Yujia Zhang

Tel.: +86 18335613725 | Email: yujia.zhang@stu.njau.edu.cn | Personal webpage: https://zhangyujia.com

EDUCATION BACKGROUND

Nanjing Agricultural University (NJAU), Nanjing, China 09/2020 – 06/2024
B.Eng. in Artificial Intelligence

- **Overall Grade: 88.7 / 100** | **Rank: 2 / 39**
- **Relevant Courses**: Mathematical Analysis (96) | Probability Theory and Mathematical Statistics (95) | Data Structure (95) | Digital Circuit (96) | Algorithm Design and Analysis (91) | Computer Vision and Graphics & Image Processing (93) | Machine Learning Lab (95) | Artificial Intelligence (100) | Natural Language Processing Lab (96) | AI and Driverless Technology (96)

PUBLICATIONS

[1] **Zhang, Y.**; Zhai, Z.* et al. ResViT-Rice: A Deep Learning Model Combining Residual Module and Transformer Encoder for Accurate Detection of Rice Diseases. *Agriculture* **2023**, *13*, 1264. [Text] – **19** citations (Google Scholar, as of 2025-09-09)

WORK EXPERIENCE

Tsinghua University - Fuzhou Joint Research Institute of Data Technology Fuzhou, China
Research Intern (remote), Medical AI 11/2024 – 05/2025

Developed a 3D breast tumor segmentation pipeline based on MRI (Magnetic Resonance Imaging) for surgical planning.

- Redesigned model architecture by replacing the 3D U-Net backbone with a ConvNeXtV2-3D encoder. Leveraged large-kernel depthwise separable 3D convolutions to expand the effective receptive field under constrained GPU memory.
- Implemented data augmentation techniques (3D random flips, constrained rotations, elastic deformations) and optimized training with focal-Tversky + surface loss to handle extreme class imbalance, AdamW with weight decay, gradient clipping, linear warm-up with cosine decay, early stopping on IoU plateau, and EMA checkpointing.
- Optimized inference with sliding-window + Gaussian blending, test-time augmentation (multi-axis flips/rotations), and confidence-weighted ensembling; refined postprocessing with largest connected component filtering and morphological operations.
- The optimized pipeline passed internal clinical validation (mean IoU > 0.91) and received positive feedback from physicians.

Nanjing Agricultural University School of Artificial Intelligence Nanjing, China
Research Assistant 07/2024 – 07/2025

Partnered with a patent firm to streamline CNIPA office-action replies with an on-prem, Chinese-first assistant built on Llama 3.1. By prioritizing efficiency via parameter-efficient training, RAG (retrieval-augmented generation), and quantized deployment, we enabled a fast, upload-to-answer workflow that greatly reduced attorneys' workload.

- Pre-trained Llama 3.1 on a Chinese patent corpora crawled from Google Patents using QLoRA (Quantized Low-Rank Adapter), internalizing claim structures, legal phrasing, and long-form technical exposition.
- Built a low-burden submission system for staff to create issue-level QA pairs in Llama-3.1 prompt format; each record includes a CNIPA objection span, matching application span, and a reference answer, enabling reliable Supervised Fine-tuning (SFT).
- Implemented RLHF (reinforcement learning with human feedback) using DPO (Direct Preference Optimization) with pairwise human judgments to enhance evidential reasoning, stepwise rebuttal structure, and output faithfulness.
- Engineered an RAG pipeline that parsed office-action issues, retrieved relevant application passages, and assembled formatted prompts via LangChain, producing issue-level replies with references, improving factual grounding and reducing hallucinations.
- Deployed an optimized GGUF (Q4) model for offline inference and delivered a FastAPI backend coupled with a React SPA behind Nginx, enabling a fast upload-to-answer workflow with rapid turnaround.

