

Yunjie Tian

tianyunjie96@gmail.com | sunsmarterjie.github.io | Google Scholar | +1 716 709 1071

Work Experience

10/25–present	Senior Research Scientist , TikTok	Seattle, WA
09/24–09/25	Postdoctoral Associate in Artificial Intelligence , University at Buffalo, SUNY	

Education

09/19–07/24	Ph.D. in Computer Vision , University of Chinese Academy of Sciences	
09/15–07/19	B.S. in Computer Science , Jilin University, China	Top 5%

Research Interests

- Visual representation learning, model design, and pre-training.
- Efficient visual perception model design.
- Post-training of large language and vision-language models.

Honors & Awards

2026	Outstanding Doctoral Dissertation Award of the Chinese Academy of Sciences (中科院“百篇优博”). National Scholarship; First-Class University Scholarship.
------	---

Open Source & Research Impact

YOLOv12	Owner; 3.0k GitHub stars.
VMamba	Main contributor; 3.2k GitHub stars.
MMPreTrain	Contributor; 3.8k GitHub stars.

Publications

Peer-Reviewed Publications

EinsPT: Efficient Instance-Aware Pre-Training of Vision Foundation Models. Zhaozhi Wang, **Yunjie Tian**, Lingxi Xie, Yaowei Wang, and Qixiang Ye. *IEEE Transactions on Image Processing (TIP)*, 2026 [**Corresponding author**].

YOLOv12: Attention-Centric Real-Time Object Detectors. **Yunjie Tian**, Qixiang Ye, and David Doermann. *NeurIPS*, 2025 [**Most Influential #2 among 5,290 papers**].

AutoEdit: Automatic Hyperparameter Tuning for Image Editing. Chau Pham, Quan Dao, Mahesh Bhosale, **Yunjie Tian**, Dimitris Metaxas, and David Doermann. *NeurIPS*, 2025 [**Corresponding author**].

PathDiff: Histopathology Image Synthesis with Unpaired Text and Mask Conditions. Mahesh Bhosale, Abdul Wasi, Yuanhao Zhai, **Yunjie Tian**, Samuel Border, Nan Xi, Pinaki Sarder, Junsong Yuan, David Doermann, and Xuan Gong. *ICCV*, 2025.

vHeat: Building Vision Models upon Heat Conduction. Zhaozhi Wang, Yue Liu, **Yunjie Tian**, Yunfan Liu, Yaowei Wang, and Qixiang Ye. *CVPR*, 2025.

ChatterBox: Multimodal Referring and Grounding with Chain-of-Questions. **Yunjie Tian**, Tianren Ma, Lingxi Xie, and Qixiang Ye. *AAAI*, 2025.

ClawMachine: Learning to Fetch Visual Tokens for Referential Comprehension. Tianren Ma, Lingxi Xie, **Yunjie Tian**, Boyu Yang, and Qixiang Ye. *ICLR*, 2025.

PLVM: A Tuning-Free Approach for Personalized Large Vision-Language Model. Chau Pham, Hoang Phan, David Doermann, and **Yunjie Tian**. *CVPR Workshops*, 2025.

Adaptive Keyframe Sampling for Long Video Understanding. Xi Tang, Jihao Qiu, Lingxi Xie, **Yunjie Tian**, Jianbin Jiao, and Qixiang Ye. *CVPR*, 2025.

Beyond Masking: Demystifying Token-Based Pre-Training for Vision Transformers. **Yunjie Tian**, Lingxi Xie, Jiemin Fang, Jianbin Jiao, and Qi Tian. *Pattern Recognition*, 2025.

VMamba: Visual State Space Model. Yue Liu, **Yunjie Tian**, Yuzhong Zhao, Hongtian Yu, Lingxi Xie, Yaowei Wang, Qixiang Ye, and Yunfan Liu. *NeurIPS*, 2024 [[Spotlight](#); [Most Influential #2 among 4,037 papers](#)].

Fast-iTPN: Integrally Pre-Trained Transformer Pyramid Network with Token Migration. **Yunjie Tian**, Lingxi Xie, Jihao Qiu, Jianbin Jiao, Yaowei Wang, Qi Tian, and Qixiang Ye. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2024.

