International Journal of Computer Vision
https://doi.org/10.1007/s11263-025-02411-9

Guest Editorial: Special Issue on Biometrics Security and Privacy

Jun Wan[1] · Arun Ross[2] · Sergio Escalera[3]

1 Introduction

This special issue is dedicated to the pivotal theme of biometric security and privacy (BSP). Biometrics, the science of identifying individuals based on their biological and behavioral traits-such as facial features, fingerprints, iris patterns, voice, gait, and palmprints-has been rapidly integrated into various applications. These applications span both identification systems, such as face recognition for mobile payments, and verification systems, such as fingerprint recognition for border control. However, the increasing deployment of biometric technologies across diverse domains has raised substantial concerns regarding the security of individuals and the privacy of biometric data itself. Moreover, recent advancements in generating synthetic images and videos, coupled with high-fidelity physical and digital presentation attacks, have further exacerbated these concerns. These innovations pose significant threats to the integrity of biometric systems, giving rise to potential risks such as identity theft, the fabrication of false identities, and heightened vulnerabilities to both system security and personal privacy.

In light of these challenges, this Special Issue presents 31 papers on the critical topic of Biometric Security and Privacy (BSP), organized into four key sections: Physical and Digital Attack Detection, Gait/Hand Gesture Authentication, Pedestrian Re-Identification, and Others. This special issue provides a comprehensive overview of the current state of research in biometric security and privacy, to advance understanding in these crucial areas and offer solutions to the increasingly complex challenges posed by the evolving landscape of biometric technologies.

✉ Sergio Escalera
sergio@maia.ub.es

Jun Wan
jun.wan@ia.ac.cn

Arun Ross
rossarun@cse.msu.edu

[1] Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

[2] Michigan State University, East Lansing, MI, USA

[3] Universitat de Barcelona and Computer Vision Center, 08007 Barcelona, Catalonia, Spain

2 Overview of Accepted Papers

2.1 Physical and Digital Attack Detection

This part of the issue consists of 23 papers.

The first article, by Ajian Liu, presents the CA-MoEiT framework, which enhances Face Anti-Spoofing (FAS) using Vision Transformers (ViT) through feature augmentation, alignment, and a semi-fixed Mixture-of-Experts mechanism for improved generalization across domains.

The second article, by Keyao Wang, Guosheng Zhang, Haixiao Yue, Yanyan Liang, Mouxiao Huang, Gang Zhang, Junyu Han, Errui Ding and Jingdong Wang, which treats domain generalization as an anomaly detection problem and uses a Dynamic Feature Queue, Domain Alignment Module, and Progressive Training Strategy for improved generalization.

The third article, by Usman Muhammad, Jorma Laaksonen, Djamila Romaissa Beddiar and Mourad Oussalah, presents a Face Presentation Attack Detection (PAD) solution that combines synthetic data generation and deep ensemble learning to enhance generalization, using diverse training subsets and a meta-model to improve performance on unseen domains.

The fourth article, by Xing Liu, Anyang Su, Minghui Wu, Zitong Yu, Kangle Wu, Da An, Jie Hao, Mengzhen Xu, Chenxu Zhao and Zhen Lei, introduces CG-FAS, a data augmentation method for Face Anti-Spoofing (FAS) that generates diverse, high-quality training data by transforming live faces into various presentation attacks and vice versa using an Interchange Bridge matrix.

The fifth article, by Yongluo Liu, Zun Li, Yaowen Xu, Zhizhi Guo, Zhaofan Zou and Lifang Wu, presents the QIDG method for Face Anti-Spoofing (FAS), using a teacher-student architecture to align liveness features into a quality-invariant space, with Dual Adversarial Learning and

Published online: 20 March 2025

Quality Feature Assembly modules to enhance domain generalization despite variations in face quality.

The sixth article, by Shiyun Mao, Ruolin Chen and Huibin Li, introduces WJDOT-FAS, an unsupervised multi-source domain adaptation method for Face Anti-Spoofing (FAS) that addresses domain bias by using joint distribution estimation, optimal transport, and domain weight optimization to align source and target domains.

The seventh article, by Fangling Jiang, Qi Li, Weining Wang, Min Ren, Wei Shen, Bing Liu and Zhenan Sun, proposes an open-set single-domain generalization approach for Face Anti-Spoofing (FAS), introducing a causal generalized representation learning framework to handle both known and unknown attacks in unseen domains through causality-inspired domain augmentation and unknown-aware probability calibration.

The eighth article, by Zitong Yu, Rizhao Cai, Yawen Cui, Xin Liu, Yongjian Hu and Alex C. Kot, explores the application of Vision Transformers (ViT) to multimodal Face Anti-Spoofing (FAS), addressing input configurations, pre-training, and fine-tuning challenges, and introducing the M^2A^2E self-supervised pre-training method for task-aware representation learning.

