

Zhicheng Du 杜知城

Ph.D. Candidate, Tsinghua University

Foundation Model Algorithms · Multimodal Learning · AI for Science

duzc21@mails.tsinghua.edu.cn | Shenzhen | dodoxb.github.io

Research Overview

With an undergraduate background in computing, I study foundation-model representations and inference efficiency, multimodal learning, and interdisciplinary applications. I combine domain knowledge with clinical and business needs to formulate predictive tasks, data representations, and experimental protocols. My experience spans algorithm research, cross-institutional collaboration, and model deployment.

- **Publications:** 35 published or accepted papers and articles, including 18 as first or co-first author.
- **Representative venues:** Science, ICLR, ICASSP, ICME, PRICAI, IEEE Communications Surveys & Tutorials, AACL-IJCNLP MultiPsyche Workshop, npj Digital Medicine, Chemical Engineering Journal, and Angewandte Chemie International Edition, among others.
- **Academic service:** Reviewer for 29 journals and 17 conferences/workshops across computing, medicine, and engineering; CCF Deadlines Organizer.

Education and Research Appointments

Tsinghua University | Ph.D. in Precision Medicine and Public Health (integrated program) 2021.09–present

Shenzhen International Graduate School; master's phase 2021–2023, doctoral phase since 2023.09; advisor: Prof. Lan Ma; GPA: 3.95/4.00

Planned early graduation: 2027.06 (maximum enrollment period through 2028.06)

Doctoral research: multimodal AI for uveal melanoma diagnosis and clinical characterization.

Xiamen University | Bachelor's degree in Cyberspace Security

2017.09–2021.06

School of Informatics; mentors: Prof. Zhirong Shen and Prof. Xiongbiao Luo

Shenzhen Bay Laboratory | Visiting Student

2025.09–present

Supervisor: Prof. Lan Ma

Foundation Models and Generative Algorithms

Continuous language diffusion: decoder diagnostics and efficient inference

First author · 2026 preprint

- Study how continuous latent states can be decoded reliably into text. Develop diagnostics for semantic recoverability, decoder compatibility, and generation trajectories to expose failure modes missed by scalar measures such as reconstruction error and perplexity.
- On generated states from public ELF models, a linear readout achieves **97.9%** agreement with native decoder decisions. Decoder-margin early exit reduces denoising steps by approximately **17%–28%** under held-out gates.

AudioCache: training-free acceleration of DiT audio generation

Co-first author · ICASSP 2025

- Analyze changes in diffusion Transformer layers across sampling steps. Design adaptive caching to reuse attention and feed-forward computations, accelerating generation without additional training.
- Achieve **2.35× acceleration** on Stable Audio Open while broadly preserving objective and subjective generation quality. This work was completed during my Huawei internship.

- **LAMPER:** Evaluate how prompting, sequence segmentation, and input length affect pretrained language models' zero-shot representations across 128 univariate time-series datasets in the UCR archive.
- **Cognitive Resilience:** Evaluate image-captioning models under different masking ratios, regions, and strategies, and analyze the relationship between salient visual regions and generated descriptions.

Biomedical and Interdisciplinary Research Collaborations

Sun Yat-sen University Cancer Center | Colorectal cancer biomarkers and risk stratification 2023.06–2025.09

First author; multimodal learning, explainable prediction, and clinical task modeling

- Develop multiscale multimodal fusion of whole-slide pathology, 3D CT, and clinical records for biomarker-status and mortality-risk prediction. Relevant models achieved over 90% accuracy in interim experiments.
- Investigate feature fusion and interpretability across clinical endpoints, exploring digital tumor biomarkers to support research on patient risk stratification and clinical decisions.
- Interim results appeared in The Lancet Regional Health–Western Pacific (explainable ensemble learning from clinical records) and Clinical Cancer Research (tri-modal contrastive learning with pathology, CT, and text).

Sir Run Run Shaw Hospital, Zhejiang University School of Medicine | Gallbladder cancer recurrence 2023.12–2025.09

First author; fusion of medical images and clinical text

- Integrate MRI, CT, and clinical records to develop multimodal fusion models for postoperative recurrence-risk prediction in gallbladder cancer. Interim experiments achieved **88% accuracy**.
- Optimize models for recurrence-risk stratification and postoperative follow-up research, gaining experience in joint modeling of imaging information and clinical endpoints.

Zhongshan Hospital, Fudan University | Liquid biopsy and multi-cancer early screening 2025.01–2025.09

First author; mass-spectrometry data, explainable deep learning, and symbolic learning

- Study early-screening and classification algorithms using laser desorption/ionization mass-spectrometry liquid-biopsy data from **seven cancer types**.
- Explore explainable deep learning, feature fusion, and symbolic learning to turn mass-spectrometry signals into useful representations for classification and analysis, gaining experience with non-image scientific data.

Multi-center collaboration | Interactive robots for adolescent mental-health screening First author · MAPS accepted screening

Partners: Shenzhen Kangning Hospital, Nanyang Fourth People's Hospital, and Lishui Central Hospital

- Use portable interactive robots to integrate images, speech, physiological measurements, and text into a multi-center multimodal dataset. Develop prediction models addressing **12 common adolescent mental-health screening targets**.
- Use complementary multimodal information for early screening, investigate comorbidity relationships and model interpretability, and incorporate affective and behavioral information into medical prediction tasks.

