更改日志

20250909

1. Developed NeuralMap Auto-Label, an offboard variant of WeRide’s BEV perception model, incorporating future temporal and spatial context from multiple frames to automatically generate high-quality training annotations for autonomous driving, with training data of xxx frames, and evaluation dataset of xxx frames.
2. Upgraded model backbone, replacing an 11M-parameter ResNet with 22M-parameter ViT-family models (Tiny-ViT, PVTv2, Swin-T) and increasing hidden dimension from 256 → 768, achieving ~5% mIoU/precision@0.5 IoU improvement on road element segmentation.
3. Enhanced BEV transformer decoder by stacking 3 additional layers, leading to a x% accuracy gain in downstream perception tasks (lane and road marker detection).
4. Implemented multi-model uncertainty estimation (predictive entropy, mutual information, variance ratio, low-margin) in C++, automatically flagging high-uncertainty logs for review; scaled pipeline to 10,000+ videos/day on x8 L20 GPUs, cutting human annotation costs by ~$25k–$50k/month.

20250219

1. Refactor Layout

20250123

1. 添加personal website，remove youtube

20250107

1. 修改姓名字号from 20
2. 添加location
3. Polish by GPT with the prompt: First, imagine you are a staff-level software enginner in computer vision and robotics and the hiring manager. Polish this for an entry-level to be professional and concise on resume. Then, bold the highlight word that you feel the most awesome, highlights should not be too much

20250104

1. 修改经历时间，删除月份
2. Skillset moves to the bottom and make it white
3. Remove 每个experience下的单独行skills

