Xiao (Brandon) **Han**

Desk 01, 01BB01, University of Surrey, Guildford, Surrey, GU2 7XH, United Kingdom

(+44) 07529989025 | xiao.han@surrey.ac.uk | brandonhan.uk | GitHub | LinkedIn | GoogleScholar

Research Interests

I am broadly interested in the field of Deep Learning. My current research interest lies in the intersection between Computer Vision and Natural Language Processing (*i.e.*, V+L). My research goal is to build multi-modal AI systems that can be used in real-world applications. My expertise includes but is not limited to V+L **pre-training** and (parameter-efficient) **adaptation**, many V+L tasks (*e.g.*, cross-modal/compositional **retrieval**, image **captioning**, text-based/guided **2D/3D** contents **generation/editing**), and some specific tasks (*e.g.*, person ReID).

Education

University of Surrey — *Guildford, UK*

Ph.D. Student — *Jan. 2021 - Present*

- Major in Vision and Signal Processing at Centre for Vision, Speech and Signal Processing (CVSSP)
- Supervisors: Prof. Tao Xiang (principal) and Prof. Yi-Zhe Song
- Fully funded by University of Surrey FEPS/iFlyTek Ph.D. Scholarship
- Thesis Title (tentative): Fine-Grained Multimodal Deep Learning

Zhejiang University — *Hangzhou, China*

Bachelor of Engineering — *Sep. 2016 - Jun. 2020*

- Major in Information Engineering at College of Information Science and Electronic Engineering (ISEE)
- Cumulative GPA: 3.92 / 4.00 (88.13 / 100), Third-year GPA: 3.94 / 4.00 (89.46 / 100)
- Thesis Title: Deep Learning-Based Features Prediction for Mass Spectrometry of Protein

Non-Degree Academic Experiences:

2020 (09-11)	Visiting student with Dr. Li Zhang at ZVG, Fudan University	*Shanghai, China*
2019 (10-05)	Research assistant with Dr. Changbin Yu at AiR, Westlake University	*Hangzhou, China*
2019 (07-10)	Visiting student with Prof. L. Jay Guo at EECS, University of Michigan	*Ann Arbor, MI, USA*
2018 (08-09)	Exchange student with iESR program at University of Notre Dame	*South Bend, IN, USA*

Work Experiences

Noah's Ark Lab — *London, UK*

Research Intern (PT) — *Dec. 2022 - Present*

- Mentor: Jiankang Deng
- Project: Text-based 2D/3D contents generation/editing via Diffusion Models

PixelShift.AI — *Shanghai, China*

Computer Vision Algorithm Intern (FT) — *May 2020 - Sep. 2020*

- Mentors: Zhiming Ma and Meng Zhang
- Project: Immersive AR application (Google MediaPipe) and deployment of generative models (TFLite)

Publications

(* Equal contribution, † Corresponding author)

Conference

HeadSculpt: Crafting 3D Head Avatars with Text — *Under Review*

Xiao Han*, Yukang Cao*, Kai Han, Xiatian Zhu, Jiankang Deng, Yi-Zhe Song, Tao Xiang†, Kwan-Yee K. Wong† — *Nov. 2023*

Controllable Person Image Synthesis with Pose-Constrained Latent Diffusion — *Under Review*

Xiao Han, Jiankang Deng, Xiatian Zhu, Yi-Zhe Song, Tao Xiang — *Oct. 2023*

FAME-ViL: Multi-Tasking Vision-Language Model for Heterogeneous Fashion Tasks *(Highlight, Top 2.5%)* — *CVPR 2023*

Xiao Han, Xiatian Zhu, Licheng Yu, Li Zhang, Yi-Zhe Song, Tao Xiang — *Jun. 2023*

FashionViL: Fashion-Focused Vision-and-Language Representation Learning *ECCV 2022*
Xiao Han, Licheng Yu, Xiatian Zhu, Li Zhang, Yi-Zhe Song, Tao Xiang *Oct. 2022*

UIGR: Unified Interactive Garment Retrieval *CVPRW 2022*
Xiao Han, Sen He, Li Zhang, Yi-Zhe Song, Tao Xiang *Jun. 2022*

Text-Based Person Search with Limited Data *BMVC 2021*
Xiao Han, Sen He, Li Zhang, Tao Xiang *Nov. 2021*