Project lead; independently developed the main algorithms using clinical data provided by collaborators

- **Doctoral research:** Use uveal melanoma as the central clinical setting to study multimodal representations, feature fusion, and explainable prediction from medical images and clinical information, organizing my doctoral work around diagnostic support and clinical-indicator modeling.
- Combine fundus images, OCT, and clinical text for tumor identification, malignancy assessment, and subtype prediction, developing an assistance system for rapid screening in primary-care settings.
- Further integrate ultrasound imaging and clinical follow-up text to study subsequent treatment-outcome prediction in UM, extending multimodal modeling from screening to follow-up.

Other Interdisciplinary Research

Contribute to ophthalmic image recognition, wearable sensing, biosensing, and materials research through algorithm experiments, data processing and analysis, and manuscript collaboration. Publications include npj Digital Medicine, Chemical Engineering Journal (co-first author), Exploration, Energy Storage Materials, and Angewandte Chemie International Edition. Articles in Science on scientific and societal issues have also developed my experience in interdisciplinary literature analysis, research communication, and English writing.

Industry Internships and Collaborations

Huawei 2012 Labs | Algorithm Research Intern

2024.06–2024.12

Research multimodal algorithm innovation and audio generation acceleration, integrating research methods with application models and contributing to outputs including AudioCache (ICASSP 2025, co-first author).

CRRC Nanjing Puzhen Co., Ltd., Digitalization Department | Research Intern

2025.06–2025.08

Shenshui Yunke (Infore Group) | Repayment intent, conversational affect, and behavior prediction

2025.10–present

- Independently lead the main algorithm development, fusing text and speech features to achieve **AUC 0.85–0.90** on internally labeled data for repayment-intent and conversational-response prediction.
- Use model outputs to improve communication scripts and follow-up strategies, prioritizing resources for cases with higher repayment intent and supporting adjustments to communication and operational investment.

Shenshui Yunke | Robot-generated and synthetic-speech detection

2025.10–present

- Develop synthetic data and detection algorithms to distinguish robot-generated or synthetic speech from human responses, enabling early identification of non-human conversations and reducing unnecessary follow-up.

Engineering Deployment and Application Development

Guokexin medical algorithms | Lingsheng supercomputing and Huawei Kunpeng CPU adaptation

2026

- Port and deliver offline inference packages for medical algorithms and ophthalmic imaging models on the Lingsheng supercomputing platform with Huawei Kunpeng CPUs (ARM64), adapting dependencies for knowledge question answering, domain-specific foundation models, and image analysis.
- Implement NUMA-aware CPU affinity, exclusion of management cores, and thread configuration. Prepare PyTorch adaptation workflows using HPCKit and Kunpeng Extension for PyTorch (KPEX), with model weights, offline dependencies, and installation and evaluation scripts.

- Develop a bilingual Chinese/English website with React, TypeScript, and Python, aggregating **over 63,000 call-for-papers records from 13 publishers and 15 disciplines**, with search, filters, countdowns, and linked SJR metrics.
- Build collection, deduplication, indexing, and presentation workflows. Use monthly sharding and lazy loading to improve access to large datasets and enable sharing of filtered views through links.

Technical Skills

Algorithms: Python, PyTorch, Transformers, LoRA/PEFT; language and multimodal representations, diffusion models, model evaluation, and explainable learning.

Engineering and applications: C/C++, Linux, ONNX Runtime, offline inference, and dependency adaptation; React, TypeScript, and Vite; data collection and processing, and research writing with $\text{\LaTeX}$.

Academic Service and Honors

Journal reviewing (29): Including npj Digital Medicine, Signal Transduction and Targeted Therapy, Journal of Hematology & Oncology, IEEE JBHI, and ACM Transactions on Computing for Healthcare.

Conference/workshop reviewing (17): Including NeurIPS, AAAI, IJCAI, ACM MM, BMVC, ICME, ICIP, IJCNN, PRICAI, CVPR MULA, and NeurIPS ODI.

Community: CCF Deadlines Organizer/Contributor.

Honors: Friends of Tsinghua–BYD Scholarship (First Prize for Comprehensive Excellence, 2025); SIGS Comprehensive Excellence Scholarship, Second Prize (2022–2024); Bronze Poster Award, Nature Conferences: Advancing Health with AI (2024); Outstanding Innovation Award, Biopharmaceutical and Health Engineering Cup (1/15, 2022).

Selected Papers and Articles

Selected references; * denotes equal first authorship. The complete list is available on my homepage.

