iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** II **Month of publication:** February 2024

DOI: https://doi.org/10.22214/ijraset.2024.58383

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue II Feb 2024- Available at www.ijraset.com

Face Anti-Spoofing Methods: A Comparative Analysis through the Lens of a Comprehensive Review

Chanchala Kamat[1], Prof. Navita Shrivastava[2]
[1]*Research Scholar,* [2]*Professor, Department of computer science, APS university, Rewa (M.P.)*

Abstract: ***In computer vision, face anti-spoofing is an important aspect that aims to differentiate genuine facial features from spoofing attempts. This review paper comprehensively explores existing methodologies, emphasising advancements in computer vision and deep learning. Diverse techniques, ranging from traditional methods like multi-scale LBPs and CNNs to recent innovations such as FeatherNet and ViT-S-Adapter-TSR, are meticulously analysed. A comparative table provides insights into different methods, highlighting their performance on various datasets like MSU-MFD, CASIA-FASD, and OULU-NPU. However, challenges like diverse datasets, varying evaluation metrics, and real-world applicability are acknowledged. The paper discusses limitations related to real-world conditions, computational efficiency, and the ever-evolving nature of spoofing techniques. It emphasises the need for ongoing collaboration and innovation in research to address challenges like dataset consistency and adaptability to emerging threats. In conclusion, while progress has been made, the paper emphasises the dynamic nature of face anti-spoofing research. The pursuit of more effective, adaptable, and computationally efficient methods continues, promising real-world impact against evolving threats.***

Keywords: ***Face anti-spoofing, Spoof detection, Deep learning, Computer vision, Dataset compilation, Evaluation metrics, Real-world scenarios, Computational efficiency, Methodological advancements, Cross-dataset evaluation.***

I. INTRODUCTION

In today's world, face recognition systems are being widely used in various areas, but there's growing concern about their vulnerability to face spoofing attacks.

Face spoofing creates challenges for face recognition systems as attackers use different methods like printing realistic images, replaying videos, and introducing three-dimensional facial mask. The emergence of deepfake technology adds another layer of complexity, introducing AIgenerated realistic facial changes in videos [1].

There are various attack methods that have been successful against different commercial face recognition systems, which highlights the need for stronger security measures. Face Anti-Spoofing (FAS) techniques are being developed to address this threat by going beyond accurate facial recognition and making systems more resistant to spoofing attempts. These techniques use smart algorithms to differentiate between real facial movements and fake ones, ultimately enhancing the overall security of these systems.

As we increasingly rely on face recognition in various aspects of our lives, it's crucial to explore and implement reliable FAS techniques [2]. This not only improves protection against deceptive attempts but also ensures the trustworthiness of these systems, especially in critical areas like finance, security, and personal device authentication. The evolving security challenges require continuous efforts to strengthen face recognition systems against a diverse range of threats.

Over the years, researchers have explored diverse methodologies to detect face-spoofing, ranging from traditional handcrafted features to advanced deep learning approaches and generative models. The initial forays into face anti-spoofing primarily concentrated on handcrafted features [3-5]. Commonly, using photos to bypass 2-D face recognition is a known threat, but counter-measures have lagged due to a lack of consensus and public databases.

The print-attack database is introduced [3], featuring a motion-based algorithm to detect correlations between head movements and scenes, aiming to address this gap and enhance face recognition security. while the problem of spoofing detection comes from a texture analysis approach based on multi-scale local binary patterns that identifies live people versus face prints but also provides features usable for face recognition without user cooperation [4]. further extended the exploration to dynamic texture for detecting face spoofing and to learn the structure and dynamics of the facial micro-textures that characterise only real faces but not fake ones. [5].

Recognising the necessity for diverse datasets, the paper addresses limitations by introducing a comprehensive database covering various attack variations and imaging qualities [6]. Investigating unconventional threats and analysing the potential of using makeup for spoofing an identity, where an individual attempts to impersonate another person's facial appearance, systematically explored the impact of makeup on automated face matchers [7]. To mitigate the spoofing risk, several video-based methods have been presented in the literature that analyse facial motion in successive video frames. However, estimating the motion between adjacent frames is a challenging task and requires high computational costs. This work rephrases the face anti-spoofing task as a motion prediction problem and introduces a deep ensemble learning model with a frame skipping mechanism [8]. Further, a novel recurrent convolutional neural network (ReCNN) architecture has been introduced, which is trained to learn a joint spectral-spatial-temporal feature representation in a unified framework for change detection in multispectral images [9].

The review paper aims to provide:

1) A comprehensive study of various face spoofing methods and explore diverse face anti-spoofing techniques;
2) A comprehensive analysis of various work done in the field of face anti spoofing;
3) A comparative study of various anti-spoofing techniques;
4) Presents a compilation of face anti-spoofing datasets captured using commercial RGB cameras;

II. FACE SPOOFING METHODS

The term "face spoofing" describes the employment of deceitful techniques to mislead facial recognition software into granting unapproved or fraudulent access. The potential of face spoofing becomes a major problem as facial recognition technology becomes more widely used in many applications, such as mobile devices, security systems, and authentication procedures. Face spoofing is achieved using a variety of techniques, which can be divided into four categories: physical attacks, morphing attacks, multimodal attacks, and deepfake generation.

A. Physical Attacks

1) *Print Attack:* In a print attack, an attacker uses a high-resolution photograph of the legitimate user's face to impersonate and gain unauthorized access. This method is simple and often effective, especially if the facial recognition system lacks antispoofing measures.
2) *Replay Attack:* In a replay attack, the attacker records a video or captures facial data during a legitimate authentication session. They then replay this pre-recorded data to the facial recognition system during subsequent login attempts, tricking the system into granting access.
3) *3D Mask Attack:* This involves the creation of a physical three-dimensional mask of the legitimate user's face. Advanced 3D printing technology allows attackers to produce realistic masks that mimic the contours and features of a real face, enabling successful spoofing.

B. Deepfake Generation (GAN-based Spoofing)

In order to produce realistic synthetic content, such as deepfake movies for face spoofing, Generative Adversarial Networks, or GANs, are extremely effective techniques. Generating artificial data and discriminating between actual and fake data are the two main components of GANs. Facial recognition algorithms find it difficult to distinguish between phoney and authentic identities when faced with face spoofing since GANs can be trained to produce realistic facial pictures or movies.

