

# LUXIN ZHANG

(+1)412-315-5820 · luxinz@alumni.cmu.edu · <https://luxinzhang-ai.github.io>

## EDUCATION

<b>Carnegie Mellon University</b> Master of Science in Computer Vision, School of Computer Science	Aug. 2018 - Dec. 2019 <i>Pittsburgh, PA</i>
<b>Peking University</b> Bachelor of Science in Intelligence Science, School of EECS	Sept. 2014 - July 2018 <i>Beijing, China</i>

## RESEARCH KEYWORDS

Multimodal LLMs (VLM), Agentic Reasoning, Post-training (SFT/RLHF), Video Generation, Foundation Models, Large-scale Data Curation, Evaluation & Benchmarking.

## EMPLOYMENT

<b>Stealth AI Startup (founded by Eric Schmidt &amp; Sebastian Thrun)</b> Member of Technical Staff	Jun. 2025 - Present <i>San Francisco, CA</i>
<ul style="list-style-type: none"><li><b>Multimodal Reasoning &amp; Personalization:</b> Architected a pipeline converting raw visual/audio signals into structured text for agentic reasoning, enabling personalized AI behavior based on semantic understanding of user content.</li><li><b>Agentic Framework Development:</b> Developed a high-level agentic system utilizing LLMs to perform complex, multi-step video editing tasks, bridging the gap between model reasoning and tool-based execution.</li><li><b>LLM Post-training &amp; Data Flywheel:</b> Engineered synthetic data pipelines for SFT to align model reasoning and outputs with complex user behaviors and personalized instructions.</li><li><b>Evaluation &amp; Benchmarking:</b> Architected comprehensive evaluation suites and automated benchmarks to quantify the performance of modules and agentic system; established rigorous baselines to drive iterative model improvements.</li></ul>	

<b>Meta</b> <i>Superintelligence Labs, Senior Research Engineer</i>	Feb. 2023 - Jun. 2025 <i>Menlo Park, CA</i>
--	--

- Research:
  - **Media Foundation Models:** Core contributor to ([Movie Gen](#), [Emu](#), [Emu-Video](#)); specialized in large-scale data curation and post-training, involving the development of automated filters to understand and rank video quality/aesthetics for training.
  - **Multimodal Llama (VLM):** Developed multimodal image generation capabilities for Llama ([Llama 4](#)), focusing on the intersection of LLM reasoning and visual synthesis.
- Product: Deployment of the foundation models into production ([Meta AI](#), Ads). Led the launch of the first video generation model into production at Meta.

<b>Meta</b> <i>Reality Lab, Research Engineer</i>	Mar. 2020 - Feb. 2023 <i>Cambridge, MA</i>
--	---

- Research: Multitask learning with auxiliary signals for Ads recommendation.
- Product & Infra: AR contents recommendation on Instagram and VR/MR glasses.

<b>Meta</b> <i>Core Infra, Engineering Intern</i>	May 2019 - Aug. 2019 <i>Cambridge, MA</i>
--	--

- Product & Infra: Internal tool of Meta's deployment services.

**Microsoft Research Asia**

*Research Intern*

Sept. 2017 - Feb. 2018

*Beijing, China*

- Research: Machine learning for multilingual language understanding.
- Product & Infra: Developed LUIS, an open-source learning-based service for personalized language understanding.

**University of Texas at Austin**

*Research Assistant*

July 2017 - Sept. 2017

*Austin, TX*

- Research: Imitation learning of human attention for visuomotor tasks.

