ABSTRACT
Object re-identification (or object re-ID) has gained significant attention in recent years, fueled by the increasing demand for advanced video analysis and safety systems. In object re-identification, a query can be of different modalities, such as an image, a video, or natural language, containing or describing the object of interest. This workshop aims to bring together researchers, practitioners, and enthusiasts interested in object re-identification to delve into the latest advancements, challenges, and opportunities in this dynamic field. The workshop covers a spectrum of topics related to object re-identification, including but not limited to deep metric learning, multi-view data generation, video-based object re-identification, cross-domain object re-identification and real-world applications. The workshop provides a platform for researchers to showcase their work, exchange ideas, and foster potential collaborations. Additionally, it serves as a valuable opportunity for practitioners to stay abreast of the latest developments in object re-identification technology. Overall, this workshop creates a unique space to explore the rapidly evolving field of object re-identification and its profound impact on advancing the capabilities of multimedia analysis and retrieval.

CCS CONCEPTS
• Computing methodologies → Appearance and texture representations; Visual content-based indexing and retrieval.

KEYWORDS
Multimedia Retrieval, Object Re-identification, Representation Learning, Deep Metric Learning, Multi-view Generation

1 BACKGROUND AND MOTIVATION
Object re-identification [18], commonly known as object re-id, has become a focal point of research and development due to the escalating demand for sophisticated video analysis and safety systems. In our daily life, an object (e.g., a vehicle, person, or fashion item) is encountered in a certain scenario, and re-identifying the same object means recognising it in an another context of a different time, location or viewpoint. Here, we usually refer to object instances instead of object categories. In the computer vision community, commonly-studied objects include persons, animals, vehicles, and landmarks, which are captured by cameras. Specifically, given an image of an object instance (or query) captured by a certain camera, we aim to match it against a database of previously captured object images to find those containing the same instance. In object re-identification, a query can be of different modalities, such as an image [5, 14, 19], a video [1, 13], point cloud [16, 20] or natural language [3, 7, 8, 12], containing or describing the object of interest. As one of the fundamental research problems, object re-identification can be applied to many real-world applications, including recognizing persons and vehicles for public safety [4, 10, 15], tracking multiple objects in a
We plan to hold a hybrid format of workshop, i.e., both onsite and online. For the onsite one at least two organizers will attend in person to host the workshop. The workshop will include two major activities, the invited keynotes, and the paper presentations. We will invite keynote presentations for a half-day workshop, following by activities, the invited keynotes, and the paper presentations. We will invite the following experts as the (senior) potential program committee (PC) members to organize the reviewing process:

(1) Dylan Campbell (Australian National University, Australia), (2) Tawfiq Salem (Purdue University, USA), (3) Julian F.P.Kooij (Delft University, Netherlands), (4) Long Chen (Hong Kong University of Science and Technology, China), (5) Yawei Luo (Zhejiang University, China), (6) Torsten Sattler (Czech Technical University in Prague, Czech Republic), (7) Laurent Kneip (ShanghaiTech University, China), (8) Yan Yan (Washington State University, USA), (9) Paul-Edouard Sarlin (ETH Zurich, Switzerland)

### 3.1 Submission Types
In this workshop, we welcome two types of submissions, all of which should relate to the topics and themes as listed in Section 2:

1. Original papers (up to 4 pages in length, plus unlimited pages for references): original solution to the tasks in the scope of workshop topics and themes.
2. Position or perspective papers (up to 4 pages in length, plus unlimited pages for references): original ideas, perspectives, research vision, and open challenges in the area of evaluation approaches for explainable multimedia systems;
3. Survey papers (up to 4 pages in length, plus unlimited pages for references): papers summarizing existing publications in leading conferences and high-impact journals that are relevant for the topic of the workshop; Page limits include diagrams and appendices. Submissions should be single-blind, written in English, and formatted according to the current ACM two-column conference format. Suitable LaTeX, Word, and Overleaf templates are available from the ACM Website (use “sigconf” proceedings template for LaTeX and the Interim Template for Word).

