Multi-Domain Feature Alignment for Face Anti-Spoofing

Subjects: Computer Science, Artificial Intelligence

Contributor: Shizhe Zhang , Wenhui Nie

Face anti-spoofing is critical for enhancing the robustness of face recognition systems against presentation attacks. Existing methods predominantly rely on binary classification tasks. An adversarial learning process is designed to narrow the differences between domains, achieving the effect of aligning the features of multiple sources, thus resulting in multi-domain alignment.

multi-domain feature alignment domain generalization (MADG) face anti-spoofing

feature alignment multiple source domain

1. Introduction

With the extensive use of deep learning in computer vision, face recognition (FR) [1][2] technology has become increasingly important in daily life, particularly in scenarios that require user identification and authorization. Despite significant progress in FR, these systems remain susceptible to various types of attacks, such as print attacks, replayed video attacks, and 3D mask attacks. To address these challenges, current state-of-the-art research has proposed various methods for face anti-spoofing (FAS) [3][4][5][6]. These methods can be broadly categorized into two groups: hand-crafted feature-based and deep learning feature-based approaches.

Despite the notable achievements of previous face anti-spoofing (FAS) methods in intra-domain testing, their performance significantly deteriorates in cross-domain testing. This is primarily due to the introduction of bias resulting from the distinct characteristics of domains and the inability to address such bias by considering their internal relationships. Consequently, the generalization effect of the model on the novel domain is insufficient. To mitigate this limitation, recent studies have utilized unsupervised-learning-based domain adaptation (DA) techniques to eliminate the domain bias between source and target domains. Nevertheless, the target domain usually denotes an unseen domain for the source domain, and acquiring an adequate amount of target domain data for training in real-world scenarios is not typically feasible.

In response to the weakness of a model’s generalization in the unseen domain, several studies related to domain generalization (DG) have been proposed. Conventional DG [7] proposed a novel multi-adversarial discriminative method to learn a discriminative multi-domain feature space and improve the generalization performance. This method aimed to find a feature space with multiple source domains aligned, assuming that the feature extracted from the unseen domain can map near that feature space, indicating that the model is well generalized over the target domain. However, Conventional DG fails to account for the difference in generalization between real and fake faces. To address this limitation, SSDG [8] proposed an end-to-end single-side method that separates the feature distributions of real and fake faces by pulling all the real faces and pushing the fake faces of different domains, resulting in a higher classification capacity than Conventional DG. However, this method is insufficient in eliminating the interference of domain-related cues in the domain generalization problem. Motivated by the above works, this study aims to align the feature spaces corresponding to multiple source domains while dividing the distribution of fake and real faces. As shown in **Figure 1**, the method aligns multiple source domains to obtain a more generalizable feature space, outperforming the baseline method in terms of generalization and classification performance.

