Ding **Xia**

Doctoral Student

UI Lab, IST, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-0033, Tokyo, Japan

(+81) 070-4035-0449 · (+86) 156-9200-5240 | dingxia1995@gmail.com | apisxia.github.io | | | |

Research Interests

HCI & Applied AI	AI-driven interface design for human-vehicle interaction, Generative AI for design automation, Interactive fabrication systems
Multimodal AI	Vision-language model applications, AI agents for embodied interaction, Multimodal knowledge augmentation
Computer Vision	3D scene understanding, Medical image analysis, Visual representation learning

Technical Skills

Programming	Python (advanced), C# (Unity3D), LaTeX, Shell scripting
ML/AI Frameworks	PyTorch, Hugging Face Transformers, LLaMA Factory, OpenAI/Anthropic APIs, LangChain, CLIP, FAISS
Computer Vision & 3D	OpenCV, PIL, scikit-image, NeRF, Point Clouds, 3D Reconstruction
Tools & Infrastructure	Git/GitHub, Docker, Linux/Unix, Pandas, Matplotlib

Education

The University of Tokyo — *Tokyo, Japan*

PhD in Creative Informatics, Department of Creative informatics, IST — *Apr. 2022 - Present (expected Sept. 2026)*

- Research focus on Large Language Models (LLMs) and Vision-Language Models (VLMs) applications for external Human-Machine Interface (eHMI) in autonomous vehicles and robotics, supervised by Prof. Takeo Igarashi
- Developed novel framework using LLMs as eHMI control action designers to solve adaptation challenges in complex real-world scenarios, implemented using Python and Blender with state-of-the-art language and vision models
- Pioneered the introduction of LLMs to eHMI research domain, facilitating practical deployment of eHMI systems in real-world applications.
- Collaborated closely with Xinyue Gui, Xi Yang, and Tsukada Lab for interdisciplinary research in LLM usage, evaluation, and fine-tuning

South China University of Technology — *Guangzhou, China*

Master of Engineering, School of Automation Science and Engineering — *Sept. 2017 - Jun. 2020*

- Graduation Thesis: Semi-supervised Classification of Brain Signals Based on VAE Framework, supervised by Prof. Zhenghui Gu
- Enhanced VAE framework with novel loss function for P300 signal classification in brain-computer interface applications, implemented using Python and TensorFlow
- Achieved improved performance over baseline by addressing limited labeled data challenges through semi-supervised learning approach
- Key Courses: Digital Signal Processing, Pattern Recognition, Computer Vision, Convex Optimization

South China University of Technology — *Guangzhou, China*

Bachelor's Degree of Engineering, School of Automation Science and Engineering — *Sept. 2013 - Jun. 2017*

- GPA: 3.65/4.0 (ranking 2/17), Outstanding Defense Award for graduation thesis
- Graduation Thesis: Research on Single Image Depth Estimation Algorithm Based on Probability Graph Model, supervised by Prof. Zhuliang Yu, implemented using MATLAB with superpixel segmentation and regression analysis
- Served as department leader in Student Union, developing leadership and organizational skills
- Key Courses: Calculus, C++, Signal Processing and Analysis

Work Experience

CyberAgent AI Lab — *Tokyo, Japan*

Research Intern — *March 2024 - Oct. 2025*

- Conducted research on color design systems for vector graphic documents under the supervision of Dr. Naoto Inoue
- Implemented and accelerated internal tools for palette-based photo recoloring (based on ACM SIGGRAPH 2015 paper) using Python, with tools later utilized in subsequent research papers and product development
- Developed "ColorGPT" framework using LangChain, advanced prompt engineering techniques, and explored fine-tuning approaches for LLM-based color recommendation systems
- Achieved superior performance with experimental results demonstrating that the LLM-based pipeline outperformed existing methods in color suggestion accuracy and color palette completion tasks
- Published research findings at ICDAR 2025 (poster presentation), demonstrating the effectiveness of LLM-based approaches for color recommendation in design workflows

Corpy & Co. *Tokyo, Japan*

Software Engineering Intern *Sept. 2023 - Feb. 2024*

- Collaborated with Shuangshuang Alice Rao on Standardization of Work Through Task Analysis Project, developing annotation tools, conducting model training, and implementing research paper methodologies
- Developed custom eye gaze tracking annotation tools by modifying Labelme using Python
- Performed extensive data annotation for task-specific industrial parts and eye gaze patterns of factory workers, supporting computer vision applications in object recognition, mask segmentation, and gaze tracking
- Delivered factory task analysis system prototype to customers as project outcome

