# Special Issue on Face Presentation Attack Detection

FACE presentation attack detection, also termed Face Anti-Spoofing (FAS) [item 1), 2) in the Appendix), is a hot and challenging research topic that has received much attention from the computer vision and pattern recognition communities in the past. Owing to the development of deep learning and big data, recent advances in this and related fields has increased considerably. However, there are still several challenging tasks that deserve attention from the community, for instance robust techniques to unknown spoofing attacks, cross-domain generalization, and multi-modal fusion in images and video sequences. We edited this special issue with the goal of compiling the latest progress in the field and identifying promising research opportunities on FAS.

The guest editors invited paper submissions for the special issue on face anti-spoofing to be published in IEEE TRANSACTIONS ON BIOMETRICS, BEHAVIOR, AND IDENTITY SCIENCE (TBIOM). The call for submissions was open from December $18^{th}$, 2020 to May $20^{st}$, 2021. A dozen submissions were received, and every submission was subject to the standard TBIOM rigorous peer-review process. This led to the set of four papers appearing in this special issue. In the remainder of this note, we briefly summarize the contributions of the articles included in the issue.

The first article of this issue, "Revisiting Pixel-Wise Supervision for Face Anti-Spoofing" [item 3) in the Appendix), first provides a comprehensive review and analysis of existing pixel-wise supervision methods for FAS. Then, it proposes a novel pyramid supervision to learn local and global semantics from multi-scale spatial contexts. The proposed pyramid supervision is able to not only improve the FAS performances beyond the existing pixel-wise supervision frameworks but also enhances the model's interpretability. Moreover, this paper also studies the efficacy of different architecture configurations with two kinds of pixel-wise supervision, such as binary mask and depth map supervisions.

The second article of this issue, "Attention-Based Spatial-Temporal Multi-Scale Network for Face Anti-Spoofing" [item 4) in the Appendix), proposes a two-stream spatial-temporal network to explore the potential depth information and multi-scale information, respectively. It first introduces a temporal shift module to extract temporal information and a scale-level attention module based on the estimated depth map to extract essential discriminative features. Finally, a fully-connected network is used to judge the face as fake or real. Extensive experiments show the performance of proposed methods on five face anti-spoofing datasets.

The third article of this issue, "Inconsistency-Aware Wavelet Dual-Branch Network for Face Forgery Detection" [item 5) in the Appendix), presents an inconsistency-aware wavelet dual-branch network for face forgery detection. This model utilizes two kinds of forgery clues: inter-image and intra-image inconsistencies. The authors firstly enhance the forgery features by using additional inputs based on stationary wavelet decomposition (SWD) and then design a dual-branch network to predict image-level and pixel-level forgery labels, respectively. The segmentation branch aims to recognize real and fake local regions, which is crucial for discovering intra-image inconsistency. The classification branch learns to discriminate the real and fake images globally, thus it can extract inter-image inconsistency. Finally, bilinear pooling is employed to fuse the features from the two branches. Experiments show the proposed method surpasses the state-of-the-art face forgery detection methods.

The fourth article of this issue, "Fighting Fake News: Two Stream Network for Deepfake Detection via Learnable SRM" [item 6) in the Appendix), first analyzes the drawbacks of Deepfake detection (rarely considering temporal information, rapidly decrease of low-quality data). To address these problems, the authors propose a two-stream network to detect Deepfake at the video level with the capability of handling low-quality data. The proposed architecture firstly divides the input video into segments and then feeds selected frames of each segment into two streams. The first stream takes RGB information as input and tries to learn the semantic inconsistency. In parallel, the second stream leverages noise features extracted by spatial rich model (learnable SRM) filters. The proposed method is evaluated on the largest Deepfake dataset up to date, FaceForensics++, and the experimental results show that the proposed model obtains state-of-the-art performance.

We thank the authors for submitting their work to TBIOM, and we thank the reviewers for their comments and suggestions during the review process. We hope the readers will find useful the insights in these papers, all of them at the cutting edge of face anti-spoofing.

JUN WAN
National Laboratory of Pattern Recognition
Institute of Automation
Chinese Academy of Sciences
Beijing 100190, China
School of Artificial Intelligence
University of Chinese Academy of Sciences
Beijing 100049, China

Digital Object Identifier 10.1109/TBIOM.2021.3089903

SERGIO ESCALERA
Computer Vision Center
Universitat de Barcelona
08007 Barcelona, Spain

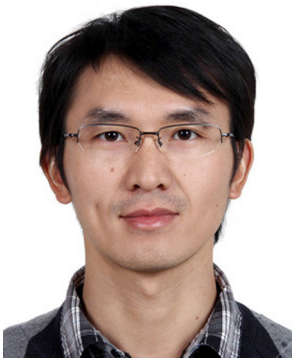
HUGO JAIR ESCALANTE
Instituto Nacional de Astrofísica, Óptica y Electrónica
Puebla 72840, Mexico

GUODONG GUO
Institute of Deep Learning
Baidu Research
Beijing 100193, China
National Engineering Laboratory for Deep Learning
Technology and Application
Beijing 100193, China

STAN Z. LI
Center for AI Research and Innovation
Westlake University
Hangzhou 310024, China
Faculty of Science and Technology
Macau University of Science and Technology
Macau, China