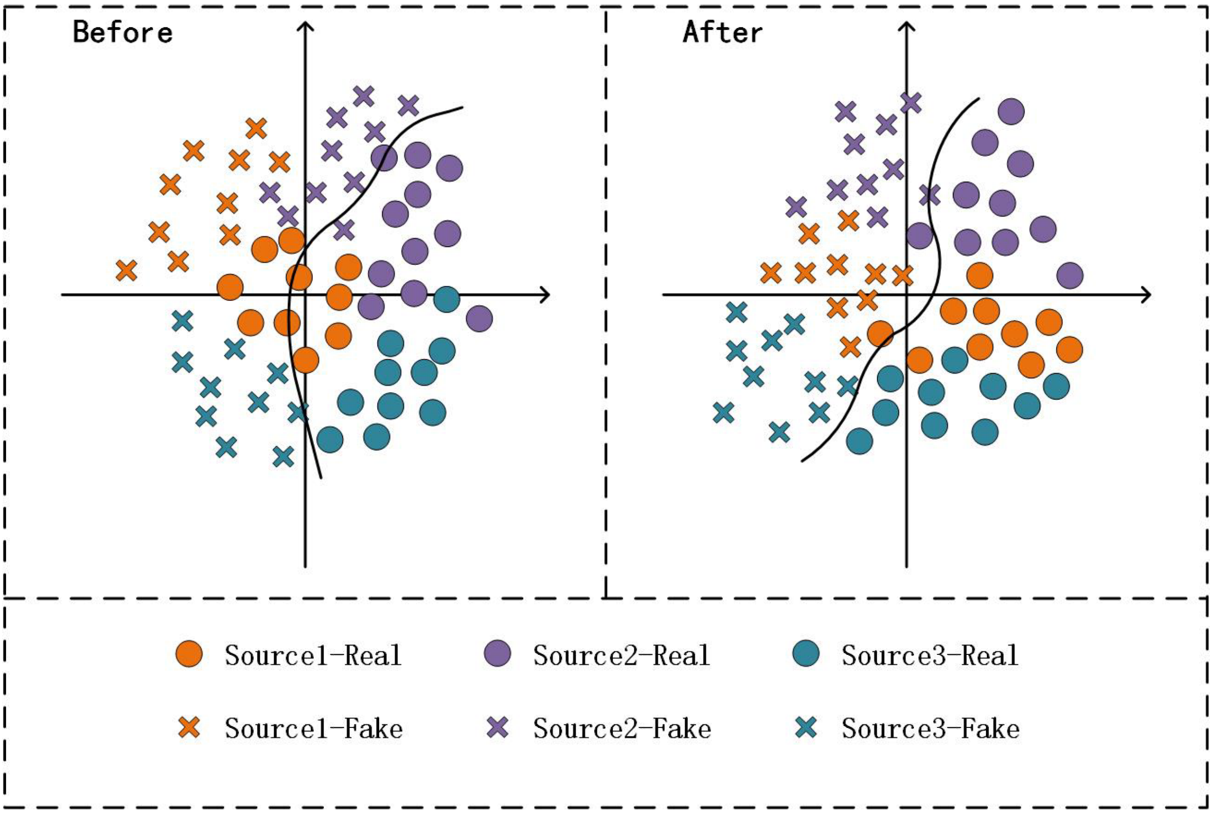

Figure 1. Schematic diagram of the multi-source domain alignment algorithm proposed in this method. The distribution plot on the left illustrates the sample distribution from the source domain before feature alignment, while the plot on the right demonstrates the distribution after feature alignment in the feature space. Notably, the original distributions from different domains exhibit distinct dissimilarities; however, the feature alignment enhances the uniformity of each domain’s distribution in the feature space.

2. Face Anti-Spoofing Methods

There exist two principal categories of conventional face anti-spoofing (FAS) techniques: appearance-based and temporal-based methods. Appearance-based methods involve the extraction of hand-crafted features for classification, such as local binary patterns (LBPs) [9][10] and scale-invariant feature transform (SIFT) [11]. However, temporal-based FAS methods detect attack faces by extracting temporal cues from a sequence of consecutive face frames. Mouth movement detection [12] and eye blink detection [13][14] are examples of the earliest dynamic texture-detection methods. However, these methods do not generalize well to cross-dataset testing scenarios due to the dissimilarities in feature spaces among diverse domains, which often lead to feature bias during generalization.

Recently, deep neural networks, specifically convolutional neural networks (CNNs), have gained widespread adoption in computer vision tasks and have been extensively applied to FAS. Yang et al. [15] were the pioneers in utilizing binary classification CNNs for FAS. Jourabloo et al. [16] proposed a face de-spoofing technique that performs fake face classification by reverse decomposition into real faces and spoof noise. Liu et al. [17] presented a CNN-RNN network that combines both appearance and temporal cues to detect spoof attacks using remote photoplethysmography (rPPG) signals. Similarly, 3D mask detection attack methods [18][19] also exploit rPPG information. Yang et al. [20] combined temporal and appearance cues to detect fake faces from real ones. Roy et al. [21] investigated frame-level FAS to enhance biometric authentication security against face-spoofing attacks. More recently, deep neural networks have been applied to FAS [4][7][8][22][23][24], achieving superior performance compared to conventional methods [15][25][26][27].