The ninth article, by Min Ren, Yunlong Wang, Yuhao Zhu, Yongzhen Huang, Zhenan Sun, Qi Li and Tieniu Tan, presents an adversarial defense method for face recognition inspired by the immune system, using sample-specific "antibodies" and self-supervised adversarial training to counter imperceptible attacks.

The tenth article, by Xiao Yang, Longlong Xu, Tianyu Pang, Yinpeng Dong, Yikai Wang, Hang Su and Jun Zhu, a method that uses 3D face modeling to generate more effective and natural adversarial patches for face recognition, overcoming the limitations of 2D and 3D attacks by simulating complex facial transformations.

The eleventh article, by YenLung Lai, XingBo Dong, Zhe Jin, Wei Jia, Massimo Tistarelli and XueJun Li, introduces U-Sketch, a method for error correction in biometric data that transforms distributions into i.i.d. data, enhancing decoder security and interpretability while ensuring optimal error tolerance and robust security for biometric storage and key derivation.

The twelfth article, by Xin Li, Rongrong Ni, Yao Zhao, Yu Ni and Haoliang Li, introduces the Multi-Teacher Universal Distillation framework, which defends against facial manipulation attacks by embedding a warning image and using multiple teacher networks to guide a student network, enhancing robustness without degrading image quality.

The thirteenth article, by Yihui Li, Yifan Zhang, Hongyu Yang, Binghui Chen and Di Huang, presents the SA^3WT method for Face Forgery Detection, combining wavelet-based attention and fine-grained sampling with self-paced auto augmentation to enhance feature extraction and model robustness.

The fourteenth article, by Junxian Duan, Yuang Ai, Jipeng Liu, Shenyuan Huang, Huaibo Huang, Jie Cao and Ran He, presents the Dynamic Dual-spectrum Interaction Network for face forgery detection, combining RGB and frequency features through attention and dynamic fusion modules, while incorporating uncertainty guidance and spatial-frequency prompt learning for improved generalization.

The fifteenth article, by Junyi Cao, Ke-Yue Zhang, Taiping Yao, Shouhong Ding, Xiaokang Yang and Chao Ma, introduces a dual-space reconstruction learning framework for face recognition, addressing both manipulated artifacts and spoofing attacks by modeling genuine faces in spatial and frequency domains, with dynamic filtering and consistency-regularized training for improved performance and generalization.

The sixteenth article, by Yuting Xu, Jian Liang, Lijun Sheng and Xiao-Yu Zhang, presents Thumbnail Layout (TALL), an efficient deepfake detection method that transforms video clips into thumbnail layouts, preserving spatial and temporal dependencies, and introduces TALL++ with Graph Reasoning and Semantic Consistency Loss for improved generalization.

The seventeenth article, by Ke Sun, Shen Chen, Taiping Yao, Xiaoshuai Sun, Shouhong Ding and Rongrong Ji, presents the Historical Distribution Preserving (HDP) framework for Continual Face Forgery Detection (CFFD), using universal adversarial perturbation and knowledge distillation to detect new and past forgery attacks while preserving face data distributions.

The eighteenth article, by Yuan Wang, Chen Chen, Ning Zhang and Xiyuan Hu, presents the WATCHER Learning framework for deepfake detection, using a Wavelet-guided AutoEncoder and multi-level attention maps to capture fine-grained features and improve generalization to realistic forgeries through semantic reasoning.

The nineteenth article, by Yang Yu, Rongrong Ni, Siyuan Yang, Yu Ni, Yao Zhao and Alex C. Kot, introduces a generalized framework for face forgery detection that captures dynamic inconsistencies using multi-timescale cues, amplifying short-term discrepancies and employing inter-group graph learning for long-term detection, with a domain alignment module to improve generalization.

The twentieth article, by Nanqing Xu, Weiwei Feng, Tianzhu Zhang and Yongdong Zhang, presents FD-GAN, a generative adversarial network for forgery detection that improves generalization and robustness by using two generators for synthetic image creation and a discriminator with spatial and frequency branches to filter adversarial corruption.

The twenty-first article, by Qilin Yin, Wei Lu, Xiaochun Cao, Xiangyang Luo, Yicong Zhou and Jiwu Huang, introduces a graph attention network for fine-grained multimodal deepfake classification, using a heterogeneous graph to represent audio-visual samples and a cross-modal interaction module to exploit synchronization patterns, improving detection and classification of multimodal forgeries.

The twenty-second article, by Huan Liu, Zichang Tan, Qiang Chen, Yunchao Wei, Yao Zhao and Jingdong Wang, presents the UFAFormer framework for multi-modal media manipulation detection, using both image and frequency domains to capture fine-grained forgery features through wavelet transform, self-attention mechanisms, and cross-modal integration for enhanced detection.

The twenty-third article, by Jingzhi Li, Changjiang Luo, Hua Zhang, Yang Cao, Xin Liao and Xiaochun Cao, introduces Anti-Fake Vaccine, a privacy-preserving framework to defend against face-swapping attacks by combining visual corruption and identity misdirection, using multi-objective optimization and an additive perturbation strategy to strengthen protection.