---

## PUBLICATIONS

- **The Llama 4 Herd: The Beginning of A New Era of Natively Multimodal AI Innovation**  
*Meta Technical Report*, 2025.  
**Luxin Zhang** as Core Contributor, Multimodal Generation
- **Movie Gen: A Cast of Media Foundation Models**  
*Meta Technical Report*, 2024.  
**Luxin Zhang** as Core Contributor, The Movie Gen team
- **MoCha: Towards Movie-Grade Talking Character Generation**  
*Conference on Neural Information Processing Systems (NeurIPS spotlight)*, 2025.  
Cong Wei, Bo Sun, Haoyu Ma, Ji Hou, Felix Juefei-Xu, Zecheng He, Xiaoliang Dai, **Luxin Zhang**, Kunpeng Li, Tingbo Hou, Animesh Sinha, Peter Vajda, Wenhui Chen
- **AVID: Any-Length Video Inpainting with Diffusion Model**  
*Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.  
Zhixing Zhang, Bichen Wu, Xiaoyan Wang, Yaqiao Luo, **Luxin Zhang**, Yinan Zhao, Peter Vajda, Dimitris Metaxas, Licheng Yu
- **Animated Stickers: Bringing Stickers to Life with Video Diffusion**  
*arXiv Preprint*, 2024.  
David Yan, Winnie Zhang, **Luxin Zhang**, Anmol Kalia, Dingkang Wang, Ankit Ramchandani, Miao Liu, Albert Pumarola, Edgar Schoenfeld, Elliot Blanchard, Krishna Narni, Yaqiao Luo, Lawrence Chen, Guan Pang, Ali Thabet, Peter Vajda, Amy Bearman, Licheng Yu
- **Cloth Region Segmentation for Robust Grasp Selection**  
*International Conference on Intelligent Robots and Systems (IROS)*, 2020.  
Jianing Qian, Thomas Weng, **Luxin Zhang**, Brian Okorn, David Held
- **Atari-HEAD: Atari Human Eye-Tracking and Demonstration Dataset**  
*Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 2020.  
Ruohan Zhang, Calen Walshe, Zhuode Liu, Lin Guan, Karl Muller, Jake Whritner, **Luxin Zhang**, Mary Hayhoe, Dana Ballard

- **Modelling Complex Perception-Action Choices**  
*Journal of Vision*, 2018.  
Ruohan Zhang, Jake Whritner, Zhuode Liu, **Luxin Zhang**, Karl Muller, Mary Hayhoe, Dana Ballard
- **Learning Attention Model from Human for Visuomotor Tasks**  
*Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 2018.  
**Luxin Zhang**, Ruohan Zhang, Zhuode Liu, Mary Hayhoe, Dana Ballard
- **AGIL: Learning Attention from Human for Visuomotor Tasks**  
*Proceedings of the European Conference on Computer Vision (ECCV)*, 2018.  
Ruohan Zhang, Zhuode Liu, **Luxin Zhang**, Jake A Whritner, Karl S Muller, Mary M Hayhoe, Dana H Ballard
- **Visual Attention Guided Deep Imitation Learning**  
*NIPS Cognitively Informed Artificial Intelligence Workshop*, 2017.  
Ruohan Zhang, Zhuode Liu, **Luxin Zhang**, Karl S Muller Mary M Hayhoe, Dana H Ballard

## ACADEMIC SERVICE

---

Program Committee/Reviewer for the following conferences/workshops.

- NeurIPS 2022 Workshop: Medical Imaging Meets NeurIPS
- ICML 2022 Workshop: Interpretable Machine Learning in Healthcare
- MICCAI 2022 Workshop: Medical Optical Imaging and Virtual Microscopy Image Analysis
- ICCV 2021 Workshop: Computer Vision for Automated Medical Diagnosis
- ICML 2021 Workshop: Interpretable Machine Learning in Healthcare
- ICML 2021 Workshop: Self-Supervised Learning for Reasoning and Perception
- ICML 2021 Workshop: Computational Approaches to Mental Health
- IJCAI 2021 Workshop: Weakly Supervised Representation Learning
- IJCAI 2021 Workshop: Long-Tailed Distribution Learning
- 2021 IEEE/CIC International Conference on Communications in China (ICCC)
- 2021 IEEE International Conference on Microwaves, Antennas, Communications and Electronic Systems (COMCAS)

## SKILLS

---

- **Programming:** Python, C/C++, C#, MATLAB, SQL, PHP, JavaScript
- **Platforms & Tools:** PyTorch, Cursor, Claude Code, Hugging Face, LangChain, verl, Wandb, Spark, Linux, Git, L<sup>A</sup>T<sub>E</sub>X