In conclusion, traditional FAS methods include appearance-based and temporal-based methods, which extract hand-crafted features and mine temporal cues, respectively. However, deep neural networks, especially CNNs, have achieved state-of-the-art performance in FAS by combining the appearance and temporal cues, using techniques such as reverse decomposition and frame-level FAS for improved biometric authentication security.

3. Multi-Domain Learning

The use of multiple datasets has recently sparked research interest in multi-domain processing. In particular, the research community has amassed several large-scale FAS datasets with rich annotations [28][29][30][31]. The work on multi-source domain processing shares similarities with domain adaptation (DA) methods [32][33][34][35][36][37][38][39] that require a retrained model to perform well on both source and target domain data. Specifically, Zhang et al. [37] introduced the concept of margin disparity discrepancy to characterize the differences between source and target domains, which has inspired the researchers' work. Ariza et al. [40] conducted a comparative study of several classification methods, which informed the researchers' experimental design. Additionally, Liu et al. [41] proposed the YOLOv3-FDL model for successful small crack detection from GPR images using a four-scale detection layer. Notably, the common approaches of Mancini et al. [34] and Rebuffi et al. [35] employ the ResNet [42] architecture, which offers benefits over architectures such as VGG [43] and AlexNet [44] by increasing abstraction through convolutional layers. Yang et al. [45] recently enriched FAS datasets from a different perspective to achieve multi-domain training, while Guo et al. [46] proposed a novel multi-domain model that overcomes the forgetting problem when learning new domain samples and exhibits high adaptability.

4. Domain Generalization

Domain adaptation (DA) and domain generalization (DG) are two fundamental methods used in FAS research. While the DG method mines the relationships among multiple domains, the DA method aims to adapt the model to a target domain. In this work, the researchers propose a novel domain generalization method that introduces a new loss function inspired by the work of Motiian et al. [47], which encourages feature extraction in similar classes. To align multiple source domains for generalization, previous works such as Ghifary et al. [48] and Li et al. [49] have proposed autoencoder-based approaches. The method follows a similar approach of learning a shared feature space across multiple source domains that can generalize to the target domain. Previous works such as Shao et al. [7], Saha et al. [50], Jia et al. [8], and Kim et al. [51] have also attempted to achieve this goal. Among them, the single-side adversarial learning method proposed in SSDG [8] is the work most related to ours. However, this end-to-end approach overlooks the relationships between different domains. To address the overfitting and generalization problems of adversarial generative networks, Li et al. [5] proposed a multi-channel convolutional neural network (MCCNN). Additionally, meta-learning formulations [52][53][54] have been utilized to simulate the domain shift during training to learn a representative feature space. However, recent works such as Wang et al. [6] have not adopted a domain-alignment approach but have instead increased the diversity of labeled data by reassembling different styles and content features in their SSAN method.

References

1. Deng, J.; Guo, J.; Xue, N.; Zafeiriou, S. Arcface: Additive angular margin loss for deep face recognition. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 4690–4699.
2. Wang, M.; Deng, W. Deep face recognition: A survey. Neurocomputing 2021, 429, 215–244.
3. Yu, Z.; Qin, Y.; Li, X.; Zhao, C.; Lei, Z.; Zhao, G. Deep learning for face anti-spoofing: A survey. IEEE Trans. Pattern Anal. Mach. Intell. 2022, 45, 5609–5631.
4. Benlamoudi, A.; Bekhouche, S.E.; Korichi, M.; Bensid, K.; Ouahabi, A.; Hadid, A.; Taleb-Ahmed, A. Face Presentation Attack Detection Using Deep Background Subtraction. Sensors 2022, 22, 3760.
5. Li, S.; Dutta, V.; He, X.; Matsumaru, T. Deep Learning Based One-Class Detection System for Fake Faces Generated by GAN Network. Sensors 2022, 22, 7767.
6. Wang, Z.; Wang, Z.; Yu, Z.; Deng, W.; Li, J.; Gao, T.; Wang, Z. Domain Generalization via Shuffled Style Assembly for Face Anti-Spoofing. In Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022; pp. 4123–4133.
7. Shao, R.; Lan, X.; Li, J.; Yuen, P.C. Multi-adversarial discriminative deep domain generalization for face presentation attack detection. In Proceedings of the 2019 IEEE/CVF Conference on

Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 10023–10031.