C. Morphing Attack (Face Morphing Attack)

Attacks known as morphing includes merging features from several faces to generate a composite image. It is challenging for face recognition algorithms to correctly identify the person because the resulting morphing image combines features from both individuals. Using this technique, attackers might be able generate a false identity that is able to get by authentication checks.

D. Multimodal Attack

1) Mask Attack: In a mask attack, an attacker wears a physical mask resembling the face of an authorized user. This can be combined with other methods, such as voice impersonation, to enhance the effectiveness of the attack.
2) Texture and Pattern Attack: Attackers may use textured materials or patterns on a mask or other spoofing medium to deceive facial recognition systems. These textures can confuse algorithms by introducing elements that resemble genuine facial features.

III. FACE ANTI SPOOFING TECHNIQUES:

Face anti-spoofing techniques use several types of approaches to identify and prevent face spoofing attacks, in which attackers trick facial recognition systems with images, videos, or other media. Based on the features they use, their type of analysis (dynamic or static), and the underlying technologies used, these approaches can be broadly defined.

1) *Feature-based Techniques:* In order to differentiate real faces from fakes, feature-based face anti-spoofing systems emphasis on identifying specific features from facial photos. Texture analysis is a popular method that looks for consistency in the patterns and imperfections in the skin's texture. Additionally, colour information is employed, examining differences in colour contrasts and patterns to find irregularities suggestive of spoofing attempts. Further, the term "liveness detection" refers to techniques used to identify indicators of life, such as head movement or eye blinking, frequently using additional sensors, such as depth or infrared ones, to improve accuracy. A key component of feature-based approaches is spectral analysis, which looks at face features in several frequency bands to spot unusual events.
2) *Motion-based Techniques:* In order to detect irregularities motion-based face antispoofing systems investigate temporal features of facial dynamics. By focusing on the temporal variations in facial characteristics, dynamic texture analysis can identify deviations in motion patterns that might point to a spoofing effort. By utilising the differences between real and artificial facial dynamics, facial dynamics analysis analyses how people's faces move and express themselves across time. Another method uses optical flow analysis to look for abnormal patterns produced on by face spoofing attacks by observing the flow of pixels between successive frames.
3) *3D-based Techniques:* These techniques leverage three-dimensional information to enhance face anti-spoofing capabilities. Depth information from 3D sensors is utilized to distinguish between real and fake faces by exploiting differences in facial feature depth. 3D face reconstruction techniques go further by creating a three-dimensional model of the face, enabling the analysis of spatial structure and identification of anomalies in facial geometry. These methods contribute an additional layer of sophistication to anti-spoofing systems.
4) *Machine Learning-based techniques:* Machine learning plays a pivotal role in face antispoofing, with various supervised and deep learning methods being employed. Supervised learning involves training models on labelled datasets containing both genuine and fake face samples. Popular algorithms include support vector machines (SVM), random forests, and neural networks. Deep learning, utilizing architectures such as convolutional neural networks and recurrent neural networks, enables automated feature learning and extraction. Transfer learning adapts pre-trained models from general face recognition tasks for face anti-spoofing, enhancing performance with fine-tuning on specific datasets.
5) *Multimodal Fusion Techniques:* Multimodal fusion techniques integrate information from multiple sources to improve overall anti-spoofing accuracy. This may involve combining RGB images with depth maps, infrared data, or other modalities. By leveraging complementary information, these techniques enhance the robustness of antispoofing systems, providing a more comprehensive understanding of the presented facial data.
6) *Biometric Fusion Techniques:* Biometric fusion techniques combine face anti-spoofing with traditional face recognition systems or other biometric modalities. Fusion with face recognition enhances security by combining liveness detection with identity verification. Voice and speech analysis may also be incorporated to create multimodal biometric systems, adding additional layers of security and making the overall system more resilient to spoofing attempts.
7) *Traditional Computer Vision Techniques:* Traditional computer vision techniques are employed in face anti-spoofing to address specific aspects of image analysis. Edge detection, for instance, focuses on identifying boundaries between facial features by examining edge patterns, while histogram analysis evaluates pixel intensity distributions for irregularities. These techniques provide classical yet effective approaches to antispoofing.
8) *Hybrid Techniques:* Hybrid techniques involve combining methodologies from different categories to create robust anti-spoofing systems. By integrating the strengths of multiple approaches, these hybrid techniques aim to improve overall performance and detection capabilities. The combination of feature extraction, motion analysis, and machine learning, for example, creates a comprehensive approach capable of detecting a wide range of face spoofing attacks. Researchers often experiment with hybrid models to achieve a balance between accuracy and computational efficiency in real-world scenarios.

IV. LITERATURE REVIEW

This comprehensive study aims to provide an overview of the existing works and approaches in the field of face spoofing detection. A lot of work has been done in this field, including:

Enoch Solomon and Krzysztof J. Cios 2023, [10] propose a face anti-spoofing system that integrates image quality features and deep learning methods. The system aims to distinguish between genuine and spoofed faces by analyzing both the inherent quality of the image and using deep learning models to learn discriminative representations. The combination of image quality features and deep learning enhances the system's ability to detect spoofing attacks, making it more robust and effective in real-world scenarios.

Deepika Sharma and Arvind selwal 2023, [11] provide a survey of the different techniques that have been proposed for face presentation attack detection (PAD). Pad is the task of detecting whether a face image or video is a genuine presentation of a person's face or a presentation attack. Presentation attacks can be carried out using a variety of methods, such as using printed images, masks, or even mobile phones. The paper concludes by discussing the future of pad. The authors believe that pad is an important research area, and that there is still much work to be done in developing effective pad techniques. They also believe that pad will become increasingly important as face recognition systems become more widely deployed.

L. Birla, P. Gupta, and S. Kumar 2023, [12] address the vulnerability of face analytic systems to face spoofing attacks in the digital world. The authors propose a face anti-spoofing method called SUNRISE (Short videos UsiNg pRe-emptIve Split and mErge) that utilizes remote Photoplethysmography (rPPG) to detect cardiovascular signals and mitigate face spoofing attacks. However, they acknowledge that the performance of rPPG-based anti-spoofing methods can be degraded due to illumination variation and face deformations, even with longer-duration face videos. To overcome this limitation, SUNRISE introduces a split and merge mechanism. It splits the video into clips, assigns low importance to clips with facial deformations, and merges the results using quality-based fusion. The proposed method utilizes statistical features of clips instead of high-dimensional features, mitigating the limitation of limited training data in existing rPPG-based methods. Experimental results on publicly available datasets demonstrate that SUNRISE outperforms well-known existing methods for short-duration videos, showcasing its effectiveness in face anti-spoofing.

