

# Shih-Yu Lai

CS GRADUATE STUDENT · AI RESEARCHER

✉ (+886) 965-513-573 | 📩 akinesia112@gmail.com | 📱 Akinesia112 | 📺 shih-yu-lai | 🎓 Shih-Yu, Lai

An AI/Robotics Scientist with International Research Experience in **ML/RL/Graph Learning** expertise, focusing on: **(1) Numerical Optimization in Graphics/Vision** (neural PDE solvers, Monte Carlo, differentiable physical simulation, **dynamics**(motion processing), VLMs based animation), **(2) Robotics** (Sim2Real, manipulation, sensing), and **(3) 2D/3D Geometry** (discrete differential geometry, digital twins).

## Summary

- Published **8 conference papers** in top-tier AI/CS conferences during my **undergraduate studies**.
- Thesis:** Neural Physics Solver in Complex Geometry Processing with Walk on Sphere-**Monte Carlo Method** for **Partial Differential Equations**.
- Co-worked** with researchers in diverse backgrounds from the **ETH Zurich, U.S. Dept. of Energy, UCLA, UCL, TU Delft, UOsaka**(Visual).

## Education

### National Taiwan University

**M.S. in Computer Science and Information Engineering (AI Program)**

Feb. 2025 – Expected Jun. 2027

- Ranked **#1** university in Taiwan, **QS 2026: #63**, Admission Rate = **3.96%**.
- Member @ Communications and Multimedia Lab | Advisor: Prof. Robin Bing-Yu Chen.

## Publications (Selected: 3/9)

- ACM SIGGRAPH'26 (Under Review)**, “EUPHORIA: Efficient Universal **Planning** via Hybrid Optimization for Robust **Industrial Robotic Assembly**” (**1st Author**, Technical Paper (Full)).
- ACM SIGGRAPH Asia'25**, “Architectural CAD/CAM in **Robotic Fabrication** via **Differentiable Energy Optimization** and Few-Shot Spatiotemporal Graph RL Policy” (**1st Author**, Technical Communications, Paper).
- ACM SIGGRAPH Asia'25**, “**Graph Q-Learning for Automatic Assembly** in Design-to-Construction” (**1st Author**, Poster, Paper).

## Work Experience

### RIKEN, Center for Computational Science @ Kobe/Osaka, Japan

*High-Performance Computing Research Intern | Large-Scale Digital Twin Team | Prof. Hirozumi Yamaguchi (UOsaka)*

Jul. 2025 – Aug. 2025

- Collaborated with international teams (**8+ nationalities**) from ETH Zurich, TU Delft, the U.S. Department of Energy, and others on high-performance computing researches, demonstrating strong cross-cultural communication and collaboration skills (Visual).
- Devised large-scale digital twin models in mobile computing, **multi-modal data-fusion sensing** with WiFi, point cloud, RGBD images, and high-performance computing on **supercomputer Fugaku**, leveraging distributed edge/cloud computing with graph federated learning foundation model and differential privacy as privacy-preserving mechanism for real/virtual human-scene interaction.

### Robot-Aided Creation & Construction Research Group

*Research Assistant | Synchronization Lab | Prof. Yang-Ting Shen, Prof. Chia-Ching Yen, Prof. Alvin Wen-Yu Su*

Apr. 2022 – Jun. 2024

- Collaborated with NVIDIA on robotic reinforcement learning using Kuka/UR10/Hiwin **manipulators** in Omniverse Isaac Gym/Lab, devising graph reinforcement learning for **geometric brick assembly** by mobile manipulators (Visual).
- Implemented robotic systems with ROS2, integrating **SLAM** for vehicle dynamics, supervised learning (YOLO, LSTM), and reinforcement learning for motion control, planning, and obstacle avoidance with **LiDAR, sensors, and map data**.

### Delta Electronics, Inc.

*Deep Learning Intern | Power and System Business Group*

Jul. 2023 – Aug. 2023

- Developed the robust reconstruction in a **generative 3D point cloud autoencoder** from 74k meshes via PyTorch3D in Docker on a V100 GPU (13 h training), enabling magnetic-core design by geometry-constraint solving and material analysis via finite-element methods.
- Optimized XGBoost to predict design parameters for inductance calculation in magnetic circuits using electromagnetic attributes from physical simulation, reducing Mean Absolute Error (MAE) to **1%** of the formula-based solution (Visual).