8. Jia, Y.; Zhang, J.; Shan, S.; Chen, X. Single-side domain generalization for face anti-spoofing. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020; pp. 8484–8493.
9. Määttä, J.; Hadid, A.; Pietikäinen, M. Face spoofing detection from single images using micro-texture analysis. In Proceedings of the 2011 International Joint Conference on Biometrics (IJCB), Washington, DC, USA, 11–13 October 2011; pp. 1–7.
10. Freitas Pereira, T.d.; Komulainen, J.; Anjos, A.; De Martino, J.M.; Hadid, A.; Pietikäinen, M.; Marcel, S. Face liveness detection using dynamic texture. EURASIP J. Image Video Process. 2014, 2014, 2.
11. Patel, K.; Han, H.; Jain, A.K. Secure face unlock: Spoof detection on smartphones. IEEE Trans. Inf. Forensics Secur. 2016, 11, 2268–2283.
12. Kollreider, K.; Fronthaler, H.; Faraj, M.I.; Bigun, J. Real-time face detection and motion analysis with application in “liveness” assessment. IEEE Trans. Inf. Forensics Secur. 2007, 2, 548–558.
13. Pan, G.; Sun, L.; Wu, Z.; Lao, S. Eyeblink-based anti-spoofing in face recognition from a generic webcamera. In Proceedings of the 2007 IEEE 11th International Conference on Computer Vision (ICCV), Rio De Janeiro, Brazil, 14–21 October 2007; pp. 1–8.
14. Sun, L.; Pan, G.; Wu, Z.; Lao, S. Blinking-based live face detection using conditional random fields. In Proceedings of the International Conference on Biometrics (ICB), Seoul, Republic of Korea, 27–29 August 2007; pp. 252–260.
15. Yang, J.; Lei, Z.; Li, S.Z. Learn convolutional neural network for face anti-spoofing. arXiv 2014, arXiv:1408.5601.
16. Jourabloo, A.; Liu, Y.; Liu, X. Face de-spoofing: Anti-spoofing via noise modeling. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 290–306.
17. Liu, Y.; Jourabloo, A.; Liu, X. Learning deep models for face anti-spoofing: Binary or auxiliary supervision. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–23 June 2018; pp. 389–398.
18. Liu, S.; Yuen, P.C.; Zhang, S.; Zhao, G. 3D mask face anti-spoofing with remote photoplethysmography. In Proceedings of the European Conference on Computer Vision (ECCV), Amsterdam, The Netherlands, 11–14 October 2016; pp. 85–100.
19. Liu, S.Q.; Lan, X.; Yuen, P.C. Remote photoplethysmography correspondence feature for 3D mask face presentation attack detection. In Proceedings of the European Conference on

Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 558–573.