The University of Tokyo *Tokyo, Japan*

Research Assistant *Nov. 2020 - Mar. 2023*

- Conducted medical image registration and computer graphics research under Prof. Takeo Igarashi and collaborated with Prof. Kin Taichi
- Developed patch-based preregistration method for multi-modality 3D medical images (CT and 3DRA) using Python/PyTorch, achieving improved accuracy compared to existing methods by addressing varying image sizes through standardized patch processing
- Created end-to-end DICOM processing pipeline facilitating clinical image registration workflows
- Published research findings at MICCAI 2022, demonstrating reliable preregistration solutions for clinical applications
- Designed interactive desktop application using Unity/C# for ongoing research on deformable registration between 2D intraoperative images and 3D brain models, with PyTorch-based ML components and Unity-based visualization

Scholarships & Awards

2022-2025	**SPRING GX Fellowship**, Fostering Advanced Human Resources to Lead Green Transformation (GX)
2017	**Outstanding Defense Award**, Bachelor's Graduation Thesis Defense
2016	**First Prize (top 8%)**, China Undergraduate Mathematical Contest in Modeling

Publications

See2Refine: Vision-Language Feedback Improves LLM-Based eHMI Action Designers

Submitted to ACL Rolling Review (ARR), January 2026 cycle

Xia D (first author), et al.

Under review

Peeking Ahead of the Field Study: Exploring VLM Personas as Support Tools for Embodied Studies in HCI

Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems

Gui X, **Xia D**, Colley M, Li Y, Chauhan V, Anubhav A, Zhou Z, Javanmardi E, Seo SH, Chang CM, Tsukada M, Igarashi T

2026

Don't Worry, Just Follow Me: Prototyping and In-the-Wild Evaluation of Smart Pole Interaction Unit with Mobility

Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems

Chauhan V, Anubhav A, Colley M, Chang CM, Gui X, **Xia D**, Javanmardi E, Igarashi T, Fujiwara K, Tsukada M

2026

EssayBench: Evaluating Large Language Models in Multi-Genre Chinese Essay Writing

Proceedings of the AAAI Conference on Artificial Intelligence

Gao F, Li D, **Xia D**, Mi F, Wang Y, Shang L, Wang B

2026

TailCue: Exploring Animal-inspired Robotic Tail for Automated Vehicles Interaction

Proceedings of the 13th International Conference on Human-Agent Interaction

Li Y*, Gui X*, **Xia D**, Colley M, Igarashi T (*co-first authors)

2025

Automating eHMI Action Design with LLMs for Automated Vehicle Communication

Findings of the Conference on Empirical Methods in Natural Language Processing

Xia D, Gui X, Gao F, Li D, Colley M, Igarashi T

2025

HealthGenie: Empowering Users with Healthy Dietary Guidance through Knowledge Graph and Large Language Models

Gao F, Zhao X, **Xia D**, Zhou Z, Yang R, Lu J, Jiang H, Park C, Li I

Proceedings of the ACM International Conference on Information and Knowledge Management Demo Track

2025

ColorGPT: Leveraging Large Language Models for Multimodal Color Recommendation

Xia D, Inoue N, Qiu Q, Kikuchi K

Proceedings of the International Conference on Document Analysis and Recognition

2025

Draw2Cut: Direct On-Material Annotations for CNC Milling

Gui X*, **Xia D***, Gao W, Dogan MD, Larsson M, Igarashi T (*co-first authors)

Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems

2025: 1-17

PairingNet: A Learning-based Pair-searching and-matching Network for Image Fragments

Zhou R, **Xia D**, Zhang Y, Pang H, Yang X, Li C

European Conference on Computer Vision

2024: 234-251

SpaceEditing: A Latent Space Editing Interface for Integrating Human Knowledge into Deep Neural Networks

Wei J, **Xia D**, Xie H, Chang CM, Li C, Yang X

Proceedings of the 29th International Conference on Intelligent User Interfaces

2024: 489-503

A Two-Step Surface-Based 3D Deep Learning Pipeline for Segmentation of Intracranial Aneurysms

Yang X, **Xia D**, Kin T, Igarashi T

Computational Visual Media

2023, 9(1): 57-69

Data-Driven Multi-modal Partial Medical Image Preregistration by Template Space Patch Mapping

Xia D, Yang X, Van Kaick O, Kin T, Igarashi T

International Conference on Medical Image Computing and Computer-Assisted Intervention

2022: 259-268

Intra: 3D Intracranial Aneurysm Dataset for Deep Learning

Yang X, **Xia D**, Kin T, Igarashi T

Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition

2020: 2656-2666 (Oral)