

# SHITONG SUN

+447423283838 · shitong.sun@qmul.ac.uk · GitHub: <https://suntongtongtong.github.io> ·

## RESEARCH INTERESTS

Multimodal Learning, Trustworthy AI, Image Retrieval, Video Understanding

## EDUCATION

<b>Queen Mary University of London, PhD</b>	2020 - 2024
<ul style="list-style-type: none"><li>• Thesis topic: Knowledge Transfer in Deep Learning Generalisation</li><li>• Supervisor: Prof. Shaogang Gong, Fellow of the Royal Academy of Engineering (FREng)</li></ul>	
<b>University of Trento, Visiting Scholar</b>	2023 - 2024
<ul style="list-style-type: none"><li>• Research topic: Trustworthy AI</li><li>• Supervisor: Prof. Nicu Sebe</li></ul>	
<b>KU Leuven, Master</b>	2017 - 2018
<ul style="list-style-type: none"><li>• Advanced Master of Artificial Intelligence</li><li>• Cum Laude (with Honor)</li></ul>	
<b>KU Leuven, Bachelor</b>	2015 - 2017
<ul style="list-style-type: none"><li>• Electronic Engineering</li><li>• Double degree program with Beijing Jiaotong University</li><li>• Cum Laude (with Honor)</li></ul>	
<b>Beijing Jiaotong University, Bachelor</b>	2013 - 2017
<ul style="list-style-type: none"><li>• Electrical Engineering and Automation</li><li>• Pilot Class: Zhou Enlai Class Member (Top 5%)</li></ul>	

## WORKING EXPERIENCE

<b>Huawei London, Research Intern</b>	2024 - Now
<ul style="list-style-type: none"><li>• <b>Smart Glasses:</b> Leveraging Multimodal Large Language Models (MLLM) for long video understanding.</li></ul>	
<b>Queen Mary University of London, PhD student</b>	2020 - 2024
<ul style="list-style-type: none"><li>• <b>Multimodal Large Language Models:</b> In published papers, extensively utilized various advanced multimodal large language models, including CLIP (image-text alignment), BLIP (vision-language understanding), LLaVA (vision-language understanding), ChatGPT (language understanding), GroundingDINO (object detection), Stable Diffusion (image generation), and Segment Anything (image segmentation).</li><li>• <b>Trustworthy AI:</b> Conducted in-depth research on privacy-preserving federated learning based on distributed learning, proposing privacy protection solutions at both image and attribute levels, systematically addressing data security challenges.</li></ul>	
<b>Queen Mary University of London, Teaching Fellow</b>	2024
<ul style="list-style-type: none"><li>• Algorithms and Data Structures</li></ul>	
<b>Queen Mary University of London, Teaching Assistant</b>	
<ul style="list-style-type: none"><li>• Machine Learning for Visual Data Analysis</li><li>• Machine Learning</li><li>• Machine Learning</li><li>• Deep Learning and Computer Vision</li><li>• Deep Learning and Computer Vision</li></ul>	
	2024 2023 2022 2022 2021
<b>Institute of Automation, Chinese Academy of Science, Algorithm Engineer</b>	2018 - 2020
<ul style="list-style-type: none"><li>• <b>Government project:</b> Reconstruction of human 3D model based on Densepose, fix it to Intelligent Scene Exploration and Evaluation (ISEE) platform</li><li>• <b>Real-world project:</b> Apply pedestrian-attribute-recognition algorithm to Skiing app</li><li>• <b>Research:</b> Develop ReID algorithm and enhance its performance on market1501, Duke datasets</li></ul>	

- **Master Topic:** Using data prototypes to avoid forgetting in continual learning
- **Supervisor:** Prof. Tinne Tuytelaars

## PUBLICATIONS

- [1] **Shitong Sun**, Chenyang Si, Guile Wu, Shaogang Gong. “Federated Zero-Shot Learning with Mid-Level Semantic Knowledge Transfer”, in **Pattern Recognition**, Vol.156, 110824, December 2024.
- [2] **Shitong Sun**, Guile Wu, Shaogang Gong. “Decentralised Person Re-Identification with Selective Knowledge Aggregation”, in **British Machine Vision Conference**, 2021.
- [3] **Shitong Sun**, Jindong Gu, Shaogang Gong. “Benchmarking Robustness of Compositional Image Retrieval”, in **NeurIPS 2023 Workshop**.
- [4] **Shitong Sun**, Fanghua Ye, Shaogang Gong. “Training-free Zero-shot Composed Image Retrieval with Local Concept Reranking”, in submission to IEEE Transactions on Circuits and Systems for Video Technology.
- [5] Jingwen Guo, Hong Liu, **Shitong Sun**, Tianyu Guo, Min Zhang, Chenyang Si. “FSAR: Federated Skeleton-based Action Recognition with Adaptive Topology Structure and Knowledge Distillation”, in **International Conference on Computer Vision**, 2023.
- [6] Jiayi Lin, **Shitong Sun**<sup>\*</sup>, Shaogang Gong. “GridCLIP: One-Stage Object Detection by Grid-Level CLIP Representation Learning”, in submission to Pattern Recognition.
- [7] Qilei Li, **Shitong Sun**<sup>\*</sup>, Weitong Cai, Shaogang Gong. “Leverage Cross-Domain Variations for Generalizable Person ReID Representation Learning”, in submission to Pattern Recognition.

## AWARDS

- |  |             |
|--|-------------|
| • Research funding for Trustworthy AI, Alan Turing Institute | £4000, 2023 |
| • Graduated with Cum Laude Master Honors, KU Leuven          | 2018        |
| • Graduated with Cum Laude Bachelor Honors, KU Leuven        | 2017        |
| • Study Scholarship, Beijing Jiaotong University             | 2015        |
| • Study Scholarship, Beijing Jiaotong University             | 2014        |

## SKILLS

- Language: English IELTS 6.5, fluent in both spoken and written English
- 7 years of coding experience, familiar with Python, PyTorch, MLLM

## ACADEMIC SERVICE

### Journal Reviewer

- International Journal of Computer Vision
- Neurocomputing
- IEEE Transactions on Emerging Topics in Computational Intelligence
- IEEE Internet of Things Journal

### Conference Reviewer

- IEEE / CVF Computer Vision and Pattern Recognition Conference (CVPR)
- Conference on Neural Information Processing Systems (NeurIPS)
- International Conference on Learning Representations (ICLR)

### Workshop Organization

- Chaired the BMVA workshop on “Trustworthy Multimodal Learning with Foundation Models”, organizing keynote talks from Samsung, Toshiba, Wayve and universities including Oxford, Glasgow and Sheffield, attracting 80+ industry professionals