### Communications & Multimedia Lab

*Research Assistant | Interactive Graphics Lab | Prof. Robin Bing-Yu Chen*

Nov. 2024 – Aug. 2027

- Developed a **Differentiable Projective Dynamics** based hyper-elastic material mesh solver with positive-definite Hessian projection, achieving **1.65x–2.26x acceleration** by introducing absolute value and clamping methods for stable and efficient **physical simulation**, enabling realistic and physically accurate modeling (Visual).

### MoonShine Animation Studio

*Vision Language Models Research Intern | Research & Development Team | Dir. Shaune Jan & Xiang 'Anthony' Chen (UCLA)*

Sept. 2025 – Jun. 2026

- Researching efficient supervisor review with **Multi-Agents MLLM-as-a-judge** by Autogen and accelerating post-production workflows, developing stable automated tools that harness controllable learning-based models for seamless design-to-production pipelines.

## Honors & Awards

### IEEE Signal Processing Society

*2023 IoT Innovation Challenge*

Feb. 2023 – Jun. 2023

- Secured **3rd Place & Sponsors Award** (\$30,000) from NXP Semiconductors (Awards).
- Built a microcontroller-based **SensorHub** (ADI/NXP/Arduino) integrating DHT11/SCD40/PMS5003T via chips communication (I<sup>2</sup>C/UART); developed calibration tests with real-time logging in HVAC/fan/lighting control linked to Unity digital twin for **vision model** based crowd regulation.

# Software Projects

---

**Generative Modeling on Synthetic Data** | Github: [Generative-Modeling-Synthetic-Data](#) | Python, PyTorch

- Implemented and benchmarked **GAN**, **diffusion models (DDPM/DDIM)**, and **Mean Flows** for One-step diffusion on a checkerboard dataset with end-to-end training, sweep-based inference (#steps,  $\eta$ ), and comprehensive evaluation using **Energy Distance** and **2-Wasserstein Distance**.
- Conducted ablation studies on **timestep embeddings** (learned vs. sinusoidal) and analyzed the **stochasticity-fidelity trade-off** in DDIM via  $\eta$  sweeps; compared sampling efficiency and geometric alignment across methods; made visualizations of distribution dynamic evolution.

**Multi-Agent RL Communication** | Github: [MARL\\_Communication](#) | Python, PyTorch, MuJoCo, Gymnasium

- Implemented a scalable **MADDPG** baseline for **multi-agent communication** in **MuJoCo locomotion** with a shared actor, central mean-pooled critic, replay buffer, and logging of return, AUC, wall-clock, and memory, including automatic plotting and video for qualitative evaluation.
- Prototyped **Attention/Transformer/GNN-based** and **PDE-inspired** multi-agent architectures on top of the MADDPG framework to study coordination and communication, and evaluated generalization across different agent counts and tasks for training and cross-task/variable- $N$  evaluation.

**Deep Reinforcement Learning Tasks** | Github: [DRL-Assignment-1](#), [MiniGrid](#), [DRL-Assignment-2](#) | Python, Gym

- Implemented deep and tabular **RL agents** for dynamic grid-worlds, including a DRL policy for **OpenAI Gym's Taxi-v3** and Tabular Q-Learning & Policy Learning with reward shaping in **MiniGrid** using PyTorch. ([Visual 1-1](#); [Visual 1-2](#))
- Developed and evaluated exploration-exploitation and tree-search methods: compared multi-armed bandit strategies (Exploration-First,  $\epsilon$ -Greedy, UCB1); engineered a **UCT-based MCTS** with TD(0) n-tuple function approximation; and deployed a self-play PUCT-MCTS framework combining learned policy and value approximators to master the **2048 game**. ([Visual 2](#))

**Multimodal LLM Agents** | Github: [Artificial-Intelligence\\_Multimodal\\_Agent](#) | Python

- Built a BLIP+Phi-4 pipeline to caption MSCOCO-Test and Flickr30k with BLEU/ROUGE/METEOR and added Stable Diffusion 3-medium **text-to-image** & SD v1.5 **style transfer**. ([Visual 1](#)); Implemented a **RAG** system using Phi-2+all-MiniLM embeddings for resume QA and pytesseract+Phi-4 **captioning** for PDF indexing. ([Visual 2](#))
- Created a GPT-4o-mini/pyautogen **multi-agent** restaurant rater mapping adjectives to scores and aggregating via geometric mean. ([Visual 3](#))

