ADGD'21: 1st Workshop on Synthetic Multimedia — Audiovisual Deepfake Generation and Detection

Stefan Winkler
dprsw@nus.edu.sg
National University of Singapore
Singapore

Weiling Chen
weiling@aisingapore.org
National University of Singapore
Singapore

Abhinav Dhall
abhinav.dhall@monash.edu
Monash University and IIT Ropar
Australia and India

Pavel Korshunov
pavel.korshunov@idiap.ch
Idiap Research Institute
Switzerland

ABSTRACT

Deepfakes, i.e. synthetic or "fake" media content generated using deep learning, are a double-edged sword. On one hand, they pose new threats and risks in the form of scams, fraud, disinformation, social manipulation, or celebrity porn. On the other hand, deepfakes have just as many meaningful and beneficial applications – they allow us to create and experience things that no longer exist, or that have never existed, enabling numerous exciting applications in entertainment, education, and even privacy.

While most work has focused on fake images and video alone, the multi-modal, audiovisual aspect is very important to both convincing generation and accurate detection of fake multimedia content. Therefore, we organize ADGD21: 1st Workshop on Synthetic Multimedia – Audiovisual Deepfake Generation and Detection so as to provide a platform for researchers and engineers to share their ideas and approaches in this field.

Related Workshop Proceedings are available in the ACM DL at: http://dl.acm.org/citation.cfm?id=3476099

CCS CONCEPTS

• **Applied computing → Computer forensics**.

ACM Reference Format:
Stefan Winkler, Weiling Chen, Abhinav Dhall, and Pavel Korshunov. 2021. ADGD'21: 1st Workshop on Synthetic Multimedia — Audiovisual Deepfake Generation and Detection. In *Proceedings of the 29th ACM Int'l Conference on Multimedia (MM '21), Oct. 20–24, 2021, Virtual Event, China.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3474085.3478578

1 INTRODUCTION

Deepfakes, that is, synthetic or "fake" media content generated using deep learning, are a double-edged sword. On one hand, they pose new threats and risks in the form of scams, fraud, disinformation, social manipulation, or celebrity porn. On the other hand, deepfakes have just as many meaningful and beneficial applications – they allow us to create and experience things that no longer exist, or that have never existed, enabling numerous exciting applications in entertainment, education, and even privacy.

As a result, methods for the generation of deepfakes as well as the detection thereof are of great interest to the community. While most work so far has focused on fake images and video alone, the multi-modal, audio-visual aspect is very important to both convincing generation and accurate detection of fake multimedia content.

ACM Multimedia has been the premier conference and a key global event to display scientific achievements and innovative solutions in the multimedia field. Fake media generation and detection as considered in the scope of this workshop includes the process and analysis of both videos and audios thus is closely related to the focus of the conference. Therefore, we decide to host 1st Workshop on Synthetic Multimedia – Audiovisual Deepfake Generation and Detection in conjunction with the 29th ACM International Conference on Multimedia.

2 OBJECTIVE AND TOPICS OF INTEREST

The purpose of this workshop is to provide a platform for researchers and engineers to share their ideas and approaches, and give some insights on fake media generation and detection to both academia and industry.

We invite submissions on a range of AI technologies and applications for media forensics domains. Topics of interest include but are not limited to the following:

- Fake image generation and/or detection
- Fake voice generation and/or detection
- Audiovisual Deepfakes and adversarial attacks
- Audiovisual Deepfakes and Fairness and Ethics
- Audiovisual Deepfakes and Data augmentation
- Audiovisual Deepfake datasets

3 WORKSHOP ORGANIZERS

3.1 Organizing Committee

Stefan Winkler is Deputy Director at AI Singapore and Associate Professor (Practice) at the National University of Singapore. Prior to that he was Distinguished Scientist and Director of the Video & Analytics Program at the University of Illinois' Advanced Digital Sciences Center (ADSC) in Singapore. He also co-founded two start-ups and worked for a Silicon Valley company. Dr. Winkler has a

MM '21, October 20–24, 2021, Virtual Event, China.

ACM ISBN 978-1-4503-8651-7/21/10.
https://doi.org/10.1145/3474085.3478578

Ph.D. degree from the Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland, and a Dipl.-Ing. (M.Eng./B.Eng.) degree from the University of Technology Vienna, Austria. He is an IEEE Fellow and has published over 150 papers. He has also contributed to international standards in VQEG, ITU, ATIS, VSF, and SCTE.

Weiling Chen is a Senior AI Engineer at AI Singapore and National University of Singapore. Prior to that she was Data Scientist at Lazada Group. She received her Ph.D. degree from Nanyang Technological University and B.Eng. in Computer Science and Technology from Shandong University. She has won Champion for SeeTrue Workshop organized by Defence Science and Technology Agency (DSTA).

Abhinav Dhall is a lecturer and co-director of the Human-Centred Artificial Intelligence lab at Monash University. He is also an Assistant Professor (on leave) at the Indian Institute of Technology Ropar. He received his Ph.D. degree from the Australian National University. His research has received awards at ACM ICMR, IEEE FG and IEEE ICME.