## APPENDIX
## RELATED WORK

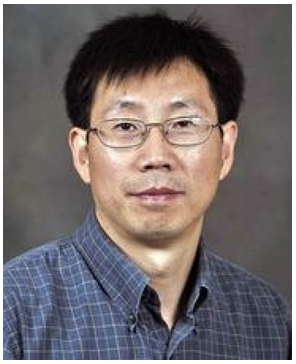
1) Z. Yu, J. Wan, Y. Qin, X. Li, S. Z. Li, and G. Zhao, "NAS-FAS: Static-dynamic central difference network search for face anti-spoofing," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Nov. 9, 2020, doi: 10.1109/TPAMI.2020.3036338.
2) A. Liu *et al.*, "Cross-ethnicity face anti-spoofing recognition challenge: A review," *IET Biometr.*, vol. 10, no. 1, pp. 24–43, 2021.
3) Z. Yu, X. Li, J. Shi, Z. Xia, and G. Zhao, "Revisiting pixel-wise supervision for face anti-spoofing," *IEEE Trans. Biom., Behav., Ident. Sci.*, early access, Mar. 11, 2021, doi: 10.1109/TBIOM.2021.3065526.
4) W. Zheng, M. Yue, S. Zhao, and S. Liu, "Attention-based spatial-temporal multi-scale network for face anti-spoofing," *IEEE Trans. Biom., Behav., Ident. Sci.*, early access, Mar. 22, 2021, doi: 10.1109/TBIOM.2021.3066983.
5) G. Jia *et al.*, "Inconsistency-aware wavelet dual-branch network for face forgery detection," *IEEE Trans. Biom., Behav., Ident. Sci.*, early access, Jun. 7, 2021, doi: 10.1109/TBIOM.2021.3086109.
6) B. Han, X. Han, H. Zhang, J. Li, and X. Cao, "Fighting fake news: Two stream network for deepfake detection via learnable SRM," *IEEE Trans. Biom., Behav., Ident. Sci.*, early access, Mar. 12, 2021, doi: 10.1109/TBIOM.2021.3065735.



**Jun Wan** (Senior Member, IEEE) received the B.S. degree from the China University of Geosciences, Beijing, China, in 2008, and the Ph.D. degree from the Institute of Information Science, Beijing Jiaotong University, Beijing, in 2015. Since January 2015, he has been a Faculty Member with the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Science, China, where he currently serves as an Associate Professor. His main research interests include computer vision and machine learning. He is an Associate Editor of the *IET Biometrics* from 2020 to 2022 and an Area Chair for ICME 2021, and has served as the Co-Editor of special issues in IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, IEEE TRANSACTIONS ON BIOMETRICS, BEHAVIOR, AND IDENTITY SCIENCE, *Machine Vision and Applications*, and *Entropy*.

**Sergio Escalera** received the Ph.D. degree in multiclass visual categorization systems from the Computer Vision Center, UAB. He leads the Human Pose Recovery and Behavior Analysis Group. He is a Full Professor with the Department of Mathematics and Informatics, Universitat de Barcelona. He is a Distinguished Professor with Aalborg University. He is also a member of the Computer Vision Center, UAB. He is a Series Editor of The Springer Series on Challenges in Machine Learning. He is the Vice-President of ChaLearn Challenges in Machine Learning, leading ChaLearn Looking at People events. He is the Co-Creator of Codalab open-source platform for challenges organization. He is also a Fellow of the ELLIS European Laboratory for Learning and Intelligent Systems working within the Human-Centric Machine Learning Program, and the Chair of IAPR TC-12: Multimedia and Visual Information Systems. His research interests include automatic analysis of humans from visual and multimodal data, with special interest in inclusive, transparent, and fair affective computing and people characterization. He obtained the 2008 Best Thesis Award. He received a CVPR Best Paper Award Nominee and a CVPR Outstanding Reviewer Award.

**Hugo Jair Escalante** received the Ph.D. degree in computer science. He has been a Research Scientist with the Instituto Nacional de Astrofisica, Optica y Electronica, INAOE, Mexico, since 2012. He is a Secretary and a member of the board of directors of ChaLearn, The Challenges in Machine Learning Organization, a nonprofit organism dedicated to organizing challenges, since 2011. He is the Chair of the IAPR Technical Committee 12. He has served as the Competition Chair of NeurIPS2019, NeurIPS2020, ICPR2020, and FG2020.



**Guodong Guo** (Senior Member, IEEE) received the B.E. degree in automation from Tsinghua University, Beijing, China, the Ph.D. degree in pattern recognition and intelligent control from the Chinese Academy of Sciences, Beijing, and the Ph.D. degree in computer science from the University of Wisconsin–Madison, Madison, WI, USA. He authored a book, *Face, Expression, and Iris Recognition Using Learning-Based Approaches* (2008), co-edited two books, *Support Vector Machines Applications* (2014) and *Mobile Biometrics* (2017), and published above 100 technical papers. His research interests include computer vision, biometrics, machine learning, and multimedia.

**Stan Z. Li** (Fellow, IEEE) received the B.Eng. degree from Hunan University, China, the M.Eng. degree from the National University of Defense Technology, China, and the Ph.D. degree from Surrey University, U.K. He was an Associate Professor with Nanyang Technological University, Singapore. He was a Professor and the Director of the Center for Biometrics and Security Research, Institute of Automation, Chinese Academy of Sciences, and was with Microsoft Research Asia, as a Researcher from 2000 to 2004. He is currently a Chair Professor with Westlake University. His research interests include pattern recognition and machine learning, image and vision processing, face recognition, biometrics, and intelligent video surveillance. He has published more than 200 papers in international journals and conferences, and authored and edited eight books. He was an Associate Editor of the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE and is acting as the Editor-in-Chief for the *Encyclopedia of Biometrics*. He was elevated to IEEE fellow for his contributions to the fields of face recognition, pattern recognition, and computer vision.