Preprint

Large-Scale Product Retrieval with Weakly Supervised Representation Learning *arXiv*
Xiao Han*, Kam Woh Ng*, Sauradip Nag, Zhiyu Qu *Aug. 2022*

Journal

Inverse Design of Metasurface Optical Filters using Deep Neural Network with High Degrees of Freedom *InfoMat*
Xiao Han*, Ziyang Fan*, Zeyang Liu*, Chao Li, and L. Jay Guo *Jun. 2020*

Open-sourced Projects

FashionMMF: A codebase for fashion-related vision-and-language research (based on Meta AI MMF)
- https://github.com/BrandonHanx/mmf and https://github.com/BrandonHanx/FAME-ViL
- Implementation and extension of our paper *FashionViL (ECCV 2022)* and *FAME-ViL (CVPR 2023)*

eBayChallenge: A modularized codebase for large-scale product retrieval (based on PyTorch Lightning and Hydra)
- https://github.com/01BB01/eBayChallenge
- Implementation of our solution for *eBay eProduct Visual Search Challenge - FGVC9 (CVPR2022)*

CompFashion: A modularized codebase for text-guided image retrieval (based on vanilla PyTorch)
- https://github.com/BrandonHanx/CompFashion
- Implementation and extension of our paper *UIGR: Unified Interactive Garment Retrieval (CVPRW 2022)*

TextReID: A modularized codebase for text-based person search (based on vanilla PyTorch)
- https://github.com/BrandonHanx/TextReID
- Implementation and extension of our paper *Text-Based Person Search with Limited Data (BMVC 2021)*

Honors & Awards

2022	2nd Place, eBay eProduct Visual Search Challenge - FGVC9 (CVPR2022)	*eBay, USA*
2021-2024	Faculty of Engineering and Physical Sciences/iFlytek Scholarship	*University of Surrey/iFlytek, UK*
2019	Chunzhen International Exchange Scholarship	*Zhejiang University, China*
2017-2019	3rd Prize, Academic & Outstanding Student Scholarship	*Zhejiang University, China*
2018	Yongping Scholarship	*Zhejiang University, China*
2018	3rd Prize, National Talent Training Base Scholarship	*Zhejiang University, China*
2018	Honorable Mention Prize, Mathematical Contest in Modeling (MCM)	*COMAP, USA*
2017	3rd Prize, Physics Innovation Competition in Zhejiang Province (Theory Part)	*Zhejiang Physical Society, China*

Services

Conference reviewer
- Conference on Neural Information Processing Systems (NeurIPS 2023)
- 6th CVPR Workshop on Computer Vision for Fashion, Art, and Design (CVFAD) 2023
- International Conference on Computer Vision (ICCV) 2023
- IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2023
- AAAI Conference on Artificial Intelligence (AAAI) 2023
- European Conference on Computer Vision (ECCV) 2022
- ACM International Conference on Multimedia (ACM MM) 2022
- IEEE International Conference on Multimedia and Expo (ICME) 2022

Journal reviewer
- IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)

Teaching assistant / Demonstrator at University of Surrey

- EEEM071 (2023 Spring) - Advanced Topics in Computer Vision and Deep Learning (under Prof. Yi-Zhe Song and Dr. Xiatian Zhu)
- Responsible for the final coursework design (Vehicle ReID) and lab tutorial preparation
- Responsible for demonstrating in lab sessions with 145+ PG/T students with mixed backgrounds
- Responsible for marking the coursework reports

Project tutor at University of Surrey

- EEEM004 (2023 Spring) - 60 Credit Standard Project (under Dr. Xiatian Zhu)
- Responsible for guiding PG/T final projects “Zero-/Few-shot learning with language supervision”

Skills

Programming Languages	Python, C/C++, MATLAB, LaTeX, Verilog, JavaScript/TypeScript, HTML/CSS
Frameworks and Tools	PyTorch, PyTorch Lightning, TensorFlow, Git, Docker, W&B
Codebases	Meta AI MMF, HuggingFace Pipeline (*e.g.*, Transformers/Diffusers), Google MediaPipe, timm
Languages	普通话 Mandarin (native), English (fluent), 한국어 Korean (beginner)

References

Will be provided upon request