**Security and Privacy of Machine Learning** | Github: [Security-and-Privacy-of-Machine-Learning-Critique](#) | Python, LaTeX

- Developed a **denoise-aware audio watermarking framework** integrating **DSP priors** and manifold learning, perturbation analysis, and lightweight neural detection to improve **robustness-efficiency trade-offs** in music provenance. Designed spectral masking strategies, adversarial attack evaluations, and **feature-importance interpretability** tools (waveform, spectrogram saliency, band-occlusion) for analyzing watermark resilience under compression, noise, filtering, and generative-model transformations.
- Performed weekly critiques and presentations of 2023–2025 research papers across the full spectrum of ML security and privacy — including adversarial and certified defenses, poisoning and backdoor attacks, **LLM and VLM jailbreaks, prompt injection**, RAG vulnerabilities, unlearning, immunization, membership inference, federated learning, memorization, and **multi-agent system** threats — emphasizing both attack vectors and defensive mechanisms.

**Collage Texture Generation via Clustered Style Transfer on 3D Model** | Github: [Visual-Symphonies](#) | Python, PyTorch

- Developed Visual Symphonies, a multimodal framework integrating collage fragments, feature clustering (InceptionV3 + GATCluster), and CycleGAN **style transfer** to enhance creative diversity and stylistic coherence in **texture generation**.
- Applied the synthesized textures to 3D model visualization, enabling computer-aided design, architectural mosaics, and digital craft pattern generation through controlled variations in color, geometry, and gradation ([Visual](#)).

**Game AI & Machine Learning** | Github: [Machine-Learning](#) | Python, NumPy, PyTorch

- Developed a **Go player identification** pipeline for SGF game records: engineered training-free opening move heat-map embeddings (black/white channels, cosine nearest neighbour) that reach  $\sim 0.83$  on the Kaggle public leaderboard, and extended it with a Supervised-Contrastive MLP encoder, class-balanced sampling, and standardized embeddings for improved few-shot style matching. ([Visual 1](#))
- Implemented core **classical ML algorithms from scratch** (Linear/Logistic Regression, SVR, Primal & Dual SVMs, SMO, CART, post-pruning, Random Forests) with NumPy and Matplotlib, met strict baselines (e.g., LR/SVR MSE thresholds, SVM  $\geq 90\%$  accuracy, RF outperforming single trees), and visualized decision boundaries, margins, pruning effects, and feature importance. ([Visual 2](#))
- Built end-to-end systems for **unsupervised learning and deep models**: K-Means and spectral clustering; PCA/ICA/t-SNE with pseudo-label pipelines; NumPy-only CNNs/ResNet/MobileNet-lite on MNIST/CIFAR-10; and PyTorch RNN/LSTM/GRU sequence models with gradient clipping, achieving target accuracies and parameter-efficiency trade-offs on standard benchmarks. ([Visual 3](#))

**Interactive 3D GUI/Rendering** | Github: [WebGL-ICG](#), [3D-Rendering](#) | C++ (OpenGL), WebGL (HTML, CSS, JavaScript)

- Devised a WebGL-based interactive 3D application demonstrating various **shading, transformation, lighting effects, texturing, animation**, multi-object rendering techniques, and user-configurable interactive features. ([Visual](#))
- Developed an OpenGL application capable of loading and rendering 3D models with textures, rotation, offset, and background color selection.
- Enhanced visualization by implementing texture filling, line-mode rendering, and user-configurable interface settings. ([Visual](#))

**CAN Bus Scheduling & ADAS Prototype** | Github: [Introduction-to-Intelligent-Vehicles](#) | Python

- Implemented **Fixed-Priority Response-Time Analysis (RTA)** on a 17-message CAN bus data set by computing blocking times and ceiling-based interference to derive per-message worst-case response times under Rate-Monotonic scheduling.
- Applied **Simulated Annealing (SA)** to optimize priority assignments for CAN bus messages, minimizing total worst-case response-time cost with unschedulability penalties via iterative swaps and Metropolis–Hastings acceptance. ([Visual 1](#))
- Developed KOSOKU, a modular **ADAS** system in **CARLA Simulator** integrating **V2X communication**, multimodal sensor fusion (LiDAR/Radar/Camera), Hybrid A\* state-search planning, and **OTA** map updates for dynamic trajectory generation and **replanning** ([Visual 2](#)).

