

MORE '24 Multimedia Object Re-ID: Advancements, Challenges, and Opportunities

Zhedong Zheng
University of Macau
Macau, China
zdzheng12@gmail.com

Yaxiong Wang
Hefei University of Technology
Hefei, China
yaxiong.wang15@gmail.com

Xuelin Qian
Fudan University
Shanghai, China
xlqian@fudan.edu.cn

Zhun Zhong
University of Nottingham
United Kingdom
zhunzhong007@gmail.com

Zheng Wang
Wuhan University
Wuhan, China
wangzwhu@whu.edu.cn

Liang Zheng
Australian National University
Australia
liang.zheng@anu.edu.au

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

ACM Reference Format:

Zhedong Zheng, Yaxiong Wang, Xuelin Qian, Zhun Zhong, Zheng Wang, and Liang Zheng. 2023. MORE '24 Multimedia Object Re-ID: Advancements, Challenges, and Opportunities. In *Proceedings of the 31st ACM International*

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

ICMR '24, June 10-13, 2024, Phuket, Thailand

© 2023 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0108-5/23/10.

<https://doi.org/10.1145/3581783.3610937>

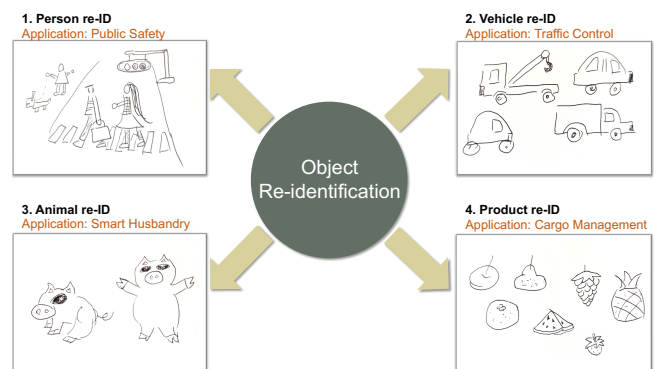


Figure 1: Example use cases of object re-identification with regard to the subjects to be re-identified [18].

Conference on Multimedia Retrieval (ICMR), June 10-13, 2024, Phuket, Thailand. ACM, New York, NY, USA, 3 pages. <https://doi.org/10.1145/3581783.3610937>

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

Table 1: Schedule of workshop activities.

Topic	Duration	Speaker	Organization
• Morning Schedule			
An opening of the workshop	5 min	Zhedong Zheng	University of Macau
Privacy-protected Person Re-identification	30 min	Yutian Lin	Wuhan University
Task Residual for Tuning Vision-Language Models	30 min	Zhihe Lu	National University of Singapore
Coffee Break			
Fine-Grained Learning from Coarse Labels	30 min	Yifan Sun	Baidu Inc.
Data-centric Learning for Object Re-ID	30 min	Liang Zheng	Australian National University
Lifelong Person Re-Identification	30 min	Zhun Zhong	University of Nottingham
<i>Round Table Discussion</i>	30 min	Workshop Host	
• Afternoon Schedule			
Paper1 Presentation	20 min	TBD	
Paper2 Presentation	20 min	TBD	
Paper3 Presentation	20 min	TBD	
...			

camera network [9, 11], smart husbandry by re-identifying animals [2], geo-localization [17] and product re-identification for cargo management [6] (refer Fig. 1).

This workshop is designed to unite researchers, practitioners, and enthusiasts passionate about object re-identification, providing a platform to delve into the latest advancements, challenges, and opportunities within this dynamic field. Encompassing a wide array of topics, the workshop covers aspects such as new datasets and benchmarks, deep metric learning, multi-view data generation, video-based object re-identification, cross-domain object re-identification, and real-world applications. For researchers, it is an opportunity to highlight their work, while practitioners can stay informed about the most recent developments in object re-identification technology. Beyond a mere gathering, this workshop establishes a unique space for exploring the swiftly evolving field of object re-identification. Its profound impact is not limited to the advancement of multimedia analysis and retrieval capabilities but extends to addressing the pressing challenges posed by the increasing complexity of safety systems.

2 TOPICS AND THEMES

Topics covered in this workshop (but not limited to) is as follows:

- New Datasets and Benchmarks
- Deep Metric Learning
- Multi-view Data Generation
- Video-based Object Re-identification
- Cross-domain Object Re-identification
- Object Re-identification Domain Adaptation / Generalization
- Single/ Multiple Object Tracking
- Object Geo-localization
- Multimedia Re-ranking

3 ACTIVITIES AND INVITED KEYNOTES

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 accepted workshop presentations. The speakers are the experts on the relevant community from different organizations globally. The schedule of the workshop activities are listed in Table 1.

