

XUESONG NIE

[🏠 Homepage](#)
[✉ xuesongnie@zju.edu.cn](mailto:xuesongnie@zju.edu.cn)
[🔍 Google Scholar](#)
[🌐 LinkedIn](#)
[🐙 Github](#)
[🐦 Twitter](#)

EDUCATION

- Zhejiang University (ZJU), China** | Supervisor: Prof. **Donglian Qi** Sep. 2022 – Mar. 2025
M.S. Student in Electronic Information Engineering
- Henan University (Henu), China** | Supervisor: Prof. **Lijia Chen** Sep. 2018 – Jun. 2022
Bachelor of Science in Communication Engineering

SELECTED PUBLICATIONS

CONFERENCE PAPERS

- PredToken: Predicting Unknown Tokens and Beyond with Coarse-to-Fine Iterative Decoding**
Xuesong Nie, Haoyuan Jin, Yunfeng Yan, Xi Chen, Zhihang Zhu, Donglian Qi
 IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**), 2024
- Wavelet-Driven Spatiotemporal Predictive Learning: Bridging Frequency and Time Variations**
Xuesong Nie, Yunfeng Yan, Siyuan Li, Cheng Tan, Xi Chen, Haoyuan Jin, Zhihang Zhu, Stan Z. Li, Donglian Qi
 AAAI Conference on Artificial Intelligence (**AAAI**), 2024
- Triplet Attention Transformer for Spatiotemporal Predictive Learning**
Xuesong Nie, Xi Chen, Haoyuan Jin, Zhihang Zhu, Yunfeng Yan, Donglian Qi
 IEEE/CVF Winter Conference on Applications of Computer Vision (**WACV**), 2024
- AMD: Towards Robust Appearance-Motion Disentanglement for Predictive Learning**
Xuesong Nie, Haoyuan Jin, Yunfeng Yan, Xi Chen, Zhihang Zhu, Donglian Qi
 European Conference on Computer Vision (**ECCV**), **Under Review**, 2024
- Object-Level Pseudo-3D Lifting for Distance-Aware Tracking**
 Haoyuan Jin*, **Xuesong Nie***, Yunfeng Yan, Xi Chen, Zhihang Zhu, Donglian Qi
 ACM International Conference on Multimedia (**ACMMM**), **Under Review**, 2024
- SAMP: Adapting Segment Anything Model for Pose Estimation**
 Zhihang Zhu, Yunfeng Yan, Yi Chen, Haoyuan Jin, **Xuesong Nie**, Donglian Qi, Xi Chen
 IEEE International Conference on Multimedia and Expo (**ICME**), 2024

JOURNAL PAPERS

* denotes equal contribution

- ScopeViT: Scale-Aware Vision Transformer**
Xuesong Nie, Haoyuan Jin, Yunfeng Yan, Xi Chen, Zhihang Zhu, Donglian Qi
 Pattern Recognition (**PR**), 2024
- AHOR: Online Multi-object Tracking with Authenticity Hierarchizing and Occlusion Recovery**
 Haoyuan Jin*, **Xuesong Nie***, Yunfeng Yan, Xi Chen, Zhihang Zhu, Donglian Qi
 IEEE Transactions on Circuits and Systems for Video Technology (**TCSVT**), 2024

RESEARCH EXPERIENCES

Generative Models for Images and Videos Alibaba, Mentor: Dr. Xi Chen	Sep. 2023 – Present
<ul style="list-style-type: none">• Topic: Explore mainstream video and image generation frameworks, such as Vector Quantization, Autoregressive, Non-Autoregressive, and Diffusion-based models.• Proposed a high-quality visual generative framework PredToken, published at CVPR 2024, which ensures low-level consistency and captures high-level dynamics by decoupling space-time tokens for iterative cascaded decoding.	
Spatiotemporal Predictive Learning Westlake University, Mentor: Prof. Stan Z. Li	Mar. 2023 – Aug. 2023
<ul style="list-style-type: none">• Topic: Investigate spatiotemporal modeling methods and introduce a new spatiotemporal predictive benchmark.• Developed OpenSTL, an open-source spatiotemporal predictive library, supports various methods and tasks from synthetic to real-world data. Presented an efficient recurrent-free model WaST, published at AAAI 2024.	
Efficient Visual Perception Backbones Alibaba, Mentor: Dr. Xi Chen	Sep. 2022 – Feb. 2023
<ul style="list-style-type: none">• Topic: Research focuses on the design of efficient visual perception backbones based on CNNs and Transformers.• Proposed a scale-aware vision transformer called ScopeViT, that implements attention at different scales within a single building block to effectively learn inter-object relationships, published at Pattern Recognition 2024.	
National Major Scientific Research Instrument Development Zhejiang University	Sep. 2022 – Present
<ul style="list-style-type: none">• Topic: Developing a video and image detection platform for defect detection in power transmission and transformation equipment. (Supervisor: Prof. Donglian Qi, Project Funding: 9.616 million CNY)• Proposed an online multi-object tracker with authenticity hierarchizing and occlusion recovery AHOR and a novel video transformer with triplet attention TAT, published in TCSVT 2024 and WACV 2024 respectively.	

HONORS AND AWARDS

National Scholarship (M.S. Student) Chinese Government	Oct. 2023
Highest scholarship awarded by Chinese Government, top 0.1%	
Golden Age Scholarship Zhejiang University	2022 – Present
Awarded to outstanding full-time students at Zhejiang University for 3 years	
Outstanding Graduate of Henan Province	Feb. 2022
Awarded to graduates with academic achievements, top 0.1%	
The Stars of Self-improvement of Chinese College Students China Youth Daily	Dec. 2021
Highest national honorary awards at the spiritual level, top 2 campus-wide	
1st Prize in the 6th China Undergraduate Physics Experiment Competition	Dec. 2020
Highest-level national physics-related competition in China	
1st Prize in the 11th “Blue Bridge Cup” National Competition	Nov. 2020
Highest-level national programming competition in China	

ACADEMIC SERVICES

Computer Vision and Pattern Recognition Zhejiang University, Teaching Assistant	Spring 2024
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Reviewer	2023
European Conference on Computer Vision (ECCV) Reviewer	2024
ACM International Conference on Multimedia (ACMMM) Reviewer	2024
IEEE International Conference on Multimedia and Expo (ICME) Reviewer	2024

TECHNICAL SKILLS

Programming Languages: Python, C/C++, Matlab, HTML/CSS, JavaScript, L ^A T _E X	
Libraries and Tools: PyTorch, Docker, Linux, Git, NumPy, Matplotlib, Pandas, Scikit-learn	