

Ze RONG

Homepage: <https://zerong7777-boop.github.io/> | Email: ZeRong7777@gmail.com

EDUCATION

Nantong University

09/2022-07/2026

- **Major: Computer Science**
- **Core Course:** *Operating Systems (95), Software Engineering (89), Computer Logic Design (85), Mathematical Modeling and Algorithm Implementation (Excellent), Java Programming Fundamentals (Excellent), University Physics B (II) (92)*

PUBLICATIONS

- Zhaoxin Wang, Wenwen Yang, Zhengyu Li, **Ze Rong**, Xing Wang, Jincong Han, Lei Ma*. *A 25-Year Retrospective of the Use of AI for Diagnosing Acute Stroke: A Systematic Review*. Journal of Medical Internet Research (JMIR), 2024.
- Zhaoxin Wang, Wenwen Yang, **Ze Rong**, Zhengyu Li, Xing Wang, Lei Ma*. *Multi-Scale Crossed Algorithm for Ultrasound Medical Image Segmentation Based on MSC-LSAM*. Journal of Data Acquisition and Processing, 2025.
- **Ze Rong**, Ziyue Zhao, Zhaoxin Wang, Lei Ma*. *FaRMamba: Frequency-based Learning and Reconstruction Aided Mamba for Medical Segmentation*. In Proceedings of the 32nd International Conference on Neural Information Processing (ICONIP 2025).
- **Ze Rong**, Xiaofeng Shen, Haoyang Qin, Yue Xu, Hongjun Li, Lei Ma*. *FIRM: Fusion-Injected Residual Memory Brings Token-Level Alignment to Unsupervised VI-ReID*. In Proceedings of the 17th Asian Conference on Machine Learning (ACML 2025).
- Xiaofeng Shen, **Ze Rong** †, Haoyang Qin, Yue Xu, Binhao Zhao, Lei Ma*. *SCORE: Soccer Curriculum for Ordinal Affect Recognition with Evidence*. International Conference on Multimedia Retrieval (ICMR 2026).
- **Ze Rong**, Xiaofeng Shen, Lei Ma*. *SpecRA: A Spectral Routing Adapter for Adaptive Multi-Transform Routing and Synergy in Medical Image Segmentation*. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2026). Under Review.
- Lei Ma, **Ze Rong** †, Xiaofeng Shen, Ting Li*. *FuzzyStateFormer: A Neuro-Fuzzy State-Aware Framework with Transformer Temporal Encoding for Reliable Non-Invasive Long-Sequence fMRI Semantic Decoding*. IEEE Transaction on Fuzzy System (IF=11.9). Under Review (Major Revision).
- **Ze Rong**, Xiaofeng Shen, Lei Ma*. *BASS: Brain-Aligned Semantic Scaffold for Temporally Structured fMRI-to-Text Decoding*. 34th ACM International Conference on Multimedia (ACMMM2026). Under Review.
- Xiaofeng Shen, **Ze Rong** †, Haoyang Qin, Yue Xu, Li Wang, Lei Ma*. *DART: Multimodal Deviation-Aware Reasoning for Grounded Tactical Explanation in Broadcast Soccer*. 34th ACM International Conference on Multimedia (ACMMM2026). Under Review.

RESEARCH EXPERIENCE

Project: FaRMamba: Frequency-based Learning and Reconstruction Aided Mamba for Medical Segmentation

Project Leader

- MSFM and SSRAE, establishing a scalable framework for frequency- and reconstruction-driven medical segmentation
- Established MRI acquisition and annotation protocols for Meniere's disease mouse cochlea, completing 583-image high-precision dataset with refined quality control.
- Proposed FaRMamba, a hybrid Vision Mamba framework integrating multi-scale frequency transforms (MSFM) and self-supervised reconstruction (SSRAE) to restore high-frequency textures and maintain spatial consistency.
- Developed an unsupervised enhancement pipeline based on NAFA-DIKNet (CN119762378A), enabling joint optimization of segmentation and reconstruction through DIP-DKP collaboration and MCMC-based sampling.
- Achieved +1.36–6.63 pp Dice/MIoU improvement across MRI and ultrasound datasets; validated superior structural recovery and boundary precision over CNN-Transformer hybrids.

- Published results at ICONIP 2025, supporting a patent application and establishing a reproducible small-animal MRI segmentation benchmark.

Project: Ophthalmic Disease Diagnosis Platform

12/2024-05/2025

Project Leader

- Defined end-to-end diagnostic workflow and metrics with clinicians, covering low-light enhancement, vessel segmentation, disease classification, and multimodal QA.
- Curated stratified datasets with long-tail labeling; analyzed color distributions (RGB/HSV) and vascular topology (centerline, bifurcation, curvature) to guide model design and augmentation.
- Developed topology-aware EfficientNet-V2 with HSV enhancement and Class-Specific Residual Attention (CSRA) for improved long-tail performance; integrated lightweight U-Net/FaRMamba for segmentation and Retinex-based enhancement.
- Delivered a full-stack Vue + ElementUI platform with RESTful backend supporting inference, visualization, and audit logging, achieving higher precision and recall in low-light, imbalanced clinical scenarios.

Project: FIRM: Fusion-Injected Residual Memory Brings Token-Level Alignment to Unsupervised VI-ReID

Project Leader

- Engineered the FIRM framework, integrating Vision–Semantic Prompt Fusion (VSPF) and Evolving Multi-view Cluster Memory (EMCM) to enable unsupervised cross-modal semantic alignment and identity consistency.
- Devised the VSPF module to inject CLIP- and LLM-derived semantic prompts into multiple Transformer layers for explicit token-level fusion and modality alignment.
- Constructed the EMCM mechanism employing optimal transport–guided clustering, confidence-gated updates, and periodic merge–split refinement to mitigate pseudo-label noise and memory fragmentation.
- Implemented OT-weighted contrastive optimization and multi-layer alignment regularization to jointly enhance cross-modal consistency and feature discriminability.
- Attained state-of-the-art results on SYSU-MM01 and RegDB benchmarks, with ablation analyses confirming the independent and synergistic efficacy of all components.

Project: MADE: Multimodal Affective Dynamics Estimation Framework for Segment-Level Team Affect Prediction

Project Leader

- Conceived the MADE framework, a vision-centric architecture for segment-level team affect prediction that integrates multi-cue representations to capture long-range emotional dynamics in full-match broadcasts.
- Formulated the Team Pose–Affect Encoder (TPAE) to generate permutation-invariant team embeddings with contextual conditioning and robust aggregation of multi-person pose features.
- Devised the Two-Stream Fusion (TSF) mechanism to align pose and facial representations within a shared latent space through reliability- and content-aware gating, variational bottlenecks, and entropy control.
- Established the Cross-Segment Emotion Dynamics (CSED) module to model event-biased temporal dependencies, enabling inertia-preserving yet event-sensitive affect trajectory estimation.
- Attained state-of-the-art accuracy on SoccerReplay-1988 and NSVA benchmarks, outperforming RGB, skeleton, and multimodal baselines while yielding interpretable team-level affective trends.

HONORS & AWARDS

- China Robot and Artificial Intelligence Competition (CRAIC) 2025 – **National Second Prize**
- China Robot and Artificial Intelligence Competition (CRAIC) 2024 – **National First Prize**
- Chinese Collegiate Computing Competition 2024 – **National Second Prize**