### 3.2 Potential Program Committee Members
We will invite the following experts as the (senior) potential program committee (PC) members to organize the reviewing process:

(1) Dylan Campbell (Australian National University, Australia), (2) Tawfiq Salem (Purdue University, USA), (3) Julian F.P.Kooij (Delft University, Netherlands), (4) Long Chen (Hong Kong University of Science and Technology, China), (5) Yawei Luo (Zhejiang University, China), (6) Torsten Sattler (Czech Technical University in Prague, Czech Republic), (7) Laurent Kneip (ShanghaiTech University, China), (8) Yan Yan (Washington State University, USA), (9) Paul-Edouard Sarlin (ETH Zurich, Switzerland)

### 4 ORGANIZER INFORMATION
Zhedong Zheng (https://zdzheng.xyz) is a tenure-track Assistant Professor at University of Macau. Previously he was a research fellow at School of Computing, National University of Singapore. He received the Ph.D. degree from the University of Technology Sydney, Australia, in 2021 and the B.S. degree from Fudan University, China, in 2016. He received the IEEE Circuits and Systems Society Outstanding Young Author Award of 2021. He has served as the reviewer and program committee (PC) member for multiple conferences and journals, including TPAMI, TMM, IJCV, CVPR, ICCV, ECCV, IJCAI, AAAI and ACM Multimedia, and organized a
special session on reliable retrieval at ICME 2022. Besides, he is also invited as a workshop keynote speaker at CVPR 2020, and 2021, and a tutorial speaker at ACM Multimedia 2022.

Yaxiong Wang [https://dlb.org/pid/202/3251.html] is an Associate Professor in Hefei University of Technology (HFUT), China. Previously, he was a research expert (PI) in Zhejiang Lab and a ‘ZJU100’ researcher in Zhejiang University (double employed). He obtained his Ph.D degree and B.S degree from Xi’an Jiaotong University and Lanzhou University in 2021 and 2015, respectively. He has been the reviewer for multiple international journals and conferences, including IEEE TIP, TNNLS, TMM, TCSVT, ACM TOMM, ICCV, ICME, ICASSP etc. His research interests include text-driven person-reid, cross-modal modeling, and generation model.

Xuelin Qian [https://naiq.github.io/] is a Postdoctoral Researcher from Tsinghua University. Before that, he received the Ph.D degree from Fudan University in 2021, and the B.S. degree from Xidian University in 2015. He has published over 10 papers in top-tier conferences and journals, and been served as a reviewer for CVPR, ICCV, TPAMI, IJCV etc. His research interests are person re-identification, image generation and 3D reconstruction.

Zhun Zhong [https://zhunzhong.site] is an Assistant Professor at the University of Nottingham. His research interests lie in computer vision and artificial intelligence. He has published over top-tier 30 papers, including CVPR, ECCV, ICCV, NeurIPS, TPAMI. He was recognized as an outstanding reviewer at CVPR 2020 and NeurIPS 2021 and is serving as an area chair for CVPR 2024, ECCV 2024, ICLR 2024, ACM MM 2024, as well as a guest editor for IJCV.

Zheng Wang [https://wangzwhu.github.io/home/] is a Professor with the National Engineering Research Center for Multimedia Software, Wuhan University, China. He has co-organized several workshops or tutorials in top-tier conferences, including CVPR 2023, ACM MM 2022, ACM MM 2020, CVPR 2020, ACM ICMR 2020 etc. His research interests focus on multimedia content analysis and retrieval.

Liang Zheng [https://zheng-lab.cecs.anu.edu.au/] is an Associate Professor (with Tenure) in the School of Computing, Australian National University (ANU). He joined ANU in 2018 and held the CS Futures Fellowship and ARC DECRA Fellowship. He received my Ph.D. (EE) from Tsinghua University in 2015, and his B.S. (Biology) from Tsinghua University in 2010. He was named Top-40 Early Achievers by The Australian.

REFERENCES