Xin Cheng et al., [13] propose an anti-spoofing method for facial recognition systems in the Internet of Things (IoT) environment. It introduces a spoofing-detection algorithm based on optical flow and texture features, utilizing a two-channel convolutional neural network (CNN) to extract and fuse facial characteristics. To enhance the optical flow field map with liveness information, a motion amplification algorithm is applied. The method was evaluated on the Replay Attack dataset, achieving a half total error rate of 0.66%. By effectively discerning real and fake faces, this approach enhances the security and reliability of facial recognition systems, making it a promising solution for combating face spoofing attacks in IoT applications.

Shizhe Zhang and Wenhui Nie 2023 [14] introduce a novel approach called Multi-Domain Feature Alignment Framework (MADG) for face anti-spoofing, which enhances the robustness of face recognition systems against presentation attacks. Existing methods face challenges in generalizing features across different domains due to distribution discrepancies. MADG addresses this issue by employing an adversarial learning process to align features from multiple source domains, narrowing the domain differences. Additionally, a multi-directional triplet loss is incorporated to improve the separation between real and fake faces in the feature space. Extensive experiments on public datasets demonstrate that MADG outperforms current state-of-the-art methods, proving its effectiveness in face anti-spoofing.

Z. Kong et al. 2023, [15] address the challenge of generalizing presentation attack detection (PAD) techniques to unknown presentation attack instruments (PAIs). It highlights the importance of model initialization for generalization, which is often overlooked. To improve generalization, the paper proposes a self-supervised learning-based method called DF-DM. This method leverages a global-local view and employs de-folding and de-mixing techniques to derive task-specific representations for PAD. During de-folding, region-specific features are learned to represent samples in a local pattern, while de-mixing drives detectors to obtain instance-specific features with global information for more comprehensive representation. The proposed DF-DM method outperforms state-of-the-art methods in face and fingerprint PAD tasks on more complex and hybrid datasets, achieving a significant reduction in equal error rate (EER).

Li, C., Li, Z., Sun, J. et al. 2023, [16] propose a face anti-spoofing algorithm that focuses on utilizing shallow features to enhance the fine-grained information of the model. It introduces a "shortcut" structure to combine shallow features with middle layer features, improving the representation ability of details. The algorithm is initialized with pre-trained model parameters and then trained on balanced samples to enhance its classification ability. Additionally, it uses an RS Block based on depthwise separable convolution to reduce model parameters and floating-point operations. Experimental results on the CASIA-SURF dataset demonstrate superior performance with a low average classification error rate and high true positive rate at a low false positive rate.

Yousef Atoum, Yaojie Liu, Amin Jourabloo, and Xiaoming Liu 2017, [17] present a novel approach for detecting face spoofing attacks. By utilizing patch-based and depth-based convolutional neural networks (CNNs), the method effectively combines local texture information and geometric cues to distinguish between genuine and fake faces. Extensive evaluations on publicly available datasets demonstrate the superiority of their approach over existing methods, showcasing its improved accuracy and robustness in detecting various spoofing attacks. The proposed technique represents a significant contribution to enhancing the security and reliability of face recognition systems by effectively countering face spoofing threats.

Chien-Yi Wang, Yujiang Lu, S. Lai, [18] present a straightforward yet effective approach for detecting face spoofing attacks. The proposed framework, PatchNet, focuses on fine-grained patch recognition within the face region. It extracts small patches from the face and classifies them as genuine or spoofed based on their distinctive features. By focusing on fine-grained details, the method aims to capture subtle cues that differentiate between real and fake faces. The paper demonstrates the efficacy of PatchNet in detecting spoof attacks through evaluations on publicly available face anti-spoofing datasets. Its simplicity and computational efficiency make it a promising approach for real-world face recognition security applications.

A.George and S. Marcel, [19] introduce a Convolutional Neural Network (CNN) based framework for detecting presentation attacks in face recognition systems. The proposed approach incorporates deep pixel-wise supervision and operates on frame-level information, making it suitable for deployment on smart devices with minimal computational and time overhead. The effectiveness of the method is demonstrated on public datasets, achieving impressive results with a 0% Half Total Error Rate (HTER) on the Replay Mobile dataset and a 0.42% Attack Classification Error Rate (ACER) on Protocol-1 of the OULU dataset, outperforming state-of-the-art methods. The paper contributes to enhancing the security and reliability of face recognition technology in unattended scenarios.

A. Günay Yılmaz, U. Turhal, and V. Nabiyev 2023, [20] present a study on face presentation attack detection using multi-block Local Binary Pattern (LBP) features on different facial regions. The authors explore the effectiveness of these features in distinguishing between genuine and spoofed faces. They evaluate the performance of the proposed method on various datasets and conduct experiments on different facial regions to analyze their impact on the overall detection accuracy. The paper aims to improve face presentation attack detection by focusing on specific facial regions and employing multi-block LBP features.

The results and findings of this research contribute to the development of robust and reliable face recognition systems in the presence of presentation attacks.

V. RESEARCH METHOD

In the research method, researchers can produce relevant search queries related to their research subjects to initiate their literature investigation. To efficiently refine search results, the procedure entails applying criteria including publication date, citation count, and author affiliations. Two popular research platforms that make it easier to navigate and explore scholarly literature are Semantic Scholar and Dimensions.ai. Scholars can take advantage of Semantic Scholar's powerful citation network functionalities, which facilitate a thorough comprehension of academic relationships and citation trends. Furthermore, by combining various facets of linked subjects, Dimensions.ai offers researchers a broader perspective and offers insightful interdisciplinary information. Scholars can improve their literature review process and obtain significant insights that are in accordance with their study objectives by leveraging these unique characteristics.

