTA)*, IEEE, Oulu, Finland, December 2016, pp. 1–6.

[22] X. Yang, W. Luo, L. Bao, Y. Gao, D. Gong, S. Zheng, Z. Li, and W. Liu, "Face Anti-Spoofing: Model Matters, So Does Data," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Long Beach, USA, June 2019.

[23] L. Feng, L.-M. Po, Y. Li, X. Xu, F. Yuan, T. C.-H. Cheung, and K.-W. Cheung, "Integration of Image Quality and Motion Cues for Face Anti-Spoofing: A Neural Network Approach," *Journal of Visual Communication and Image Representation*, Vol. 38, 2016, pp. 451–460.

[24] Z. Wang, Z. Yu, C. Zhao, X. Zhu, Y. Qin, Q. Zhou, F. Zhou, and Z. Lei, "Deep Spatial Gradient and Temporal Depth Learning for Face Anti-Spoofing," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Seattle, USA, June 2020.

[25] R. Huang and X. Wang, "Face Anti-Spoofing Using Feature Distilling and Global Attention Learning," *Pattern Recognition*, Vol. 135, 2023, Article No. 109147.

[26] Z. Yu, J. Wan, Y. Qin, X. Li, S. Z. Li, and G. Zhao, "NAS-FAS: Static-Dynamic Central Difference Network Search for Face Anti-Spoofing," *IEEE Transactions on Pattern*

Analysis and Machine Intelligence, Vol. 43, No. 9, September 2021, pp. 3005–3023.

[27] R. Cai, H. Li, S. Wang, C. Chen, and A. C. Kot, "DRL-FAS: A Novel Framework Based on Deep Reinforcement Learning for Face Anti-Spoofing," *IEEE Transactions on Information Forensics and Security*, Vol. 16, 2020, pp. 937–951.

[28] H.-H. Chang and C.-H. Yeh, "Face Anti-Spoofing Detection Based on Multi-Scale Image Quality Assessment," *Image and Vision Computing*, Vol. 121, 2022, Article No. 104428.

[29] Y. Jia, J. Zhang, S. Shan, and X. Chen, "Single-Side Domain Generalization for Face Anti-Spoofing," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Seattle, USA, June 2020.

[30] Z. Wang, Z. Wang, Z. Yu, W. Deng, J. Li, T. Gao, and Z. Wang, "Domain Generalization via Shuffled Style Assembly for Face Anti-Spoofing," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, New Orleans, USA, June 2022.

[31] S. Liu, K.-Y. Zhang, T. Yao, K. Sheng, S. Ding, Y. Tai, J. Li, Y. Xie, and L. Ma, "Dual Reweighting Domain Generalization for Face Presentation Attack Detection," *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, AAAI Press, Virtual Conference, February 2021.

[32] W. Wang, P. Liu, H. Zheng, and R. Ying, "Domain Generalization for Face Anti-Spoofing via Negative Data Augmentation," *IEEE Transactions on Information Forensics and Security*, 2023.

[33] Y. Liu, Z. Li, Y. Xu, Z. Guo, Z. Zou, and L. Wu, "Quality-Invariant Domain Generalization for Face Anti-Spoofing," *International Journal of Computer Vision*, Vol. 132, No. 11, 2024, pp. 5239–5254.

[34] B. Sun and K. Saenko, "Deep CORAL: Correlation Alignment for Deep Domain Adaptation," *Proceedings of the European Conference on Computer Vision (ECCV)*, Springer, Amsterdam, The Netherlands, October 2016, pp. 443–450.

[35] M. Liu, J. Mu, Z. Yu, K. Ruan, B. Shu, and J. Yang, "Adversarial Learning and Decomposition-Based Domain Generalization for Face Anti-Spoofing," *Pattern Recognition Letters*, Vol. 155, 2022, pp. 171–177.

[36] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image Style Transfer Using Convolutional Neural Networks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Las Vegas, USA, June 2016, pp. 2414–2423.

[37] X. Pan, X. Zhan, J. Shi, X. Tang, and P. Luo, "Switchable Whitening for Deep Representation Learning," *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, IEEE, Seoul, South Korea, October 2019, pp. 1863–1871.

[38] W. Cho, S. Choi, D. K. Park, I. Shin, and J. Choo, "Image-to-Image Translation via Group-Wise Deep Whitening-and-Coloring Transformation," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Long Beach, USA, June 2019, pp. 10639–10647.

