

Xi Fang

Machine Learning & Research Engineer | Computer Vision | Embodied AI

Cupertino, CA • (518) 960-2197 • xifang96@gmail.com • fang-xi.com • LinkedIn • GitHub • Google Scholar

SUMMARY

Machine Learning / Research Engineer with a Ph.D. in Biomedical Engineering and 3 years of industry experience building production computer vision and multimodal ML systems for medical robotics. Hands-on experience spanning VLA post-training, sim-to-real robotic manipulation, 3D / multimodal registration, segmentation, human-in-the-loop data pipelines, and TensorRT deployment. Published at MedIA, TMI, and MICCAI, with a track record of translating research algorithms into deployed robotic and medical imaging products.

EXPERIENCE

Machine Learning Engineer — Noah Medical, San Carlos, CA Dec 2023 – Present

- Built an AI-assisted fluoroscopy annotation and human-in-the-loop refinement pipeline that cut per-video labeling time by **83%** (12h → 2h), then trained and deployed an nnU-Net for metal artifact reduction using the resulting scalable dataset.
- Led development and production deployment of an airway segmentation system, from annotation / QA strategy through customized U-Net training and TensorRT inference; replaced a third-party API, achieved **>95% overall recall**, extended prediction into deeper airway branches, and reduced segmentation-related complaints by **90%**.
- Led multimodal registration between pre-operative CT and intra-operative tomography, achieving **~2.3 mm median error** across **7** different C-Arm types and integrating the optimized model into the core production C++ system to improve lesion visibility for real-time image-guided biopsy.
- Evaluated forward-looking ML applications with product leadership, including AI agents for complaint-investigation automation, learned world models for tomographic image quality, and VLA-based robotic autonomy.

Immersive Media Research Intern — Tencent America, Immersive Media Lab Feb 2022 – May 2022

- Built panoptic segmentation and scene-retrieval pipelines for Tencent's 3D Virtual Tour product using MSeg, panorama-to-perspective projection, unsupervised domain adaptation, and semantic-distribution retrieval across **100+ indoor and outdoor classes**.

Deep Learning Research Intern — Infervision, Beijing Mar 2017 – May 2017

- Adapted Faster R-CNN for lung-nodule detection in a clinical collaboration with Tongji Hospital, adding a dedicated false-positive class and achieving **90% detection accuracy** with reduced false positives.

SELECTED ROBOTICS & MACHINE LEARNING PROJECTS

Vision-Language-Action Post-Training & Sim-to-Real Robotic Manipulation May 2026 – Present

- Benchmarked Isaac GR00T-based VLA post-training with full fine-tuning and LoRA under DeepSpeed, tuning the embodiment-specific projector and diffusion module while freezing the LLM and vision encoder; achieved **90% simulation success** on LIBERO.
- Fine-tuned the best post-trained policy using **100** self-collected demos, then deployed it to an SO-101 robotic arm with TensorRT for sim-to-real pick-and-place and cover-the-cube execution, achieving **72% real-world task success**; documented results across CALVIN, LIBERO and an Open-H-Embodiment endoscope dataset.

Learning-Based Soft-Tissue Simulation & Surgical Planning May 2021 – Mar 2023

- Developed an attention-based correspondence model for facial soft-tissue simulation after bony movement, matching finite-element-method accuracy while reducing runtime from **5 hours to 1.2 minutes**.
- Developed a soft-tissue-driven planning framework that used the learned forward model to evaluate and optimize bony movement, enabling one of the first outcome-driven approaches to orthognathic surgical planning.

EDUCATION

Ph.D., Biomedical Engineering — Rensselaer Polytechnic Institute Aug 2018 – Nov 2023

Dissertation: *Deep Learning-Based Orthognathic Surgical Planning*

B.S., Computing — Wuhan University Sep 2014 – Jun 2018

SELECTED PUBLICATIONS & HONORS

Correspondence Attention for Facial Appearance Simulation. Fang X et al. *Medical Image Analysis*, 2024.

Soft-tissue Driven Craniomaxillofacial Surgical Planning. Fang X et al. *MICCAI*, 2023. **Early Acceptance (Top 13.6%).**

Deep Learning-Based Facial Appearance Simulation Driven by Surgically Planned Craniomaxillofacial Bony Movement. Fang X et al. *MICCAI*, 2022. **Oral; Young Scientist Award (Top 0.3%); Best of MICCAI (Top 1).**

Multi-organ Segmentation over Partially Labeled Datasets with Multi-scale Feature Abstraction. Fang X, Yan P. *IEEE Transactions on Medical Imaging*, 2020.

Additional publication recognition: NVIDIA Bench-to-Bedside Award Runner-up (2024); MICCAI Student Travel Award (2022). Full publication list available on Google Scholar.

SKILLS

Robotics / Multimodal: Vision-Language-Action Models, Embodied AI, Multimodal Learning, Post-training, Sim-to-Real

Computer Vision: Segmentation, Object Detection, Tracking, Action / Phase Recognition, Semantic Correspondence, Multimodal Registration

ML Systems: PyTorch, Python, TensorRT, DeepSpeed, Transformers, Self-Supervised Learning