RESEARCH EXPERIENCE

***INSPECT*: Crop Pest and Disease Identification based on Multi-modal and Few-Shot Learning** 10/2023 – 05/2024
NJAU Outstanding Undergraduate Thesis | Advisors: Prof. Junxian Huang and Prof. Shuxin Zhu, NJAU

Addressed the challenges in real-world crop-pest recognition tasks under severe background clutter, scarce samples, large intra-class variability, and high inter-class similarity by developing LIIP102, to our knowledge, the first large-scale image-text dataset for insects and pests, and a multi-modal framework INSPECT, surpassing strong unimodal models (~8%) in accuracy and CLIP variants while remaining robust in extreme few-shot settings.

- Constructed LIIP102 (Language-Image Insect Pest 102) dataset by generating high-quality natural language descriptions for 75,222 IP102 images using LLMs (GPT-4V, Gemini) with five-annotator manual review; further augmented the data using InstructPix2Pix to generate scene variants under different weather conditions (e.g., sunny, cloudy, or rainy)
- Engineered INSPECT (**I**mage-Text **N**etwork for **S**ystematic **P**est **C**lassification **T**ask), a cross-modal architecture that integrates a custom-designed CNN/Transformer hybrid image encoder (ConViT) with a bidirectional cross-attention multimodal encoder.
- Trained the model under dual supervision with image-text contrastive loss (ITC) for alignment and multimodal classification loss (CLS) for discrimination; compared multiple text encoders (BERT, RoBERTa, Sentence-T5) under different initialization strategies (CLIP-style vs. supervised) to identify the optimal model configuration.
- The model attained 79.38% accuracy on pre-training tasks, outperforming current SOTA unimodal baseline (ConvNeXtBase) by ~8%. In extreme few-shot settings (e.g., 6-shot) the model maintained 58.24% accuracy. Among various ablation studies, removing ITC loss led to a 5-pt drop in accuracy, which validated the align-then-fusion paradigm.
- Developed a React front end with Flask APIs to deploy INSPECT; incorporated Grad-CAM visualization to highlight biologically relevant regions, showing model interpretability and alignment with agronomic knowledge.

***ResViT-Rice*: A Hybrid CNN-Transformer Model for Accurate Rice Disease Detection** 10/2022 – 08/2023
Advisor: Prof. Zhaoyu Zhai, NJAU

Developed ResViT-Rice, a hybrid CNN-Transformer architecture that fuses CNN's spatial inductive bias with Transformer's long-range context, achieving 97.84% accuracy (+4.81% pts vs. strongest baseline) on rice disease classification.

- Proposed the hybrid architecture of ResViT-Rice, integrating bottleneck blocks with custom Transformer-based ResViT-Rice blocks. Each block projects channels (1×1), applies a lightweight Transformer encoder for long-range context, then fuses via residual connection and a 3×3 convolution to stabilize local texture features; a CBAM (Convolutional Block Attention Module) provides modest channel/spatial reweighting to emphasize lesion regions while suppressing background noise.
- Processed the Kaggle Rice Leaf Disease dataset (leaf blast, brown spot and healthy; 516 images per class) through Gaussian filtering, contrast/brightness/color augmentation, random rotation/translation/scaling/flipping; expanded each class to 1,648 training samples and 104 testing samples to ensure model generalizability and reduce overfitting.
- Benchmarked ResViT-Rice against strong CNN/Transformer baselines (e.g., ResNet, Swin Transformer) with metrics such as accuracy, precision, recall, F1-score, AUC, confusion matrices, and ROC curves. Achieved 97.84% mean accuracy, outperforming the strongest baseline by 4.81%.
- Conducted ablation studies to quantify the contribution of the proposed ResViT-Rice Block and CBAM. Built Grad-CAM to show attention concentrated on lesion regions aligning with expert views. Further demonstrated model robustness beyond the core dataset through experiments on a more complex multi-category dataset.