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 of Technology, 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://dblp.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 in the School of Data Science, Fudan 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

- [1] Shutao Bai, Bingpeng Ma, Hong Chang, Rui Huang, and Xilin Chen. 2022. Salient-to-broad transition for video person re-identification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 7339–7348.
- [2] Rotimi-Williams Bello, Ahmad Mohamed, and Abdullah Talib. 2022. Smart animal husbandry: A review of its data, applications, techniques, challenges and opportunities. *Applications, Techniques, Challenges and Opportunities (May 8, 2022)* (2022).
- [3] Shuang Li, Tong Xiao, Hongsheng Li, Bolei Zhou, Dayu Yue, and Xiaogang Wang. 2017. Person search with natural language description. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 1970–1979.
- [4] Xuelin Qian, Yanwei Fu, Tao Xiang, Yu-Gang Jiang, and Xiangyang Xue. 2019. Leader-based multi-scale attention deep architecture for person re-identification. *IEEE transactions on pattern analysis and machine intelligence* 42, 2 (2019), 371–385.
- [5] Xuelin Qian, Wenxuan Wang, Li Zhang, Fangrui Zhu, Yanwei Fu, Tao Xiang, Yu-Gang Jiang, and Xiangyang Xue. 2020. Long-term cloth-changing person re-identification. In *Proceedings of the Asian Conference on Computer Vision*.
- [6] Chenggui Sun, Li Bin Song, and Lihang Ying. 2022. Product Re-identification System in Fully Automated Defect Detection. In *International Conference on Smart Multimedia*. Springer, 144–156.
- [7] Yaxiong Wang, Hao Yang, Xiuxiu Bai, Xueming Qian, Lin Ma, Jing Lu, Biao Li, and Xin Fan. 2021. PFAN++: Bi-Directional Image-Text Retrieval With Position Focused Attention Network. *IEEE Trans. Multim.* 23 (2021), 3362–3376. <https://doi.org/10.1109/TMM.2020.3024822>
- [8] Yaxiong Wang, Hao Yang, Xueming Qian, Lin Ma, Jing Lu, Biao Li, and Xin Fan. 2019. Position Focused Attention Network for Image-Text Matching. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019, Macao, China, August 10-16, 2019*, Sarit Kraus (Ed.). ijcai.org, 3792–3798. <https://doi.org/10.24963/IJCAI.2019/526>
- [9] Zhongdao Wang, Liang Zheng, Yixuan Liu, Yali Li, and Shengjin Wang. 2020. Towards real-time multi-object tracking. In *European Conference on Computer Vision*. Springer, 107–122.
- [10] Longhui Wei, Xiaobin Liu, Jianing Li, and Shiliang Zhang. 2018. VP-ReID: Vehicle and person re-identification system. In *Proceedings of the 2018 ACM on International Conference on Multimedia Retrieval*. 501–504.
- [11] Fan Yang, Zheng Wang, Yang Wu, Sakriani Sakti, and Satoshi Nakamura. 2022. Tackling multiple object tracking with complicated motions—re-designing the integration of motion and appearance. *Image and Vision Computing* 124 (2022), 104514.
- [12] Shuyu Yang, Yinan Zhou, Zhedong Zheng, Yaxiong Wang, Li Zhu, and Yujiao Wu. 2023. Towards Unified Text-based Person Retrieval: A Large-scale Multi-Attribute and Language Search Benchmark. In *Proceedings of the 31st ACM International Conference on Multimedia, MM 2023, Ottawa, ON, Canada, 29 October 2023- 3 November 2023*, Abdulmoteleb El-Saddik, Tao Mei, Rita Cucchiara, Marco Bertini, Diana Patricia Tobon Vallejo, Pradeep K. Atrey, and M. Shamim Hossain (Eds.). ACM, 4492–4501. <https://doi.org/10.1145/3581783.3611709>
- [13] Liang Zheng, Zhi Bie, Yifan Sun, Jingdong Wang, Chi Su, Shengjin Wang, and Qi Tian. 2016. Mars: A video benchmark for large-scale person re-identification. In *European conference on computer vision*. Springer, 868–884.
- [14] Liang Zheng, Liyue Shen, Lu Tian, Shengjin Wang, Jingdong Wang, and Qi Tian. 2015. Scalable person re-identification: A benchmark. In *Proceedings of the IEEE international conference on computer vision*. 1116–1124.
- [15] Zhedong Zheng, Tao Ruan, Yunchao Wei, Yi Yang, and Tao Mei. 2020. VehicleNet: Learning robust visual representation for vehicle re-identification. *IEEE Transactions on Multimedia* 23 (2020), 2683–2693.
- [16] Zhedong Zheng, Xiaohan Wang, Nenggan Zheng, and Yi Yang. 2022. Parameter-efficient person re-identification in the 3d space. *IEEE Transactions on Neural Networks and Learning Systems* (2022).
- [17] Zhedong Zheng, Yunchao Wei, and Yi Yang. 2020. University-1652: A multi-view multi-source benchmark for drone-based geo-localization. In *Proceedings of the 28th ACM international conference on Multimedia*. 1395–1403.
- [18] Zhedong Zheng and Liang Zheng. 2024. Object Re-identification: Problems, Algorithms and Responsible Research Practice. In *The Boundaries of Data*. Amsterdam University Press.
- [19] Zhedong Zheng, Liang Zheng, and Yi Yang. 2017. Unlabeled samples generated by gan improve the person re-identification baseline in vitro. In *Proceedings of the IEEE international conference on computer vision*. 3754–3762.
- [20] Zhedong Zheng, Jiayin Zhu, Wei Ji, Yi Yang, and Tat-Seng Chua. 2022. 3D Magic Mirror: Clothing Reconstruction from a Single Image via a Causal Perspective. *arXiv preprint arXiv:2204.13096* (2022).