Pavel Korshunov is a research associate at the Idiap Research Institute (Martigny, Switzerland). He currently works on detection of deepfakes and audio-visual manipulations, age detection in images and voice, and speech anti-spoofing. He received a Ph.D. in Computer Science from School of Computing, National University of Singapore in 2011 and was a postdoctoral researcher at EPFL (Lausanne, Switzerland) and Idiap Research Institute. During his past tenures, he worked on problems related to high dynamic range(HDR) imaging, crowdsourcing, visual privacy in video surveillance. He has over 70 papers with one ACM TOMM journal best paper award(2011), two top 10% best paper awards in MMSP 2014, and a top 10% best paper awards in ICIP 2014. He is also a co-editor of JPEG XT standard for HDR images.

3.2 Program Committee

We would like to thank the members of the program committee for their time and effort in providing helpful reviews:

Sunpreet Arora, Visa, United States
Jyoti Joshi, Kroop AI, India
Siwei Lyu, SUNY Buffalo
Yisroel Mirsky, Georgia Institute of Technology, United States
Ramanathan Subramanian, University of Canberra, Australia
Ruben Tolosana, Autonomous University of Madrid, Spain
Luisa Verdoliva, University Federico II of Naples, Italy

4 WORKSHOP PROGRAM

4.1 Invited Keynote Speakers

Siwei Lyu is an Empire Innovation Professor at the Department of Computer Science and Engineering and the founding Director of UB Media Forensic Lab (UB MDFL) of the University at Buffalo, State University of New York. Before joining UB, Dr. Lyu was an Assistant Professor from 2008 to 2014, a tenured Associate Professor from 2014 to 2019, and a Full Professor from 2019 to 2020, at the Department of Computer Science, University at Albany, State University of New York. From 2005 to 2008, he was a Post-Doctoral Research Associate at the Howard Hughes Medical Institute and the Center for Neural Science of New York University. He was an Assistant Researcher at Microsoft Research Asia (then Microsoft Research China) in 2001. Dr. Lyu received his Ph.D. degree in Computer Science from Dartmouth College in 2005, and his M.S. degree in Computer Science in 2000, and B.S. degree in Information Science in 1997, both from Peking University, China. Dr. Lyu's research interests include digital media forensics, computer vision, and machine learning. Dr. Lyu has published over 150 refereed journal and conference papers. He is the recipient of the IEEE Signal Processing Society Best Paper Award (2011), the National Science Foundation CAREER Award (2010), SUNY Albany's Presidential Award for Excellence in Research and Creative Activities (2017), SUNY Chancellor's Award for Excellence in Research and Creative Activities (2018) and Google Faculty Research Award (2019).

Sergey Tulyakov is a Principal Research Scientist heading the Creative Vision team at Snap Research. His work focuses on creating methods for manipulating the world via computer vision and machine learning. This includes human and object understanding, photorealistic manipulation and animation, video synthesis, prediction and retargeting. He pioneered the unsupervised image animation domain with MonkeyNet and First Order Motion Model that sparkled a number of startups in the domain. His work on Interactive Video Stylization received the Best in Show Award at SIGGRAPH Real-Time Live! 2020. He has published 30+ top conference papers, journals and patents resulting in multiple innovative products, including Snapchat Pet Tracking, OurBaby Snappable, Real-time Neural Lenses (gender swap, baby face, aging lens, face animation) and many others. Before joining Snap Inc., Sergey was with Carnegie Mellon University, Microsoft, NVIDIA. He holds a PhD degree from the University of Trento, Italy.

Jianfei Cai is a Professor at Faculty of IT, Monash University, where he currently serves as the Head for the Data Science & AI Department. Before that, he was a full professor, a cluster deputy director of Data Science & AI Research center (DSAIR), Head of Visual and Interactive Computing Division and Head of Computer Communications Division in Nanyang Technological University (NTU). His major research interests include computer vision, multimedia and visual computing. He has published more than 200 technical papers in international conferences and journals. He is a co-recipient of paper awards in ACCV, ICCM, IEEE ICIP and MMSP. He has served as an Associate Editor for IEEE T-IP, T-MM, T-CSVT and Visual Computer as well as serving as Area Chair for CVPR, ICCV, ECCV, ACM Multimedia, IJCAI, ICME and ICIP. He was the Chair of IEEE CAS VSPC-TC during 2016-2018. He had also served as the leading TPC Chair for IEEE ICME 2012 and the best paper award committee chair/co-chair for IEEE T-MM 2020/2019. He is a Fellow of IEEE.

4.2 Accepted Papers

All accepted papers received at least two reviews from the reviewers. The program includes 5 technical papers distributed over two sessions, namely deepfake detection and deepfake generation.

ACKNOWLEDGMENTS

This work is supported by the National Research Foundation, Singapore under its AI Singapore Programme (AISG Award No: AISG-GC-2019-001).