A. Study Selection Procedure:

A subject search was carried out on face anti spoofing techniques and its synonyms in these databases: Web of Science, Research Gate, ACM, IEEE Xplore, IEEE IJCB, and Springer ICB. Using "face and spoofing" as the keywords in the search domain helps to preclude any non-face spoof detection. The search scope was also limited to English literature only. The selection included searching the literature source, then screening and filtering the obtained articles. Irrelevant and duplicate articles are being removed by checking the title and abstract [44]. B. Search: The search for literature in the aforementioned databases was done using the keywords "face recognition" with the "AND" operator and "spoof detection" with different synonyms such as anti-spoof, liveness detection, etc., as seen in figure 3 showing the query text [44].

```
("face recognition" AND "spoof detection") OR ("facial recognition" AND "spoofing")  OR ("biometric spoof detection") OR
    ("facial anti-spoofing") OR ("deep learning" AND "face spoof detection")
```

Fig. 3: Search query

B. Comparison Table

Table 1. Comparison Table for Different Existing Methods

Ref.	Year	Method	Dataset	Result	Remarks
[22]	2011	Multi-scale LBPs	NUAA	AUC = 0.99	analyses the texture of the facial images using multiscale local binary patterns (LBP)
[23]	2012	DoG	CASIA-FASD	EER = 0.17	explores the high frequency information in the facial region to determine the liveness.
[24]	2014	CNN	Replay Attack CASIA-MFSD	Intra DB: HTER=6.25.68, EER=6.23 on CASIA; HTER=2.68, EER=4.32 on CASIA Cross DB: HTER= 41.36, EER= 42.48 on Replay Attack HTER= 42.04, EER=41.86 on CASIA MFASD	the deep convolutional neural network is relied on to learn features of high discriminative ability in a supervised manner and combined with some data pre-processing, the face antispoofing performance improves drastically.
[25]	2017	Patch and Depthbased CNN	CASIA-FASD, MSU-USSA, and Replay Attack	AUC = 0.997 on Replay -Attack EER= 2.67%, HTER=2.27% on CASIA-FAS Database EER= 0.79%, HTER= 0.72% on Replay-Attack EER= 0.35 ± 0.19, HTER= 0.21 ± 0.21 on MSU-USSA	extracting the local features and holistic depth maps from the face images, which facilitate CNN to discriminate the spoof patches independent of the spatial face areas.
[26]	2018	CNN+RNN (Dept + rPPG)	SiW, Olulu-NPU, MSU-MFSD, Replay-Attack	Intra DB: ACER= 3.58 on SiW DB; ACER= 1.6 on OULU-NPU CrossDB: HTER= 27.6% for MSU-MFSD & HTER= 28.4% for Replay-Attack	introduces a new face anti-spoofing database that covers a large range of illumination, subject, and pose variations.
[27]	2018	Temporal and depth information	SiW database OULU-NPU CASIA-MFSD	Intra DB: ACER= 0.73 on SiW; ACER= 1.3 on OULU-NPU Cross DB: HTER = 17.5 on Repaly Attack DB; HTER = 24.0 on CASIA-MFSD	develop a new method to estimate depth informatio n from multiple RGB frames and propose a depth-supervised architecture which can efficiently encodes spatiotemporal informatio n for presentation attack detection.

[28]	2019	FeatherNet A/B	CASIA-SURF (MMFD	ACER=0.0013, TPR=0.999 @FPR=10e-2 TPR=0.998 @FPR=10e-3 TPR=0.9814 @FPR=10e-4	a novel fusion procedure with "ensemble + cascade" structure is presented to satisfy the performance preferred use cases.
[29]	2020	MRCNN	Print Attack DB, Replay Attack DB OULU-NPU DB and Spoof in the Wild (SiW) DB	HTER = 0.71 on Print Attack Database HTER= 1.6 on Replay Attack DB ACER= 1.9 on OULUNPU; ACER= 1.3 on SiW database	introduces the concept of local classification loss to local patches, so as to utilize the input information in the entire face region and to avoid over-emphasizing certain local areas.
[30]	2021	ANRL+ AFNM	OULU-NPU (O), CASIA-FASD (C), Idiap Replay-Attack (I), and MSU-MFSD (M).	HTER = 16.03% and AUC = 91.04% on O&C&M to I HTER = 10.83% and AUC = 96.75% on O&C&I to M HTER = 17.85% and AUC = 89.26% on O&M&I to C HTER = 15.67% and AUC = 91.90% on I&C&M to O	emphasizing the significance of normalization selection during feature extraction.
[31]	2021	D^2AM	OULU-NPU (O), CASIA-FASD (C), Idiap Replay-Attack (I), and MSU-MFSD (M).	HTER = 15.43% and AUC = 91.22% on O&C&M to I HTER = 12.70% and AUC = 95.66% on O&C&I to M HTER = 20.98% and AUC = 85.58% on O&M&I to C HTER = 15.27% and AUC = 90.87% on I&C&M to O	extract discriminative domain fea tures for clustering and designs a generalizable face antispoofing with meta-learning to enhance the interpretability through visualization.
[32]	2022	ICT+ ICT-Ref	MS-Celeb-1M	AUC= 87.01% on ICT AUC = 96.34% on ICT-Ref	Identity Consistency Tran sformer exhibits superior generalization ability not only across different datasets but also across various types of image degradation forms found in real-world applications including deepfake video.