- [1] Qingyang Shi*, **Zhicheng Du***, Jiasheng Lu, et al. AudioCache: Accelerate Audio Generation With Training-Free Layer Caching. *ICASSP 2025*, pp. 1–5, 2025.
- [2] **Zhicheng Du**, Zhaotian Xie, Yan Tong, and Peiwu Qin. LAMPER: LanguAge Model and Prompt EngineeRing for zero-shot time series classification. *The Second Tiny Papers Track at ICLR 2024*. OpenReview.net, 2024.
- [3] **Zhicheng Du**, Zhaotian Xie, Huazhang Ying, Likun Zhang, and Peiwu Qin. Cognitive resilience: Unraveling the proficiency of image-captioning models to interpret masked visual content. *The Second Tiny Papers Track at ICLR 2024*. OpenReview.net, 2024.
- [4] **Zhicheng Du**, Changyue Liu, Wenji Xi, et al. Auditing 3D Shape Retrieval without Supervised Training Using Diffused Geodesic Moments. *PRICAI 2026*, to appear.
- [5] Yiran Wang, Jiasheng Lu, Jun Chen, Xinyu Zhang, Yingshan Liang, **Zhicheng Du**, Qingyang Shi, and Shao-Lun Huang. Content-Style Disentangled Audio Style Transfer via Diffusion Model. *ICME 2025*, pp. 1–6, 2025.
- [6] **Zhicheng Du** and Lan Ma. Continuous Language Diffusion as a Decoder-Interface Problem. arXiv:2606.08810, 2026.
- [7] Yang Liu, Keming Zhao, Lihui Luo, et al. Diagnosing pathologic myopia by identifying morphologic patterns using ultra widefield images with deep learning. *npj Digital Medicine*, 8(1):435, 2025.
- [8] Zhenglin Chen*, Jiaqi Yang*, **Zhicheng Du***, Jiansong Ji*, et al. A conductive and anti-freezing gelatin-PAA-based organic hydrogel (PC-OH) with high adhesion and self-healing activities for wearable electronics. *Chemical Engineering Journal*, 492:152465, 2024.
- [9] Zhaotian Xie, Xin He, Wentao Zhang, **Zhicheng Du**, et al. Unlocking Latent Titanium Active Centers Through the Synergistic Effect of Oxygen Vacancies and Boron Doping for Accelerated Sulfur Redox Reactions in Silicon-Sulfur Batteries. *Angewandte Chemie International Edition*, 65(17):e4505904, 2026.
- [10] Zhaotian Xie, Wentao Zhang, Xin He, Ziyao Gao, **Zhicheng Du**, et al. Crossing the capacity threshold in Si-S batteries through mud-crack electrodes. *Energy Storage Materials*, 75:104046, 2025.
- [11] **Zhicheng Du**, Changyue Liu, Zhaotian Xie, Hui-Yan Luo, Wentao Zhang, and Xin He. Address Arctic shipping crisis at COP30. *Science*, 388(6752):1143–1144, 2025.
- [12] **Zhicheng Du**, Hui-Yan Luo, Xuefei Li, et al. Performance of explainable ensemble learning for mortality risk stratification and multimodal biomarker prediction in colorectal cancer: a retrospective, database cohort study. *The Lancet Regional Health–Western Pacific*, 55(Suppl.1):101459, 2025.
- [13] **Zhicheng Du**, Hui-Yan Luo, Lijin Lian, Vijay Kumar. Pandey, Jiansong Ji, and Peiwu Qin. Abstract A049: Development of a tri-modal contrast learning model integrating pathology-text and CT-text for clinical oncology tasks. *Clinical Cancer Research*, 31(13_Supplement):A049, 2025.
- [14] Yu Zheng, Kezhi Wang, **Zhicheng Du**, Wenji Xi, Gang Yu, Jinbo Hou, Haonan Hu, Jie Zhang. Point Cloud for Wireless Communications: A Comprehensive Survey on Fundamentals, Techniques, and Applications. *IEEE Communications Surveys & Tutorials*, 2026.
- [15] **Zhicheng Du**, Yang Liu, Zheng Chen, et al. MAPS: Robot-Mediated Multimodal Mental Health Screening for Chinese Adolescents. *AACL-IJCNLP 2026 MultiPsyche Workshop*, accepted, to appear.
- [16] **Zhicheng Du**, Zhuo Deng, Lan Ma. OphJev: endpoint-aware System One decisions for ophthalmic imaging and clinical text. *medRxiv preprint*, doi:10.64898/2026.10.03.26364645, 2026.
- [17] Zheng Chen, **Zhicheng Du**, Haoxuan Li, Peiwu Qin. Syndrome, Synergy, and Safety: Structured Reasoning and Knowledge-Driven Alignment for TCM Prescription Generation. *arXiv preprint*, arXiv:2609.25755, 2026.
- [18] Zheng Chen, **Zhicheng Du**, Haoxuan Li, Yingshan Liang, Peiwu Qin. LingLan: An Advancing Traditional Chinese Medicine Diagnosis LLM with Multimodal Data. *arXiv preprint*, arXiv:2609.25715, 2026.
- [19] Yang Liu, Yibing Shen, Keming Zhao, Cenk Jiang, Zhenghang Qian, **Zhicheng Du**, et al. An Ultra-Widefield Swept-Source OCTA Dataset and a Polar-Gated Mamba Network for Retinal Vessel Segmentation. *arXiv preprint*, arXiv:2609.12574, 2026.