

# Xin WEN

The University of Hong Kong

(+86) 180-1908-1031

wenxin@eee.hku.hk

[github.com/xwen99](https://github.com/xwen99)

## EDUCATION

### The University of Hong Kong

Ph.D. in Electrical and Electronic Engineering

Hong Kong, China

Sept. 2021 - July 2025 (expected)

- Advisor: [Prof. Xiaojuan Qi](#) and [Prof. Jiaya Jia](#).
- Research area: Computer Vision, Deep Learning, and Machine Learning.

### Tongji University

B.Eng. in Computer Science

Shanghai, China

Sept. 2017 - July 2021

- I joined the Deep Learning Lab of Tongji University lead by [Prof. Yin Wang](#) at Sept. 2018.

B.Sc. in Mathematics (Minor)

Sept. 2017 - Jan. 2019

- Mathematical Intensive Training Class (Innovation Experimental Area)

### The University of Manchester

United in Manchester International Summer School

Manchester, UK

July 2019 - Aug. 2019

## PUBLICATIONS

\*: Equal contribution

1. *Temporal Context Aggregation for Video Retrieval with Contrastive Learning*. Jie Shao\*, **Xin Wen\***, Bingchen Zhao, Xiangyang Xue. WACV 2021. [\[PDF\]](#) [\[code\]](#)
2. *Distilling Visual Priors from Self-Supervised Learning*. Bingchen Zhao, **Xin Wen**. ECCV-W 2020. [\[PDF\]](#) [\[code\]](#)
3. *Self-Supervised Visual Representation Learning with Semantic Grouping*. **Xin Wen**, Bingchen Zhao, Anlin Zheng, Xiangyu Zhang, Xiaojuan Qi. (Under review)
4. *Gaussian Mixture Model for Generalized Category Discovery with Prototypical Contrastive Learning*. Bingchen Zhao, **Xin Wen**, Kai Han. (Under review)

## RESEARCH EXPERIENCES

### CVMI Lab, The University of Hong Kong

Research Assistant

Hong Kong, China

Jan. 2021 - Present

- Advised by [Prof. Xiaojuan Qi](#), I am working on a project of joint scene-decomposition and self-supervised visual representation learning from scene-centric images.

### Visual Computing Group, ByteDance AI Lab

Research Intern

Shanghai, China

Jan. 2020 - June 2021

- Worked on Video Representation Learning and Video Retrieval, advised by [Dr. Jie Shao](#) and [Prof. Xiangyang Xue](#).
- We proposed a **self-attention** based feature-aggregation method with **supervised contrastive learning** to better exploit temporal correlations within video frames. The proposed method TCA [\[1\]](#) deliver competitive results with faster inference time.

### ECCV 2020 Workshop VIPriors Image Classification Challenge

Core Member

Virtual

June. 2020 - July 2021

- In view of the data-efficient image classification setting, we first leverage self-supervised **contrastive learning** to train a generalizable teacher model, then distill the knowledge from the teacher model to the student model in a **self-distillation** manner while fine-tuning with cross-entropy loss. This work [\[2\]](#) won **2nd** place in the ECCV-W 2020 VIPriors challenge. [\[Ranking\]](#)

## ACADEMIC SERVICES

- Reviewer for ECCV 2022, CVPR 2022, and WACV 2022.

## SELECTED AWARDS

- Shanghai Outstanding Graduates (June 2021)
- Ranked **2nd** place in the ECCV 2020 Workshop VIPriors Image Classification Challenge (July 2020)
- Qidi Scholarship of Tongji University (**Top 1%**, June 2020)
- **Regional Champion (China)** of the Covestro International Data Science Hackathon (Nov. 2019)
- **Silver Medal** of the 43rd ACM International Collegiate Programming Contest (ICPC) Asia-East Continent Final (Dec. 2018)

## SKILLS

- **Programming:** C/C++, Python, MATLAB, Verilog HDL, HTML, JavaScript, React.js
- **Platform and Toolkits:** Linux, Bash, Git, LaTeX, PyTorch, TensorFlow, OpenCV, OpenGL, Vim, MySQL, FFmpeg
- **Multimedia:** Adobe Photoshop, Illustrator, Premiere Pro, Audition, After Effects, Acrobat DC, XD