



# JUNFEI XIAO

Baltimore, Maryland

<https://lambert-x.github.io>

✉ [xiaojf97@gmail.com](mailto:xiaojf97@gmail.com)  

## Education

---

<b>Johns Hopkins University</b> <i>Ph.D. in Computer Science</i>	<b>2025</b> <i>Baltimore, Maryland</i>
<b>Johns Hopkins University</b> <i>M.S.E in Computer Science, GPA: 4.0/4.0</i>	<b>2021</b> <i>Baltimore, Maryland</i>
<b>Beihang University</b> <i>B.E. in Mechanical Engineering, Double Degree in Mathematics, GPA: 3.8/4.0 (Top 2%)</i>	<b>2019</b> <i>Beijing, China</i>

## Experience

---

<b>xAI</b> <i>Member of Technical Staff, Grok Imagine</i>	<b>Oct. 2025 – Present</b> <i>Palo Alto, California</i>
<b>ByteDance Seed, Video Generation Team</b> <i>Research Intern</i>	<b>May 2024 – May 2025</b> <i>San Jose, California</i>
<b>Google Research</b> <i>Student Researcher</i>	<b>July 2023 – Feb. 2024</b> <i>Mountain View, California</i>
<b>CCVL Lab, Johns Hopkins University</b> <i>Research Assistant, Advisor: Alan Yuille</i>	<b>Sep. 2021 – Present</b> <i>Baltimore, Maryland</i>
<b>CCVL Lab, Johns Hopkins University</b> <i>Research Intern, Advisor: Alan Yuille</i>	<b>June 2020 – Sep. 2021</b> <i>Baltimore, Maryland</i>

## Publications and Manuscripts

---

- [20] Yu-Cheng Chou, Xingrui Wang, Yitong Li, Jiahao Wang, Hanting Liu, Cihang Xie, Alan Yuille, and **Junfei Xiao**. “Captain Safari: A World Engine”. In: *CVPR*. 2026, (**Project Lead**).
- [19] Tiezheng Zhang, Yitong Li, Yu-Cheng Chou, Jieneng Chen, Alan L. Yuille, Chen Wei, and **Junfei Xiao**. “VLV: Vision-Language-Vision Auto-Encoder: Scalable Knowledge Distillation from Diffusion Models”. In: *NeurIPS*. 2025, (**Project Lead**).
- [18] Yunfei Xie, Yinsong Ma, Shiyi Lan, Alan Yuille, **Junfei Xiao**, and Chen Wei. “Play to Generalize: Learning to Reason Through Game Play”. In: *ICLR*. 2025, (**Project Lead**).
- [17] **Junfei Xiao**, Ceyuan Yang, Lvmin Zhang, Shengqu Cai, Yang Zhao, Yuwei Guo, Gordon Wetzstein, Maneesh Agrawala, Alan Yuille, and Lu Jiang. “Captain Cinema: Towards Short Movie Generation”. In: *ICLR*. 2025, (**Project Lead**).
- [16] **Junfei Xiao**, Feng Cheng, Lu Qi, Liangke Gui, Jiepeng Cen, Zhibei Ma, Alan Yuille, and Lu Jiang. “VideoAuteur: Towards Long Narrative Video Generation”. In: *ICCV*. 2025.
- [15] Shengqu Cai, Ceyuan Yang, Lvmin Zhang, Yuwei Guo, **Junfei Xiao**, Ziyang Yang, Yinghao Xu, Zhenheng Yang, Alan Yuille, Leonidas Guibas, Maneesh Agrawala, Lu Jiang, and Gordon Wetzstein. “Mixture of Contexts for Long Video Generation”. In: *ICLR*. 2025.