[39] Y. Li, C. Fang, J. Yang, Z. Wang, X. Lu, and M.-H. Yang, "Universal Style Transfer via Feature Transforms," *Advances in Neural Information Processing Systems (NeurIPS)*, MIT Press, Long Beach, USA, December 2017, pp. 385–395.

[40] S. Roy, A. Siarohin, E. Sangineto, S. R. Bulo, N. Sebe, and E. Ricci, "Unsupervised Domain Adaptation Using Feature-Whitening and Consensus Loss," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Long Beach, USA, June 2019, pp. 9471–9480.

[41] S. Choi, S. Jung, H. Yun, J. T. Kim, S. Kim, and J. Choo, "RobustNet: Improving Domain Generalization in Urban-Scene Segmentation via Instance Selective Whitening," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Nashville, USA, June 2021, pp. 11580–11590.

[42] M. P. Swapna, D. Rajeev, J. Ramkumar, and S. Chandran, "Unveiling Cybercrime Patterns in Kerala: A Machine Learning Approach," in Lecture Notes in Networks and Systems, 2026, pp. 111–122.

[43] R. Jaganathan, S. Mehta, and R. Krishan, "Preface," Bio-Inspired Intell. Smart Decis., pp. xix–xx, 2024.

[44] R. Jaganathan, S. Mehta, and R. Krishan, "Preface," Intell. Decis. Mak. Through Bio-Inspired Optim., pp. xiii–xvi, 2024.

[45] J. Joy, S. Devaraju, and J. Ramkumar, "Utilizing Fuzzy-Identity-Based Encryption With Proxy-Re-Encryption For Data Sharing,"

J. Theor. Appl. Inf. Technol., vol. 103, no. 18, pp. 7469–7479, 2025.

[46] R. Jaganathan, S. Rajagopal, and K. Rajendran, "Cultural Intelligence in the AI Era-Enhancing Transitional Higher Education," in Bridging Global Divides for Transnational Higher Education in the AI Era, 2024, pp. 273–292.

[47] J. Ramkumar and V. Valarmathi, "Harnessing AI-Driven Models for Sustainable Development in Business Management," in World Sustainability Series, 2025, pp. 217–238.

[48] X. Guo, Y. Liu, A. Jain, and X. Liu, "Multi-Domain Learning for Updating Face Anti-Spoofing Models," *Proceedings of the European Conference on Computer Vision (ECCV)*, Springer, Tel Aviv, Israel, October 2022.

[49] Z. Boulkenafet, J. Komulainen, L. Li, X. Feng, and A. Hadid, "OULU-NPU: A Mobile Face Presentation Attack Database with Real-World Variations," *Proceedings of the 12th IEEE International Conference on Automatic Face and Gesture Recognition (FG)*, IEEE, Washington, USA, May 2017, pp. 612–618.

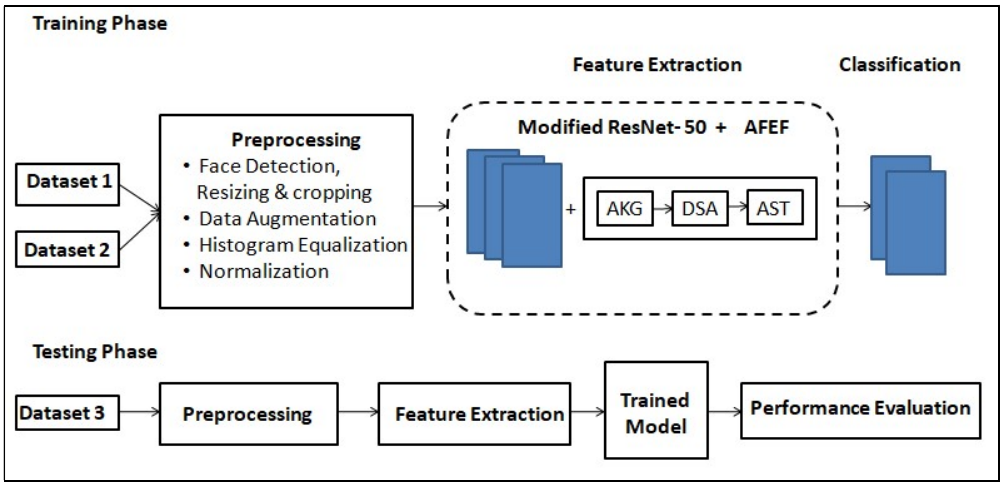

Figure 1: Proposed Methodology

Table 1: Details Of Datasets Used

Dataset	Capturing devices	Attack Types	Background scenes	Training videos	Testing videos	Development videos
Replay-Attack	Webcams, Mobile phones	Print & Video attacks	Controlled environment	360	480	360
SiW-Mv2	Mobile phones	Print, Replay & 3D mask attacks	Varied	1650	700	350
OULU-NPU	Mobile phones	Print, Replay & Video attacks	Varied	4944	1614	NA