Measurement of Rapeseed Leaf Phenotypic Parameters from Images 02/2022 – 06/2023
Advisor: Prof. Shuxin Zhu, NJAU
Built an end-to-end pipeline that ingests single/batch leaf images and outputs leaf length/width, petiole length, curl grade, and notch grade in one-click, helping standardize field and lab phenotyping.
- Implemented Mask R-CNN to segment leaf blade and petiole; applied principal-axis pose normalization to make downstream measurements orientation-invariant; computed axis-aligned leaf length/width and petiole length from clean masks.
- Trained a lightweight CNN (MobileNetV2) with targeted data augmentation to classify leaf-curl grades, achieving 93% test accuracy while supporting low-FLOP on-device inference.
- Built a leaf-edge notch detection pipeline: used grayscale conversion, morphological operations for denoising, contour extraction, and polygonal approximation to obtain stable vertices; computed adjacent-segment turning angles via arctan2 to form 1D feature sequences and classified notch grade using a linear SVM.
- Developed a PyQt-based desktop GUI with visual overlays and CSV export to enable fast and reliable phenotyping workflows for agronomy research teams.

***Nong Cha*: A Self-service Beverage Vending Machine** 09/2021 – 06/2023
Advisor: Prof. Wei Lu, NJAU
Collaborated on the design and building of a fruit juice vending machine covering embedded on-device control, an on-site human-machine interface (HMI), and a cloud back end.
- Engineered the cloud back-end with Spring Boot 2, MyBatis, and MySQL; exposed REST APIs and implemented WeChat OAuth for login plus JWT-backed stateless API auth across HMI and mobile clients.
- Developed a .NET/C# WinForms supervisory HMI on an industrial PC; enabled real-time coordination over Modbus/RS-485, driving actuators for fruit retrieval, juicing, blending, sealing, and delivery with bidirectional state synchronization.
- Built a WeChat Mini Program (mobile client) for ordering, beverage configuration, order management, and user profiles using native Mini Program APIs.
- Contributed to the mechanical design and prototyping of the vending machine structure in SolidWorks.
- Successfully deployed a prototype on NJAU campus, which has served over 12,000 orders (~500 orders monthly) to date.

RoboCup China Open 2022 09/2021 – 12/2022
Advisor: Prof. Wei Lu, NJAU
Developed a dual-arm mushroom harvesting robot with real-time detection and 3D localization for autonomous picking. Led the design of the robot's perception pipeline and its integration with the grasp planning system
- Led the engineering of a YOLOv3-based perception module for mushroom detection and 3D localization by fusing RGB and RGB-D camera data.
- Integrated the detection module into a ROS pipeline to provide continuous feedback to the dual-arm grasp planner.
- Transformed perception outputs into the robot base frame to enable non-destructive grasping and re-boxing, supporting fully autonomous mushroom harvesting.

EXTRACURRICULAR ACTIVITIES

Student Art Troupe of Nanjing Agricultural University Pukou Campus 10/2022 – 06/2024
President
- Managed the art troupe's daily operations and budgets; oversaw rehearsal, stage design, and technical support for performances.
- Spearheaded the planning, coordination, and promotion of 6 on-campus cultural events and annual graduation galas.
- Organized touring performances in high schools across multiple cities, promoting the "Elegant Art on Campus" initiative and improving the university's public image.

HONORS & AWARDS

Nanjing Agricultural University Outstanding Undergraduate Thesis (**Top 1%**) 07/2024
Nanjing Agricultural University First-Class Scholarship (**Top 7%**) 2021, 2022, 2023
Ceaiya Scholarship (**Top 2% in the University**) 12/2020
Provincial Silver Award, China College Students "Internet+" Innovation and Entrepreneurship Competition 09/2024
National Third Prize, RoboCup China Open 2022 12/2022

SKILLS

Programming Languages: Python | C# | C++ | SQL | JavaScript
Tools & Frameworks: PyTorch | LangChain | Hugging Face | Git | Linux | SolidWorks
Languages: Mandarin (Native), English (TOEFL 106) | **Hobbies:** Parkour and Somersaults (front, back, spin, group side, etc.)