

# 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.*, vision-language). My research goal is to build multi-modal AI systems that can be used in real-world applications (*e.g.*, e-commerce platform). My expertise includes but is not limited to vision-language pre-training and (parameter-efficient) adaptation, many vision-language tasks (*e.g.*, cross-modal retrieval, text-guided image retrieval, image captioning, text-based/guided image generation/editing), and some specific tasks (*e.g.*, person ReID).

## Education

### University of Surrey

PH.D. STUDENT

Guildford, UK

Jan. 2021 - Present (Jan. 2024)

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

### Zhejiang University

BACHELOR OF ENGINEERING

Hangzhou, China

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

RESEARCH INTERN (PT)

London, UK

Dec. 2022 - Present

- Mentor: Jiankang Deng
- Project: Vision-Language Prompt Tuning, Diffusion-based Image Generation/Editing

### PixelShift.AI

COMPUTER VISION ALGORITHM INTERN (FT)

Shanghai, China

May 2020 - Sep. 2020

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

## Publications

### Conference

#### FAME-ViL: Multi-Tasking Vision-Language Model for Heterogeneous Fashion Tasks

Under Review

Xiao Han, Xiatian Zhu, Licheng Yu, Li Zhang, Yi-Zhe Song, Tao Xiang

Nov. 2022

#### 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

**Large-Scale Product Retrieval with Weakly Supervised Representation Learning**

Xiao Han\*, Kam Woh Ng\*, Sauradip Nag, Zhiyu Qu

arXiv

Aug. 2022

**Journal****Inverse Design of Metasurface Optical Filters using Deep Neural Network with High Degrees of Freedom**

Xiao Han\*, Ziyang Fan\*, Zeyang Liu\*, Chao Li, and L. Jay Guo

InfoMat

Jun. 2020

**Open-sourced Projects****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)*

**FashionViL: A codebase for fashion-related vision-and-language research (based on Meta AI MMF)**

- <https://github.com/BrandonHanx/mmf>
- Implementation and extension of our paper *FashionViL: Fashion-Focused Vision-and-Language Representation Learning*

**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*

**TextRelD: A modularized codebase for text-based person search (based on vanilla PyTorch)**

- <https://github.com/BrandonHanx/TextRelD>
- Implementation and extension of our paper *Text-Based Person Search with Limited Data*

**Honors & Awards**

2022	2 <sup>nd</sup> 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	3 <sup>rd</sup> Prize, Academic & Outstanding Student Scholarship	Zhejiang University, China
2018	Yongping Scholarship	Zhejiang University, China
2018	3 <sup>rd</sup> Prize, National Talent Training Base Scholarship	Zhejiang University, China
2018	Honorable Mention Prize, Mathematical Contest in Modeling (MCM)	COMAP, USA
2017	3 <sup>rd</sup> Prize, Physics Innovation Competition in Zhejiang Province (Theory Part)	Zhejiang Physical Society, China

**Services****Conference reviewer**

- IEEE 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)

**Skills**

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

**References**

Will be provided upon request