20. Yang, X.; Luo, W.; Bao, L.; Gao, Y.; Gong, D.; Zheng, S.; Li, Z.; Liu, W. Face anti-spoofing: Model matters, so does data. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 3507–3516.
21. Roy, K.; Hasan, M.; Rupty, L.; Hossain, M.S.; Sengupta, S.; Taus, S.N.; Mohammed, N. Bi-fpnfas: Bi-directional feature pyramid network for pixel-wise face anti-spoofing by leveraging fourier spectra. Sensors 2021, 21, 2799.
22. Liu, Y.; Stehouwer, J.; Jourabloo, A.; Liu, X. Deep tree learning for zero-shot face anti-spoofing. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 4680–4689.
23. Liu, Y.; Stehouwer, J.; Liu, X. On disentangling spoof trace for generic face anti-spoofing. In Proceedings of the European Conference on Computer Vision (ECCV), Glasgow, UK, 23–28 August 2020; pp. 406–422.
24. Zhang, K.Y.; Yao, T.; Zhang, J.; Tai, Y.; Ding, S.; Li, J.; Huang, F.; Song, H.; Ma, L. Face anti-spoofing via disentangled representation learning. In Proceedings of the European Conference on Computer Vision (ECCV), Glasgow, UK, 23–28 August 2020; pp. 641–657.
25. Feng, L.; Po, L.M.; Li, Y.; Xu, X.; Yuan, F.; Cheung, T.C.H.; Cheung, K.W. Integration of image quality and motion cues for face anti-spoofing: A neural network approach. J. Vis. Commun. Image Represent. 2016, 38, 451–460.
26. Li, L.; Feng, X.; Boulkenafet, Z.; Xia, Z.; Li, M.; Hadid, A. An original face anti-spoofing approach using partial convolutional neural network. In Proceedings of the 2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA), Oulu, Finland, 12–15 December 2016; pp. 1–6.
27. Patel, K.; Han, H.; Jain, A.K. Cross-database face antispoofing with robust feature representation. In Proceedings of the Biometric Recognition: 11th Chinese Conference (CCBR), Chengdu, China, 14–16 October 2016; pp. 611–619.
28. Chingovska, I.; Anjos, A.; Marcel, S. On the effectiveness of local binary patterns in face anti-spoofing. In Proceedings of the 2012 BIOSIG-Proceedings of the International Conference of Biometrics Special Interest Group (BIOSIG), Darmstadt, Germany, 6–7 September 2012; pp. 1–7.
29. Zhang, Z.; Yan, J.; Liu, S.; Lei, Z.; Yi, D.; Li, S.Z. A face antispoofing database with diverse attacks. In Proceedings of the 2012 5th IAPR International Conference on Biometrics (ICB), New Delhi, India, 29 March–1 April 2012; pp. 26–31.
30. Wen, D.; Han, H.; Jain, A.K. Face spoof detection with image distortion analysis. IEEE Trans. Inf. Forensics Secur. 2015, 10, 746–761.

31. Boulkenafet, Z.; Komulainen, J.; Li, L.; Feng, X.; Hadid, A. OULU-NPU: A mobile face presentation attack database with real-world variations. In Proceedings of the 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017), Washington, DC, USA, 30 May–3 June 2017; pp. 612–618.

32. Rebuffi, S.A.; Bilen, H.; Vedaldi, A. Learning multiple visual domains with residual adapters. arXiv 2017, arXiv:1705.08045.

33. Li, H.; Li, W.; Cao, H.; Wang, S.; Huang, F.; Kot, A.C. Unsupervised domain adaptation for face anti-spoofing. IEEE Trans. Inf. Forensics Secur. 2018, 13, 1794–1809.

34. Mancini, M.; Ricci, E.; Caputo, B.; Rota Bulo, S. Adding new tasks to a single network with weight transformations using binary masks. In Proceedings of the European Conference on Computer Vision (ECCV) Workshops, Munich, Germany, 8–14 September 2018.

35. Rebuffi, S.A.; Bilen, H.; Vedaldi, A. Efficient parametrization of multi-domain deep neural networks. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–23 June 2018; pp. 8119–8127.

36. Tu, X.; Zhang, H.; Xie, M.; Luo, Y.; Zhang, Y.; Ma, Z. Deep transfer across domains for face antispoofing. J. Electron. Imaging 2019, 28, 043001.

37. Zhang, Y.; Liu, T.; Long, M.; Jordan, M. Bridging theory and algorithm for domain adaptation. In Proceedings of the International Conference on Machine Learning (ICML), Long Beach, CA, USA, 9–15 June 2019; pp. 7404–7413.

38. Wang, G.; Han, H.; Shan, S.; Chen, X. Improving cross-database face presentation attack detection via adversarial domain adaptation. In Proceedings of the 2019 International Conference on Biometrics (ICB), Crete, Greece, 4–7 June 2019; pp. 1–8.