- [14] Taiming Lu, Tianmin Shu, **Junfei Xiao**, Luoxin Ye, Jiahao Wang, Cheng Peng, Chen Wei, Daniel Khashabi, Rama Chellappa, Alan Yuille, and Jieneng Chen. “GenEx: Generating an Explorable World”. In: *arXiv preprint arXiv:2412.09624* (2024), (**\*: Core Contributor**).
- [13] Xianhang Li, Haoqin Tu, Mude Hui, Zeyu Wang, Bingchen Zhao, **Junfei Xiao**, Sucheng Ren, Jieru Mei, Qing Liu, Huangjie Zheng, et al. “What If We Recaption Billions of Web Images with LLaMA-3?”. In: *ICML*. 2025.
- [12] **Junfei Xiao**, Zheng Xu, Alan Yuille, Shen Yan, and Boyu Wang. “PaLM2-VAdapter: Progressively Aligned Language Model Makes a Strong Vision-language Adapter”. In: *arXiv preprint arXiv:2402.10896* (2024).
- [11] **Junfei Xiao**, Ziqi Zhou, Wenxuan Li, Shiyi Lan, Jieru Mei, Zhiding Yu, Bingchen Zhao, Alan Yuille, Yuyin Zhou, and Cihang Xie. “A Semantic Space is Worth 256 Language Descriptions: Make Stronger Segmentation Models with Descriptive Properties”. In: *ECCV*. 2024.
- [10] Sucheng Ren, Xiaoke Huang, Xianhang Li, **Junfei Xiao**, Jieru Mei, Zeyu Wang, Alan Yuille, and Yuyin Zhou. “Medical Vision Generalist: Unifying Medical Imaging Tasks in Context”. In: *arXiv preprint arXiv:2406.05565* (2024).
- [9] Sucheng Ren, Zeyu Wang, Hongru Zhu, **Junfei Xiao**, Alan Yuille, and Cihang Xie. “Rejuvenating Image-GPT as Strong Visual Representation Learners”. In: *ICML*. 2024.
- [8] Jie Liu, Yixiao Zhang, Jie-Neng Chen, **Junfei Xiao**, Yongyi Lu, Bennett A. Landman, Yixuan Yuan, Alan Yuille, Yucheng Tang, and Zongwei Zhou. “CLIP-Driven Universal Model for Organ Segmentation and Tumor Detection”. In: *ICCV*. 2023, pp. 21152–21164.
- [7] Qixin Hu, Yixiong Chen, **Junfei Xiao**, Shuwen Sun, Jie-Neng Chen, Alan Yuille, and Zongwei Zhou. “Label-Free Liver Tumor Segmentation”. In: *CVPR*. 2023.
- [6] Yutong Bai, Zeyu Wang, **Junfei Xiao**, Chen Wei, Huiyu Wang, Alan Yuille, Yuyin Zhou, and Cihang Xie. “Masked Autoencoders Enable Efficient Knowledge Distillers”. In: *CVPR*. 2023, (**Technical Lead**).
- [5] **Junfei Xiao**, Yutong Bai, Alan Yuille, and Zongwei Zhou. “Delving into Masked Autoencoders for Multi-Label Thorax Disease Classification”. In: *WACV*. 2023.
- [4] Qixin Hu, **Junfei Xiao**, Yixiong Chen, Shuwen Sun, Jie-Neng Chen, Alan Yuille, and Zongwei Zhou. “Synthetic Tumors Make AI Segment Tumors Better”. In: *NeurIPS Workshop* (2022).
- [3] **Junfei Xiao**, Lequan Yu, Zongwei Zhou, Yutong Bai, Lei Xing, Alan Yuille, and Yuyin Zhou. “CateNorm: Categorical Normalization for Robust Medical Image Segmentation”. In: *MICCAI Workshop on Domain Adaptation and Representation Transfer*. Springer. 2022, (**Best Paper Honorable Mention**).
- [2] **Junfei Xiao**, Longlong Jing, Lin Zhang, Ju He, Qi She, Zongwei Zhou, Alan Yuille, and Yingwei Li. “Learning from Temporal Gradient for Semi-supervised Action Recognition”. In: *CVPR*. 2022.
- [1] Siqi Wang, Lei Li, Yufeng Chen, Yueping Wang, Wenguang Sun, **Junfei Xiao**, Dylan Wainwright, Tianmiao Wang, Robert J. Wood, and Li Wen. “A Bio-Robotic Remora Disc with Attachment and Detachment Capabilities for Reversible Underwater Hitchhiking”. In: *2019 International Conference on Robotics and Automation (ICRA)*. IEEE. 2019.

## Honors and Awards

---

1st Place in Robust Vision Challenge - Semantic Segmentation Track (ECCV 2022)	2022
Best Paper Honorable Mention - DART (MICCAI Workshop)	2022
Academic Excellence Award	2016, 2017, 2018
JJWorld Scholarship	2017
National Scholarship ( <b>Top 0.1%</b> national wide)	2016

## Talks

---

<b>ByteDance Seed</b> (Topic: Captain Cinema: Towards Short Movie Generation)	May, 2025
<b>ByteDance Seed</b> (Topic: VideoAuteur: Towards Long Narrative Video Generation)	Jan, 2025
<b>JHU Computer Vision Seminar</b> (Topic: ProLab: Vision Semantic Taxonomy in the LLM Era)	Oct, 2024
<b>Google Research Vision Reading Group</b> (Topic: PaLM2-VAdapter)	April, 2024
<b>Google Research</b> (Topic: PaLM2-VAdapter)	Feb, 2024
<b>MICCAI Workshop - DART</b> (Topic: CateNorm)	Sept, 2022
<b>JHU Computer Vision Seminar</b> (Topic: Learning from Temporal Gradient)	Mar, 2022
<b>JHU Computer Vision Seminar</b> (Topic: CateNorm)	Mar, 2021