**Visual-Oriented Music Mixer AR** | Github: [Stereo](#) | C# (Unity)

- Designed and implemented a visual-oriented **music mixer AR** as an interactive **audio-visual game** in Unity, integrating auditory and visual elements to enhance user engagement and creativity.
- Enabled users to create personalized remixes via **audio effect** processors and mixers, fostering community-driven media interaction.

**Medical Images Segmentation** | Github: [Medical-Images-Segmentation](#) | Python, PyTorch

- Implemented a medical image segmentation pipeline with **CNN-UNet**, **Transformer-UNet** (ViT bottleneck), and **self-supervised pretraining**, using a custom Weighted Dice-BCE Loss and strong data augmentation on the Kaggle competition dataset.
- Conducted ablations on augmentation, loss functions, and architectures, logged learning curves with Weights & Biases, and achieved up to 0.84 validation DSC and 0.82 public DSC with the Transformer + SSL model ([Visual](#)).

## Publications (Others: 6/9)

---

- **ACM SIGSPATIAL'25**, “M3: Recommendation via Attention-Graph Cluster Q-Learning with Multi-Scale Spatial Heterogeneity for Multi-Purpose, Multi-Stakeholder Green Attractions in Transportation”, A Rank in Spatiotemporal Systems (**1st Author**, Research Track, **26%** Acceptance, Paper).
- **ACM CIKM'24**, “LINKin-PARK: Land Valuation Information and Knowledge in Predictive Analysis and Reporting Kit via Dual Attention-DCCNN”, A Rank Conference in Data Mining and Information Retrieval (**2nd Author**, Demo Track, **47%** Acceptance, Paper).
- **ACM SIGSPATIAL'24**, “GreenSpot: Improving Public Transport with GIS-Based **AR** and Cluster-GCN Recommendation”, A Rank in Spatiotemporal Systems (**1st Author**, Demo Track, Paper).
- **ACM SIGSPATIAL'24**, “ACCEPT: A Context-Sensitive, Configurable, and Extensible Prediction Tool using Grid-based Data Processing and Neural Networks in Geospatial Decision Support”, A Rank in Spatiotemporal Systems (**2nd Author**, Demo Track, Paper).
- **ACM CHI'24 -EA**, “Footprints of Travel: AloT and **AR** Enhanced Tourist Gaming Experience in Unmanned Cultural Sites”, Top Conference, A+ Rank in Human-Computer Interaction (**1st Author**, Student Game Competition, **26.4%** Acceptance, Paper).
- **ACIS'23**, “Enhancing Urban Crowd Monitoring through Predictive Modelling System with Diverse Geospatial Datasets” in Australasian Conference on Information Systems (ACIS), A Rank in Information System (**3rd Author**, Proceeding, Paper).

## Community Service

---

**ACM @ CHI'25, HRI'25, CSCW'24, WWW'25** | Reviewer

Jan. 2024 – Jan. 2025

- Reviewed **8** full- and short-paper submissions in human-computer/robot interaction and web retrieval.

**NTU AI Club** @ Taipei, Taiwan | Lead Technical Specialist

Jan. 2025 – Jun. 2027

- Spearheaded academic research initiatives and lectures on multi-modal learning and reinforcement learning, **successfully mentoring undergrads to submit and accepted by CACS'25**; designed and mentored advanced technical workshops in AI research.

**Taiwan Air Force** @ Taipei Songshan Airport | Mandatory Military Service

Jul. 2024 – Nov. 2024

- Developed rapid-feedback mechanisms for a **150-person** unit supporting telecom infrastructure maintenance and repair.

## Professional Skills

---

**Coding** Python (PyTorch/TensorFlow/PyTorch Geometric), C/C++, C#, Matlab, LaTeX, OpenCV, **OpenGL**, **WebGL**, HTML, CSS, JavaScript.

**Tools** ROS 1/2, NVIDIA Omniverse Isaac Sim/Gym, CARLA, Unity 3D, Stable Baselines, Gym, Pybullet, Geopandas, **Linux**, **Docker**, **Git**.

**Languages** Mandarin (native speaker), English (professional working proficiency).