39. Wang, G.; Han, H.; Shan, S.; Chen, X. Unsupervised adversarial domain adaptation for cross-domain face presentation attack detection. IEEE Trans. Inf. Forensics Secur. 2020, 16, 56–69.

40. Ariza-Colpas, P.; Piñeres-Melo, M.; Barceló-Martínez, E.; De la Hoz-Franco, E.; Benitez-Agudelo, J.; Gelves-Ospina, M.; Echeverri-Ocampo, I.; Combita-Nino, H.; Leon-Jacobus, A. Enkephalon-technological platform to support the diagnosis of alzheimer’s disease through the analysis of resonance images using data mining techniques. In Proceedings of the Advances in Swarm Intelligence: 10th International Conference, ICSI 2019, Chiang Mai, Thailand, 26–30 July 2019; pp. 211–220.

41. Liu, Z.; Gu, X.; Chen, J.; Wang, D.; Chen, Y.; Wang, L. Automatic recognition of pavement cracks from combined GPR B-scan and C-scan images using multiscale feature fusion deep neural networks. Autom. Constr. 2023, 146, 104698.

42. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Las

Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

43. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv 2014, arXiv:1409.1556.
44. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. Commun. ACM 2017, 60, 84–90.
45. Yang, J.; Lan, G.; Xiao, S.; Li, Y.; Wen, J.; Zhu, Y. Enriching facial anti-spoofing datasets via an effective face swapping framework. Sensors 2022, 22, 4697.
46. Guo, X.; Liu, Y.; Jain, A.; Liu, X. Multi-domain Learning for Updating Face Anti-spoofing Models. In Proceedings of the European Conference on Computer Vision (ECCV), Tel Aviv, Israel, 23–27 October 2022; pp. 230–249.
47. Motiian, S.; Piccirilli, M.; Adjeroh, D.A.; Doretto, G. Unified deep supervised domain adaptation and generalization. In Proceedings of the 2017 IEEE/CVF International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 5715–5725.
48. Ghifary, M.; Kleijn, W.B.; Zhang, M.; Balduzzi, D. Domain generalization for object recognition with multi-task autoencoders. In Proceedings of the 2015 IEEE/CVF International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; pp. 2551–2559.
49. Li, H.; Pan, S.J.; Wang, S.; Kot, A.C. Domain generalization with adversarial feature learning. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–23 June 2018; pp. 5400–5409.
50. Saha, S.; Xu, W.; Kanakis, M.; Georgoulis, S.; Chen, Y.; Paudel, D.P.; Van Gool, L. Domain agnostic feature learning for image and video based face anti-spoofing. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, Seattle, WA, USA, 13–19 June 2020; pp. 802–803.
51. Kim, T.; Kim, Y. Suppressing spoof-irrelevant factors for domain-agnostic face anti-spoofing. IEEE Access 2021, 9, 86966–86974.
52. Shao, R.; Lan, X.; Yuen, P.C. Regularized fine-grained meta face anti-spoofing. In Proceedings of the AAAI Conference on Artificial Intelligence (AAAI), New York, NY, USA, 7–12 February 2020; Volume 34, pp. 11974–11981.
53. Chen, Z.; Yao, T.; Sheng, K.; Ding, S.; Tai, Y.; Li, J.; Huang, F.; Jin, X. Generalizable representation learning for mixture domain face anti-spoofing. In Proceedings of the AAAI Conference on Artificial Intelligence (AAAI), Vancouver, BC, Canada, 2–9 February 2021; Volume 35, pp. 1132–1139.
54. Wang, J.; Zhang, J.; Bian, Y.; Cai, Y.; Wang, C.; Pu, S. Self-domain adaptation for face anti-spoofing. In Proceedings of the AAAI Conference on Artificial Intelligence (AAAI), Vancouver, BC,

Canada, 2–9 February 2021; Volume 35, pp. 2746–2754.