[33]	2022	PatchNet	OULU-NPU (O), SiW (S), CASIAFASD (C), Replay-Attack (I), MSUMFSD (M)	HTER = 7.10% and AUC = 98.46% on O&C&M to I HTER = 11.33% and AUC = 94.58% on O&C&I to M HTER = 13.4% and AUC = 95.67% on O&M&I to C HTER = 11.82% and AUC = 95.07% on I&C&M to O	shows that the model is capable of recognizing unseen spoof types robustly by only looking at local regions.
[34]	2023	DGUA-FAS	CASIA-FASD, MSU-MFSD, Idiap Replayattack, OULU-NPU, CelebA-Spoof, and WMCA	AUC = 97.678% HTER = 7.1%	conduct the experiments on the effectiveness of in-distribution samples and out-of-distribution samples.
[35]	2023	IFAST	BNI-FAS	ACC = 99.41%, AUC = 0.99990%, EER = 0.3657%, TPR = 99.85% @ FPR= 1., TPR = 99.75% @ FPR= 0.5, TPR = 98.18%@ FPR= 0.1, TPR = 96.70% @ FPR= 0.05, TPR = 91.52% @ FPR= 0.01, TPR = 79.74% @ FPR= 0.001	proving the effectiveness of the single-shot FAS based on binocular NIR images.
[36]	2023	FLIP Cross-domain FAS (FLIP-Vision (FLIP-V) + FLIPImage-Text Similarity (FLIP-IT) + FLIP-Multimodal-Contrastive-Learning (FLIP-MCL))	MSU-MFSD, CASIA-MFSD, Idiap Replay Attack, and OULUNPU.	Avg. HTER = 3.48% IN FLIP-V, Avg. HTER = 3.06% IN FLIP-IT, Avg. HTER = 3.01 %IN FLIP-MCL	aligning the image representation with an ensemble of class descriptions (based on natural language semantic s) improves FAS generalizability in lowdata regimes, and a multimodal contrastive learning strategy is proposed to boost feature generalization further and bridge the gap between source and target domains.
[37]	2023	Distributional Estimation (DisE)	SuHiFiMask	AUC = 97.42 @ initial learning rate = 0.01% on SuHiFiMask dataset	a method that converts traditional FAS point estimation to distrib utional estimation by modeling data uncertainty during training, including feature (mean) and uncertainty (variance) .

[38]	2023	semi-supervised learning + LSTM	CASIA(C), Idiap REPLAY-ATTACK(I), OULU-NPU (O), and MSU-MFSD(M)	HTER = 3.10% and AUC = 99.95% on O&C&I to M HTER = 4.12% and AUC = 98.49% on O&M&I to C HTER = 7.37% and AUC = 99.18% on O&C&M to I HTER = 28.92% and AUC = 96.44% on I&C&M to O	Leveraging the benefits of semi-supervised learning, which considers both labeled and unlabeled apex frames to effectively discriminate between live and spoof classes.
[39]	2023	ViT-S-Adapter-TSR	CASIA-FASD (C), IDIAP REPLAY ATTACK (I), MSU MFSD (M), and OULU-NPU (O)	HTER = 3.43% and AUC = 99.50% on C&I&O to M HTER = 6.32% and AUC = 97.82% on O&M&I to C HTER = 7.16% and AUC = 97.61% on O&C&M to I HTER = 7.21% and AUC = 98.00% on I&C&M to O	S-Adapter employs the histogram information of transformer tokens and incorporates our proposed Token Style Regularization (TSR) to learn more domain-invariant feature representations.
[40]	2023	ViT-Euc-Hyp + MCDeepPixBiS-Euc-Hyp and ViT-HypHCL	Intra and Cross datasets: WMCA, PADISI-Face, and SiW-M, MSU-MFSD(M), IDIAP REPLAY-ATTACK(I), CASIA-FASD©, and OULU-NPU(O)	Intra-DB: ViT-Euc-Hyp model can increase AUC% from 94.90±1.49% to 96.10±0.82%. And MCDeepPixBiS-Euc-Hyp decreases HTER% from 17.04±2.07% to 14.76±0.88%. Cross-DB: AUC = 78.81% HTER = 29.97% on M&I to C AUC = 75.42% HTER = 32.09% on M&I to O	In a novel multimodal FAS framework consisting of Euclidean multimodal feature decomposition and hyperbolic multimodal feature fusion & classification is designed.
[41]	2023	CNN-RNN deep learning architecture	MSU-MFSD (M), Idiap Replay-Attack (I), CASIA (C) and OULU-NPU (O), intradataset: SMFMVD	Cross-DB: HTER= 23.88%, AUC = 99.78% on M&I to C HTER= 16.87%, AUC = 94.05% on M&I to O Intra-DB: APCER = 3.5%, BPCER = 0.1%, ACER = 2.3%, EER= 0.7% on SMFMVD Dataset	combining the strength of CNN in capturing spatial features and the temporal modeling capabilities of the recurrent neural network (RNN), in particular a Gated Recurrent Unit (GRU).
[42]	2024	DTDA	OULU-NPU (O), CASIA-FASD (C), Idiap Replay-Attack (I), MSU-MFSD (M) and CelebA-Spoof (S)	Protocol 1: Cross-DB: HTER = 9.66%, and AUC =95.45% Intra-DB: AUC = 99.75, APCER = 2.43, BPCER = 2.40, ACER = 2.42 Protocol 2: Cross-DB: HTER = 22.67%, and AUC = 82.43% on O&I to C, HTER = 28.22 %, and AUC = 79.44% on M&I to O Intra-DB: AUC = 99.89±0.16, APCER = 0.89±1.06, BPCER = 0.91±1.09, ACER = 0.90±1.07	The model can extract domain-invariant features for FAS.

[43]	2024	AAViT (MLP or AAMLP + TRANSFORMER ENCODER)	Replay-attack	HTER = 1.71% on replay-attack	Traditional vision transformer consists of two parts: transformer encode r and multi-layer perception (MLP). The former plays the role of feature learning to obtain better representation, while the latter plays the role of classification.

Face anti-spoofing datasets play a pivotal role in the development and evaluation of models designed to distinguish between genuine facial features and spoofing attempts. These datasets typically consist of a diverse collection of images and videos that simulate various types of attacks, such as using printed photos, replayed videos, or 3D masks.

Table 2. Compilation of datasets captured using commercial RGB cameras.[21]

Dataset	Year	Live/Spoof	Subject	File type	Setup	Spoofing Types
NUAA	2010	5105/7509	15	Image	N/R	Print(flat, wrapped)
YALE Recaptured	2011	640/1920	10	Image	50cm-distance from 3 LCD minitors	Print(flat)
CASIAMFSD	2012	150/450	50	Video	7 scenarios and 3 image quality	Print(flat, wrapped, cut), Replay(tablet)
REPLAY-ATTACK	2012	200/1000	50	Video	Lighting and holding	Print(flat), Replay(tablet, phone)
Kose and Dugelay	2013	200/198	20	Video	N/R	Mask(hard resin)
MSU-MFSD	2014	70/210	35	Video	Indoor scenario; 2 types of cameras	Print(flat), Replay(tablet, phone)
UVAD	2015	808/16268	404	Video	Different lighting, background and places in two sections	Replay(monitor)
REPLAY-Mobile	2016	390/640	40	Video	5 lighting conditions	Print(flat), Replay(monitor)
HKBU-MARs V2	2016	504/504	12	Video	7 cameras from stationary and mobile devices and 6 lighting settings	Mask(hard resin) from Thatsmyface and REAL-f
MSU USSA	2016	1140/9120	1140	Image	Uncontrolled; 2 types of cameras	Print(flat), Replay(laptop, tablet, phone)
SMAD	2017	65/65	-	Video	Color images from online resources	Mask(silicone)
OULU-NPU	2017	720/2880	55	Video	Lighting & background in 3 sections	Print(flat), Replay(phone)
Rose-Youtu	2018	500/2850	20	Video	5 front-facing phone camera; 5 different illumination conditions	Print(flat), Replay(monitor, laptop),Mask(paper, croppaper)
SiW	2018	1320/3300	165	Video	4 sessions with variations of distance, pose, illumination and expression	Print(flat, wrapped), Replay(phone, tablet, monitor)
WFFD	2019	2300/2300(I) 140/145(V)	745	Image	Collected online; superrealistic; removed lowquality faces	Waxworks(wax)