2.2 Gait/Hand Aesture Authentication

This part of the issue consists of 2 papers.

The first article, by Saihui Hou, Zengbin Wang, Man Zhang, Chunshui Cao, Xu Liu and Yongzhen Huang, proposes an imperceptible adversarial attack strategy for deep gait recognition, injecting noise into silhouette edges to create sparse attacks across spatial and temporal dimensions, ensuring high success rates with flexibility in untargeted and targeted modes.

The second article, by Wenwei Song, Wenxiong Kang, Adams Wai-Kin Kong, Yufeng Zhang and Yitao Qiao, introduces the Linear Adaptive Additive Angular Margin (L3AM) loss function for hand gesture authentication, enhancing margin-based Softmax losses with a linear similarity function and an adaptive margin scheme for more stable and efficient training.

2.3 Pedestrian Re-identification

This part of the issue consists of 2 papers.

The first article, by Junyao Gao, Xinyang Jiang, Shuguang Dou, Dongsheng Li, Duoqian Miao and Cairong Zhao, presents a novel membership inference attack method for person re-identification, analyzing distribution shifts in inter-sample similarity to assess privacy risks, with strategies for both "one-to-one" and “one-to-any” attack scenarios.

The second article, by Yukang Zhang, Yan Yan, Yang Lu and Hanzi Wang, introduces the Adaptive Middle-modality Alignment Learning (AMML) method for visible-infrared person re-identification, reducing modality gaps by using an Adaptive Middle-modality Generator and aligning feature distributions with adaptive losses to enhance cross-modality knowledge.

2.4 Others

This part of the issue consists of 4 papers.

The first article, by Kaiduo Zhang, Muyi Sun, Jianxin Sun, Kunbo Zhang, Zhenan Sun and Tieniu Tan, presents Human-Diffusion, a diffusion-based framework for open-vocabulary, text-driven human image generation, incorporating Stylized Memory Retrieval and Multi-scale Feature Mapping modules to improve accuracy and handle arbitrary text inputs for diverse, high-fidelity human images.

The second article, by Ziyuan Yang, Andrew Beng Jin Teoh, Bob Zhang, Lu Leng and Yi Zhang, introduces PSFed-Palm, a physics-driven federated learning method for palmprint verification that ensures spectrum consistency and privacy protection by aligning local models with spectrum-specific anchors while preventing model drift.

The third article, by Jiyang Guan, Jian Liang, Yanbo Wang and Ran He, introduces SAC-JC, a novel model stealing detection method for deep face recognition that utilizes pairwise sample relationships and the correlation matrix of JPEG-compressed outputs to detect attacks, offering robustness against adversarial defenses and transfer learning.

The fourth article, by Xiao Guo, Xiaohong Liu, Iacopo Masi and Xiaoming Liu, presents HiFi-Net++, a hierarchical fine-grained representation learning framework for unified image forgery detection and localization (IFDL), which addresses the challenge of significant forgery attribute differences by incorporating a multi-branch feature extractor, CLIP-based multi-modal inputs, and pixel-level detection.

3 Conclusion

The aim of this guest editorial is to introduce the special issue on Biometric Security and Privacy (BSP). The papers accepted for this issue cover four key subtopics of BSP: Physical and Digital Attack Detection, Gait/Hand Gesture Authentication, Pedestrian Re-Identification, and Others. These papers are presented in the context of recent advancements in the field, as reviewed in this editorial.

Limitations and Challenges of BSP. While the accepted papers advance the state of the art in Biometric Security and Privacy (BSP), several limitations and challenges remain. Firstly, Vision Foundation Models (VFM), such as the Stable Diffusion model, have demonstrated impressive results in BSP research. However, while fine-tuning VFMs for downstream tasks has been widely explored, their application in Physical and Digital Attack Detection remains relatively underdeveloped. Secondly, the current adversarial attack

strategies, while effective at imperceptibly disrupting recognition systems, still exhibit a strong dependence on specific model architectures. A key challenge moving forward is to maintain the effectiveness of these attacks across different gait recognition models, particularly when confronted with more robust systems. This is an area that requires further investigation.

Finally, a critical challenge in the field is addressing the wide range of real-world attack strategies and spoofing techniques. For instance, Physical and Digital Attack Detection systems must continually adapt to emerging spoofing methods, such as high-quality deepfakes and 3D masks, which may evade detection by conventional models. Similarly, gait and hand gesture authentication methods remain vulnerable to environmental variations, such as changes in lighting or camera angles, which can negatively impact accuracy. Developing robust models that maintain performance under these real-world variations, while ensuring both accuracy and privacy, remains a significant and unresolved challenge.

We anticipate that the papers presented in this special issue, along with the comprehensive review of recent advancements discussed in this editorial, will offer a holistic overview of the current state of the art in the field of BSP. This collective effort seeks to promote further progress and stimulate innovation in this dynamic and rapidly evolving domain.