Artemis: Towards Referential Understanding in Complex Videos. Jihao Qiu, Yuan Zhang, Xi Tang, Lingxi Xie, Tianren Ma, Pengyu Yan, David Doermann, Qixiang Ye, and **Yunjie Tian**. *NeurIPS*, 2024 [[Corresponding author](#)].

Exploring Complicated Search Spaces with Interleaving-Free Sampling. **Yunjie Tian**, Lingxi Xie, Jiemin Fang, Jianbin Jiao, Qixiang Ye, and Qi Tian. *IEEE Transactions on Neural Networks and Learning Systems (TNNLS)*, 2024.

Spatial Transform Decoupling for Oriented Object Detection. Hongtian Yu, **Yunjie Tian**, Qixiang Ye, and Yunfan Liu. *AAAI*, 2024 [[Co-first author](#)].

Integrally Pre-Trained Transformer Pyramid Networks. **Yunjie Tian**, Lingxi Xie, Zhaozhi Wang, Longhui Wei, Xiaopeng Zhang, Jianbin Jiao, Yaowei Wang, Qi Tian, and Qixiang Ye. *CVPR*, 2023.

HiViT: A Simpler and More Efficient Design of Hierarchical Vision Transformer. Xiaosong Zhang, **Yunjie Tian**, Lingxi Xie, Wei Huang, Qi Dai, Qixiang Ye, and Qi Tian. *ICLR*, 2023 [[Co-first author](#); [Spotlight](#)].

GraFormer: Graph-Oriented Transformer for 3D Pose Estimation. Weixi Zhao, Weiqiang Wang, and **Yunjie Tian**. *CVPR*, 2022.

Discretization-Aware Architecture Search. **Yunjie Tian**, Chang Liu, Lingxi Xie, Qixiang Ye, et al. *Pattern Recognition*, 2021.

Adaptive Linear Span Network for Object Skeleton Detection. Chang Liu, **Yunjie Tian**, Zhiwen Chen, Jianbin Jiao, and Qixiang Ye. *IEEE Transactions on Image Processing (TIP)*, 2021 [[Co-first author](#)].

Preprints

Self-Supervised Visual On-Policy Distillation. Yijiang Li, Yijun Liang, **Yunjie Tian**, Bingyang Wang, Ke Zhang, Zhenfei Yin, Di Fu, Philip Torr, and Nuno Vasconcelos. *arXiv preprint arXiv:2608.14144*, 2026.

On-Policy Self-Distillation without Any Supervision. Yijiang Li, Bingyang Wang, Yijun Liang, **Yunjie Tian**, Di Fu, and Nuno Vasconcelos. *arXiv preprint arXiv:2608.06296*, 2026.

Visual Contrastive Self-Distillation. Yijun Liang, **Yunjie Tian**, Yijiang Li, Yuqi Jia, Furong Huang, Tianyi Zhou, and Di Fu. *arXiv preprint arXiv:2607.21556*, 2026.

On-Policy Distillation with Best-of-N Teacher Rollout Selection. Ke Zhang, **Yunjie Tian**, Dongdi Zhao, Yijiang Li, Yuanye Liu, Vishal M. Patel, and Di Fu. *arXiv preprint arXiv:2605.09725*, 2026.

From Imitation to Intuition: Intrinsic Reasoning for Open-Instance Video Classification. Ke Zhang, Xiangchen Zhao, **Yunjie Tian**, Jiayu Zheng, Vishal M. Patel, and Di Fu. *arXiv preprint arXiv:2603.10300*, 2026.

Academic Service

Conference reviewer: CVPR (2024–2026), ICCV (2023, 2025), ECCV (2024, 2026), NeurIPS (2024–2026), ICLR (2025–2026), ICML (2025), AAAI (2025), WACV (2026).

Journal reviewer: TPAMI, IJCV, TIP, TCSVT, TITS, TMM, IJIG.

Research Leadership & Technical Experience

- Led a research team, overseeing project planning and execution.
- Developed “Dasheng,” one of the earliest 1B-parameter visual models, as a primary contributor to model training and implementation.
- Advanced remote-sensing road detection and video matting projects toward real-world application during Ph.D. research.
- 7 years of hands-on Python and PyTorch experience, including training on large-scale GPU clusters.