SiW-M	2019	660/968(V)	493	Video	Indoor environment with pose, lighting and expression variations	Print(flat), Replay, Mask(hard resin, plastic, silicone, paper, Mannequin), Makeup(cosmetics, impersonation, Obfuscation), Partial(glasses, cut paper)
Swax	2020	Total 1812(I) 110(V)	55	Image and video	Collected online; captured under uncontrolled scenarios	Waxworks(wax)
CelebA-Spoof	2020	156384/469153(I)	10177	Video	4 illumination conditions; indoor & outdoor; rich annotations	Print(flat, wrapped), Replay(monitor tablet, phone), Mask(paper)
RECOD-Mtablet	2020	450/1800(V)	45	Image, audio, and video	Outdoor environment and low-light & dynamic sessions	Print(flat), Replay(monitor)
CASIA-SURF 3DMask	2020	288/864(V)	48	Video	High-quality identitypreserved; 3 decorations and 6 environments	Mask(mannequin with 3D print)
HiFiMask	2021	13650/40950(V)	75	Video	three mask decorations; 7 recording devices; 6 lighting conditions; 6 scenes	Mask(transparent, plaster, resin)
SiW-M v2	2022	785/915 (V)	1093(493/600)	Video	Both indoor and outdoor, diverse age and enthnicity, 7 illumiations	IAPRA-verified 14 spoof attacks (4 coverings, 3 makeups, 3 masks, 2 human models, replay and print)
SuHiFiMask	2022	10195/10195 (V)	101	Video	Long distance using Surveillance cameras, recording in 3 scenes, and 3 lightings, 4 whethers	2D image, Video replay, 3D Mask with materials Resin, Plaster, Silicone, Paper
WFAS	2023	529,571/ 853,729 (I)	469,920	image	Internet, unconstrained settings	17 PAs, Print(newspaper, poster, photo, album, picture book, scan photo, packging, cloth), Display(phone, tablet, TV, computer), Mask, 3D Model(garage kit, doll, adult doll, waxwork)

Table 2 showcases a compilation of datasets captured using commercial RGB cameras that is integral to the development and analysis of models designed for face anti-spoofing.

VI. LIMITATIONS AND DISCUSSION

While the reviewed literature presents various approaches to face anti-spoofing, there are several limitations and considerations that need to be acknowledged. Firstly, the diversity in datasets used across different studies may impact the generalisation of proposed methods. Some models might excel on specific datasets but struggle with others, highlighting the importance of cross-dataset evaluation for robustness. The evaluation metrics also vary among studies, with some emphasising false acceptance rates (FAR), false rejection rates (FRR), or area under the curve (AUC). This makes direct comparisons challenging, and a standardised evaluation framework could enhance the comparability of different methods.

Additionally, the performance of face anti-spoofing systems can be influenced by factors such as environmental conditions, illumination, and pose variations.

Many existing works focus on specific aspects of spoofing attacks, and their effectiveness in handling diverse real-world scenarios remains an open question. Another aspect worth considering is the trade-off between accuracy and computational efficiency. Deep learning models, especially convolutional neural networks (CNNs), may demand substantial computational resources, limiting their practicality in real-time applications or resource-constrained devices. Moreover, the continuous evolution of spoofing techniques poses a challenge for the development of robust countermeasures. As spoofing methods advance, the effectiveness of existing anti-spoofing approaches may diminish, necessitating ongoing research and adaptation.