|  |  |  |
| --- | --- | --- |
| **EDUCATION** |  | |
| **University of California, Berkeley** | M.ENG., EECS | GPA - 3.86/4 | *Berkeley, CA* | 2024-2025 |
| Advisor: Prof. Masayoshi Tomizuka, Dr. Wei Zhan  Capstone Project - Optimization of Modularized Robot Design in Complex Scenes.  Coursework - Computer Vision(Jitendra Malik), Deep Learning for Computer Vision, Introduction to Robotics, Data Science, Agentic LLMs | | |
| **Shanghai University** | B.ENG., Computer Science | Rank - Top 1 | GPA - 92.59/100 | *Shanghai, China* | 2020-2024 |
| **WORK EXPERIENCE** |  | |
| **WeRide** |Machine Learning Engineer | May 2025 - Present | |
| * Built perception and auto-labeling systems for L3/L4 **self driving**, covering onboard real-time models and large-scale offline pipelines * Upgraded Bird Eye View road marker image backbone from an 11M-parameter ResNet with 22M-parameter **Vision Transformer-family models** (Tiny-ViT, PVTv2, Swin-T), achieving ~5%mIoU(mean Intersection over Union) improvement on segmentation task * Built C++ uncertainty estimation for auto-labeling, scaling to **10K+ videos/day** on 8× NVIDIA L20 GPUs and **saving $30K/month** * Currently fine-tuning Qwen3 vision-language model (LoRA, 32× L20 GPUs) on 16K in-house annotated clips + 40K ChatGPT QAs, enabling a **GPT-style Q&A system for risky driving behaviors** across ~1M hours of video | | |
| **Momenta** |Software Engineer Intern | Jan 2024 - Jul 2024 | |
| * Delivered the first version of an **autopilot reversing** feature for **GM Cadillac LYRIQ** in parking lots, covering the **full** **product** **lifecycle** * Developed a **stuck-state** detection **benchmark** stack for reversing maneuver, integrating classification and precision-recall analysis * **Evaluated 800+** real parking test cases in 30+ garages over 4 months, enhancing collision-free reversing performance by 3% * Optimized threshold-based stuck-state detection by integrating **SVM**, refining borderline cases and reducing **false positives** by 15% * Slashed **workload** for 4 product managers across 4 vehicle projects from **120min/day to 10 min/day** by automating data processing and report generation for 36,000+/day simulation test records evaluation, streamlining success rate analysis | | |
| **DJI RoboMaster** | **[Git repo](https://github.com/SRM-Vision/SRM-Vision-2022)** | **[Video](https://www.youtube.com/watch?v=4uyBBJRXUTg)** | | |
| *Team Lead* | 2022-2024 | |
| * Led a 40-student team to deliver **8** fully functional **robots** **from scratch**, winning the **3rd place** in RoboMaster 2023 * Gained **US$20,000** in sponsorship by improving shooting, motion control, and detection through **7000+ test iterations** | | |
| *Co-Head of Robot Computer Vision* | 2020-2022 | |
| * Developed an adaptive **vision pipeline** for **real-time** auto aiming, integrating camera calibration, image preprocessing, YOLOv7 detection, object tracking, and fire control, earning **promotion** for code contributions and troubleshooting * Boosted detection accuracy by 25%, upgrading from pure OpenCV detection to an OpenCV + **YOLOv7** hybrid system * Initiated a **non-linear least squares** based tracking algorithm for planar rotating object in **parametric sinusoidal speed**, integrating pitch-yaw angle computation based on geometry for precise targeting, outperforming **90%** competitors in hit rate * Refactored the system from **Python to C++** with 5 peers, achieving **60 fps** on NVIDIA NX and **80 fps** on NVIDIA AGX | | |
| **FIRST Tech Challenge** | **[Team Documentary](https://www.youtube.com/watch?v=ZKn0rDUpNfY) |** Team Lead | 2017-2020 | |
| * Bent the performance of a 15-member team, leading to **2 FIRST World Championships admissions**(top 2 % out of 7500 teams globally), **1 Inspire Award**(1st out of 40 teams), and **3 Connect Awards**(top 8% out of 60 teams) * Developed an autonomous system using multiple sensors and **OpenCV**/**TensorFlow SDK**, achieving 95%+ detection accuracy and securing the highest Regional score with rule-based human driver imitating strategies | | |
| **RESEARCH & SIDE PROJECTS** | | |
| **FlexTok for Video Reconstruction** | **[Report](https://drive.google.com/file/d/18d3QyqxoFDTTRwvBRvx2NwESfzbkK1qc/view)** | **[Git repo](https://github.com/LIYunzhe1408/video-tokenizer)** |Instructed by Prof. Jitendra Malik & Prof. Angjoo Kanazawa | | |
| * Extended a 1D image tokenizer to the video domain, evaluating reconstruction on **exocentric** (UCF101), **egocentric** (EGTEA Gaze+), and **synthetic** (CATER) video datasets with **1.8K–3K+ clips** across 16–256 token lengths * Demonstrated **64 tokens** as an optimal trade-off, achieving up to **2× lower FVD (Fréchet Video Distance)** vs. **16 tokens** and preserving motion and spatial structure, while complex tasks required 256 tokens for full fidelity * Benchmarked against VidTok (video-native tokenizer): FlexTok-256 achieved **FVD 28.1** vs. VidTok’s **12.7**, highlighting strengths in spatial fidelity but limitations in temporal consistency | | |
| **MealMate: From Cravings to Carts** | **[Git repo](https://github.com/LIYunzhe1408/MealMate)** | **[Video](https://youtu.be/bAT-jZhDtCw?si=HPL83vIrPcu6HJY9)** | LLM, Flask, React.js, Python, HTML&CSS |  | |
| * Designed an LLM-powered assistant to generate tailored shopping lists using user preferences and real-time inventory * Benchmarked GPT-4o-mini against GPT-4 and GPT-3.5-turbo for LLM agent performance, demonstrating 20% higher recipe match precision and 40% suggestion accuracy, and 42% faster processing time * Built a showcase using React.js for the frontend and Flask for the backend with effective prompts and transaction logic | | |
| **Visual Explainer For Deep Learning Image Classification** | **[Git repo](https://github.com/LIYunzhe1408/Visual-Explainer-For-Deep-Learning-Image-Classification)** | **[Video](https://youtu.be/9Tyr53aP-r8?si=XDSroY1QU7EiMskY)** | | |
| * Designed a 2-stage **image segmentation** pipeline and an **AutoEncoder** with tree constraints, using Shapley Values to extract and rank concepts by importance, boosting explanation consistency score by 35% on 1000+ images from 20 **ImageNet** classes * Developed a **Django backend APIs** for page navigation, handling **GET** and **POST** requests, and efficient data retrieval * Built a **Vue.js** **frontend** with user authentication, image segmentation, and contribution heatmap visualization | | |
|  | | |