Algorithm 1: The Proposed Adaptive Feature Extraction Framework (AFEF)

Input: Preprocessed face image I

Initialization:

Initialize static and dynamic convolution branches in AKG
Initialize fully-connected style mapping layer in DSA
Initialize instance-normalization and style-transfer parameters in AST

Perform:

Extract features using static convolution path
Extract adaptive features using dynamic convolution path
Fuse static and dynamic features through interpolation and concatenation
Generate style vector from fused feature maps
Apply softmax to obtain style attention weights
Modulate feature maps using style attention
Normalize features and apply learned style-transfer parameters

At the j-th iteration:

Update module parameters using back-propagation to refine AKG, DSA, and AST outputs

Stopping criteria:

If convergence achieved or maximum iterations reached → Stop
Otherwise, $j = j + 1$ and continue

Output: Enhanced feature representation $F_{enhanced}$ for robust face anti-spoofing classification

Table 2: Single-Source Domain Performance Comparison

Train → Test	Model	Accuracy (%)	HTER (%)
SiW-Mv2 → SiW-Mv2	ResNet-18	96.56	0.05
	ResNet-18 + AFEF	97.29	0.40
Replay → Replay	ResNet-18	72.14	0.42
	ResNet-18 + AFEF	99.88	0.10
OULU → OULU	ResNet-18	97.76	0.578
	ResNet-18 + AFEF	98.77	0.21

Table 3: Cross-Domain Training And Testing Results

Train → Test	Model	Accuracy (%)	HTER (%)
Replay → SiW-Mv2	ResNet-18	60.06	2.445
	ResNet-18 + AFEF	70.68	1.908
OULU → Replay	ResNet-18	79.09	13.45
	ResNet-18 + AFEF	81.8	20.98
SiW-Mv2→ OULU	ResNet-18	76.02	9.80
	ResNet-18 + AFEF	72.77	11.08

Table 4: Cross-Domain Training And Testing Results

Train → Test	Model	Accuracy (%)	HTER (%)
Replay + OULU → Replay + OULU	ResNet-18	96	0.45
	ResNet-18 + AFEF	**98.89**	**0.32**
SiwMv2 + Replay → SiwMv2 + Replay	ResNet-18	91.93	5.52
	ResNet-18 + AFEF	**98.35**	**0.72**
OULU + SiwMv2 → OULU + SiwMv2	ResNet-18	94	0.90
	ResNet-18 + AFEF	**97.22**	**0.71**

Table 5: Multi-Domain Cross-Evaluation Results

Train → Test	Model	Accuracy (%)	HTER (%)
Replay + OULU → SiwMv2	ResNet-18	72.3	90.7
	ResNet-18 + AFEF	**86.90**	**45.7**
SiwMv2 + OULU → Replay	ResNet-18	68.2	79.8
	ResNet-18 + AFEF	**91.30**	**58.7**
Replay + SiwMv2 → OULU	ResNet-18	76.8	69.7
	ResNet-18 + AFEF	**94**	**9.2**

Table 6: Comparison To The-State-Of-Art FAS Methods On Limited Source Domains

Method	Training Datasets	Testing Dataset	Accuracy (%)	HTER (%)
AMEL (Zhou, 2022)	M & I	C	85.17	23.33
	M & I	O	87.01	19.68
ANRL (Liu, 2021)	M & I	C	72.12	31.06
	M & I	O	74.10	29.88
SSDG-M (Jia, 2020)	M & I	C	71.29	31.89
	M & I	O	36.01	66.88
MADGG (Shao, 2019)	M & I	C	64.33	41.02
	M & I	O	65.10	39.35
Ours (AFEF)	Replay-Attack & OULU-NPU	SiW-Mv2	86.90	45.70
	SiW-Mv2 & OULU-NPU	Replay-Attack	91.30	58.70
	Replay-Attack & SiW-Mv2	OULU-NPU	94.00	9.20