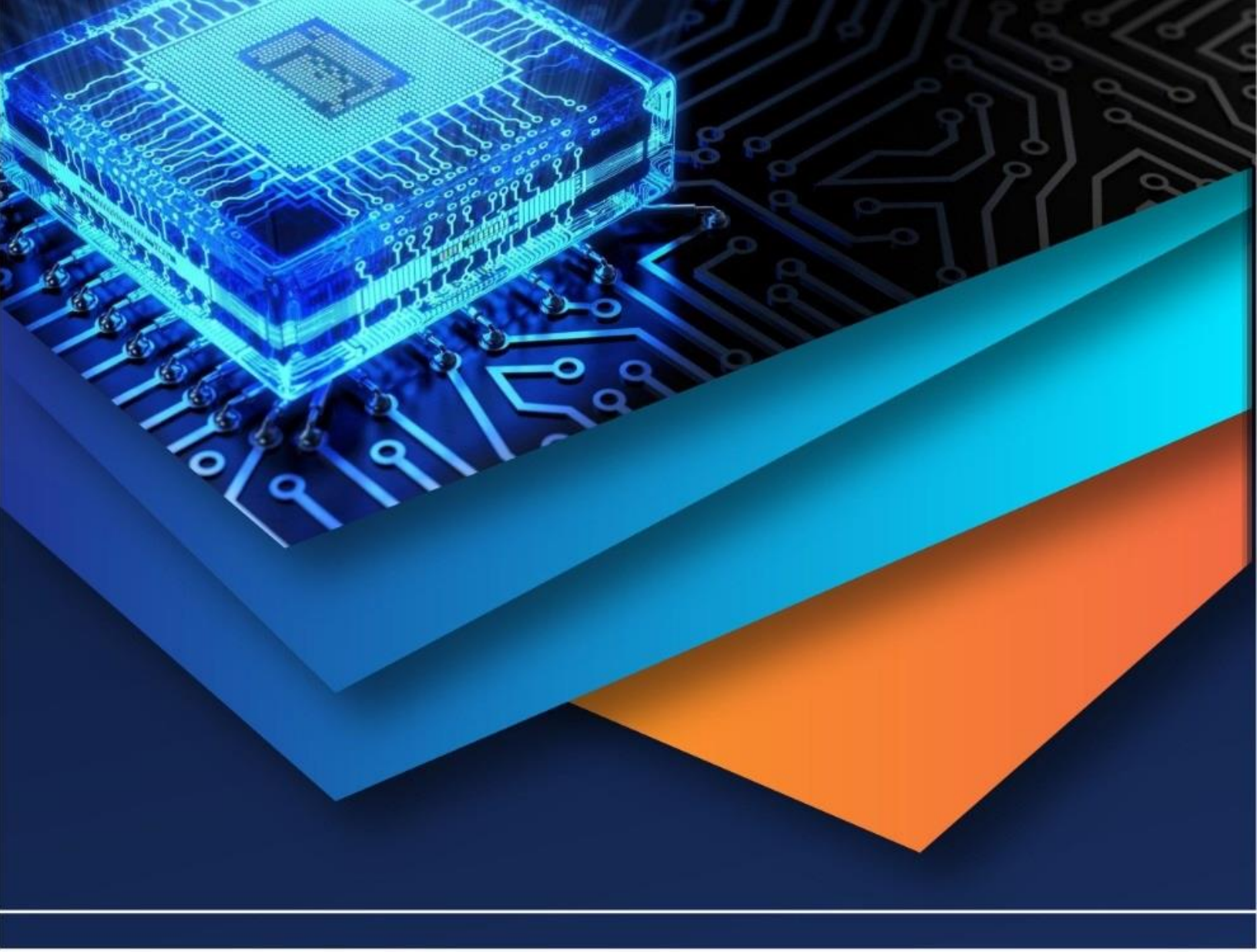
VII. CONCLUSION

In conclusion, the review of existing face anti-spoofing methods reveals a diverse landscape of techniques that leverage advancements in computer vision and deep learning. The comparative table provides insights into the strengths and limitations of different approaches, highlighting the importance of considering factors such as dataset diversity, evaluation metrics, and realworld applicability. The introduction of novel architectures, feature extraction techniques, and fusion procedures have enabled in considerable advancements in the field. However, the limitations that have been mentioned highlight the need for further study to deal with issues with dataset standardisation, cross-dataset assessment, resilience in diverse environmental settings, and computational effectiveness. As spoofing attacks continue to evolve, ongoing collaboration and innovation within the research community are essential to developing face anti-spoofing solutions that are not only effective against current threats but also adaptable to emerging challenges. The quest for more robust, generalizable, and computationally efficient face anti-spoofing methods remains a dynamic area of research with the potential for real-world impact.

REFERENCES

[1] Dong, Xiaoyi et al. “Protecting Celebrities from DeepFake with Identity Consistency Transformer.” 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2022): 9458-9468.
[2] Anthony, Peter et al. “A Review of Face Anti-spoofing Methods for Face Recognition Systems.” 2021 International Conference on INnovations in Intelligent SysTems and Applications (INISTA) (2021): 1-9.
[3] Anjos, André, and Sébastien Marcel. “Counter-measures to photo attacks in face recognition: A public database and a baseline.” 2011 International Joint Conference on Biometrics (IJCB) (2011): 1–7.
[4] Määttä, Jukka et al. “Face spoofing detection from single images using micro-texture analysis.” 2011 International Joint Conference on Biometrics (IJCB) (2011): 1-7.
[5] Komulainen, Jukka et al. “Face Spoofing Detection Using Dynamic Texture.” ACCV Workshops (2012).
[6] Zhang, Zhiwei et al. “A face antispoofing database with diverse attacks.” 2012 5th IAPR International Conference on Biometrics (ICB) (2012): 26-31.
[7] Chen, Cunjian et al. “Spoofing faces using makeup: An investigative study.” 2017 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA) (2017): 1-8.
[8] Mou, Lichao et al. “Learning Spectral-Spatial-Temporal Features via a Recurrent Convolutional Neural Network for Change Detection in Multispectral Imagery.” IEEE Transactions on Geoscience and Remote Sensing 57 (2018): 924-935.
[9] Muhammad, Usman et al. “Deep Ensemble Learning with Frame Skipping for Face Anti-Spoofing.” 2023 Twelfth International Conference on Image Processing Theory, Tools and Applications (IPTA) (2023): 1-6.
[10] Solomon, Enoch & Cios, Krzysztof. (2023). FASS: Face Anti-Spoofing System Using Image Quality Features and Deep Learning. Electronics. 12. 2199.10.3390/electronics12102199.
[11] Sharma, Deepika & Selwal, Arvind. (2023). A survey on face presentation attack detection mechanisms: hitherto and future perspectives. Multimedia Systems. 29. 1-51. 10.1007/s00530-023-01070-5.
[12] L. Birla, P. Gupta and S. Kumar, "SUNRISE: Improving 3D Mask Face Anti-Spoofing for Short Videos Using Pre-Emptive Split and Merge," in IEEE Transactions on Dependable and Secure Computing, vol. 20, no. 3, pp. 1927-1940, 1 May-June 2023, doi: 10.1109/TDSC.2022.3168345.
[13] Xin Cheng, Jingmei Zhou, Xiangmo Zhao, Hongfei Wang, and Yuqi Li. 2023. A presentation attack detection network based on dynamic convolution and multi-level feature fusion with security and reliability. Future Gener. Comput. Syst. 146, C (Sep 2023), 114–121. https://doi.org/10.1016/j.future.2023.04.012
[14] Zhang S, Nie W. Multi-Domain Feature Alignment for Face Anti-Spoofing. Sensors. 2023; 23(8):4077. https://doi.org/10.3390/s23084077
[15] Z. Kong et al., "Taming Self-Supervised Learning for Presentation Attack Detection: De-Folding and De-Mixing," in IEEE Transactions on Neural Networks and Learning Systems, doi: 10.1109/TNNLS.2023.3243229.
[16] Li, C., Li, Z., Sun, J. et al. Middle-shallow feature aggregation in multimodality for face anti-spoofing. Sci Rep 13, 9870 (2023). https://doi.org/10.1038/s41598-023-36636-w
[17] Yousef Atoum, Yaojie Liu, Amin Jourabloo, and Xiaoming Liu. Face anti-spoofing using patch and depth-based cnns. In 2017 IEEE International Joint Conference on Biometrics (IJCB), pages 319–328, 2017.
[18] Chien-Yi Wang, Yujiang Lu, S. Lai. "PatchNet: A Simple Face Anti-Spoofing Framework via Fine-Grained Patch Recognition." Computer Science. In Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 27 March 2022.
[19] A. George and S. Marcel, "Deep Pixel-wise Binary Supervision for Face Presentation Attack Detection," 2019 International Conference on Biometrics (ICB), Crete, Greece, 2019, pp. 1-8, doi: 10.1109/ICB45273.2019.8987370.
[20] Günay Yılmaz, Asuman & Turhal, Ugur & Nabiyev, Vasif. (2023). Face presentation attack detection performances of facial regions with multi-block LBP features. Multimedia Tools and Applications. 1-25. 10.1007/s11042-023-14453-7.

[21] https://github.com/ZitongYu/DeepFAS?tab=readme-ov-file#data
[22] J. Määttä, A. Hadid, and M. Pietikäinen, "Face spoofing detection from single images using micro-texture analysis," in 2011 International Joint Conference on Biometrics (IJCB), Oct. 2011, pp. 1–7, doi: 10.1109/IJCB.2011.6117510.
[23] Z. Zhang, J. Yan, S. Liu, Z. Lei, D. Yi, and S. Li, "A face anti spoofing database with diverse attacks," 2012 5th IAPR Int. Conf. Biometrics, pp. 26–31, 2012.
[24] J. Yang, Z. Lei, and S. Li, "Learn Convolutional Neural Network for Face AntiSpoofing," ArXiv, vol. abs/1408.5, 2014.
[25] Y. Atoum, Y. Liu, A. Jourabloo, and X. Liu, "Face anti-spoofing using patch and depthbased CNNs," in 2017 IEEE International Joint Conference on Biometrics (IJCB), Oct. 2017, pp. 319–328, doi: 10.1109/BTAS.2017.8272713.
[26] Y. Liu, A. Jourabloo, and X. Liu, "Learning Deep Models for Face Anti-Spoofing: Binary or Auxiliary Supervision," Jun. 2018.
[27] Z. Wang, C. Zhao, Y. Qin, Q. Zhou, and Z. Lei, "Exploiting temporal and depth information for multi-frame face anti-spoofing," ArXiv, vol. abs/1811.0, 2018.
[28] P. Zhang et al., "FeatherNets: Convolutional neural networks as light as feather for face anti-spoofing," in IEEE/CVF Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2019, vol. 2019-June, pp. 1574–1583, doi: 10.1109/CVPRW.2019.00199.
[29] Y. Ma, L. Wu, Z. Li, and F. liu, "A novel face presentation attack detection scheme based on multi-regional convolutional neural networks," Pattern Recognit. Lett., vol. 131, pp. 261–267, 2020.
[30] Liu, Shubao et al. "Adaptive Normalized Representation Learning for Generalizable Face Anti-Spoofing." Proceedings of the 29th ACM International Conference on Multimedia (2021): n. pag.
[31] Chen, Z., Yao, T., Sheng, K., Ding, S., Tai, Y., Li, J., Huang, F., & Jin, X. (2021). Generalizable Representation Learning for Mixture Domain Face AntiSpoofing. ArXiv, abs/2105.02453.
[32] Dong, X., Bao, J., Chen, D., Zhang, T., Zhang, W., Yu, N., Chen, D., Wen, F., & Guo, B. (2022). Protecting Celebrities from DeepFake with Identity Consistency Transformer. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 9458-9468.
[33] Wang, Chien-Yi et al. "PatchNet: A Simple Face Anti-Spoofing Framework via FineGrained Patch Recognition." 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2022): 20249-20258.
[34] Hong, Z., Lin, Y., Liu, H., Yeh, Y., & Chen, C. (2023). Domain-Generalized Face AntiSpoofing with Unknown Attacks. 2023 IEEE International Conference on Image Processing (ICIP), 820-824.
[35] Huang, J., Zhou, D., & Chen, S. (2023). IFAST: Weakly Supervised Interpretable Face Anti-spoofing from Single-shot Binocular NIR Images. ArXiv, abs/2309.17399.
[36] Srivatsan, K., Naseer, M., & Nandakumar, K. (2023). FLIP: Cross-domain Face Antispoofing with Language Guidance. ArXiv, abs/2309.16649.
[37] Huang, M. (2023). Distributional Estimation of Data Uncertainty for Surveillance Face Anti-spoofing. ArXiv, abs/2309.09485.
[38] Muhammad, U., Oussalah, M., & Laaksonen, J.T. (2023). Semi-Supervised learning for Face Anti-Spoofing using Apex frame. ArXiv, abs/2309.04958.
[39] Cai, R., Yu, Z., Kong, C., Li, H., Chen, C., Hu, Y., & Kot, A. (2023). S-Adapter: Generalizing Vision Transformer for Face Anti-Spoofing with Statistical Tokens. ArXiv, abs/2309.04038.
[40] Han, S., Cai, R., Cui, Y., Yu, Z., Hu, Y., & Kot, A.C. (2023). Hyperbolic Face AntiSpoofing. ArXiv, abs/2308.09107.
[41] Muhammad, U., Oussalah, M., Hoque, M.Z., & Laaksonen, J.T. (2023). Saliencybased Video Summarization for Face Anti-spoofing. ArXiv, abs/2308.12364.
[42] Kong, Z., Zhang, W., Wang, T., Zhang, K., Li, Y., Tang, X., & Luo, W. (2024). Dual Teacher Knowledge Distillation with Domain Alignment for Face Anti-spoofing.
[43] Yang, J., Chen, F., Das, R., Zhu, Z., & Zhang, S. (2024). Adaptive-avg-pooling based Attention Vision Transformer for Face Anti-spoofing.
[44] Anthony, P., Ay, B., & Aydin, G. (2021). A Review of Face Anti-spoofing Methods for Face Recognition Systems. 2021 International Conference on INnovations in Intelligent SysTems and Applications (INISTA), 1-9.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY