Yuriel Wang Jun Long Ryan

Email | Portfolio | LinkedIn | GitHub | +65 9799 9098

Education

Singapore University of Technology and Design (SUTD) Sep 2021 – Sep 2026
MEng (Research), Artificial Intelligence

- cGPA: *5.00/5.00*
- AI Singapore Accelerated Masters Program (2024 - 2026)
- DSO-AISG Incentive (Research) Award (2025-2026)

BEng (Computer Science) | Minor in Artificial Intelligence

- cGPA: 4.70/5.00, Honours with Highest Distinction
- SUTD Undergraduate Merit Scholarship (2021 - 2024)

Relevant Coursework: Machine Learning, Deep Learning, Computer Vision, Statistical NLP, Discrete Mathematics

Publications

Humor in Pixels: Benchmarking Large Multimodal Models Understanding of Online Comics *Nov 2025*
Findings of the Association for Computational Linguistics (EMNLP 2025)
Yuriel Ryan[*], Rui Yang Tan[*], Kenny Tsu Wei Choo, Roy Ka-Wei Lee

Self-Captioning Multimodal Interaction Tuning: Amplifying Exploitable Redundancies for *Under Review*
Robust Vision Language Models
Yuriel Ryan, Ip Hei Man, Adriel Kuek, Paul Pu Liang, Roy Ka-Wei Lee

Position: Balance Human Agency & AI Assistance in the Tussle for the Right to * *Under Review*
Zi-Yu Khoo, **Yuriel Ryan**, Nicole Heng Yim Oo, Hui En Pang, Eric J. W. Orlowski, Hakim Norhashim,
Ruth Wan Theng Chew, Rachael Hwee Ling Sim, Simon Chesterman, Jungpil Hahn, Bryan Kian Hsiang Low

[*] *Co-first authors.*

Work Experience

Research Assistant, Social AI Lab (SUTD) June 2024 – Dec 2024

- Automated scripts with **Selenium** and **Beautiful Soup** to collect 28,000 web comics for analysis.
- Recruited and managed 8 participants to annotate 2,800 comics with Label Studio to curate a benchmark.
- Evaluated large vision language models (e.g., Qwen2-VL-72B) to reveal their weaknesses in sequence recognition and biases against dark humor, leading to a publication in EMNLP Findings.
- Maintained a GitHub codebase for the collected data and ensure reproducible pipelines for evaluating VLMs on PixelHumor.

LLM Research Intern, DSO National Laboratories Aug 2023 – Dec 2023

- Applied **Graph of Thoughts** reasoning framework with GPT-4 to detect fragile functions within a 3-layer call stack, reducing incurred API costs by 25%.
- Implemented a Retrieval-Augmented Generation (RAG) system using **LangChain** with Llama 2 and Code Llama for deeper contextual code understanding.

Founding AI Engineer, Pallo (Iterative W25) Jan 2025 – Mar 2025

- Built a RAG system to ground large language models (LLMs) on Singapore exam syllabi, securing **pre-seed funding** from Iterative VC (Winter 2025).
- Built a document processing pipeline with open-source Vision Language Models, contributing 8,700 high-quality questions to enable RAG.
- Deployed **DeepSeek R1 models** to **Google Cloud Run** combined with RAG for solving Singapore GCE A-Level math problems, increasing accuracy scores by 36%.

Relevant Projects

Vision & Text Modalities | Self-Captioning MIT For Robust VLMs (MEng Thesis)   2026

- Analyzed vision language instruction datasets through an information theoretic lens to motivate hypotheses for Multimodal Interaction Tuning (MIT): adjusting redundant interactions (overlapping information) between modalities to address hallucination and robustness issues.
- Operationalize the analysis, using **vLLM**, to curate (e.g. cleaning, deduplication) three training sets (984,000 samples) of varying redundant interactions for further experiments.
- Test the hypotheses by fine-tuning VLMs—ranging from 256M (SmolVLM) to 8B (LLaVa-OneVision) parameters—with low rank adapters on the prepared training sets with additional redundancies, leading to a 38.3% decrease in visual-induced hallucinations and 16.8% gain in consistency.

Vision & Text Modalities | Augment and Think Before Answering: VLM Agentic Workflow   [GitHub], 2025

- Designed an agentic workflow to integrate external metadata (e.g., user comments) and slow down video inputs, exposing fine-grained temporal cues often missed by fixed-rate sampling.
- Implemented structured chain of thoughts using Google Gemini Pro 2.5 that enforces an "Augment Think (ANT) Before Answering" strategy to mitigate misleading queries.
- **Outperformed GPT-4o** across multiple question types, achieving a 58.07% Correctness Score (vs 45.2%) and a 15.3% Robustness Score (vs 8%) through the proposed Augment and Think pipeline, securing the **1st Runner-Up** award in the National AI Student Challenge (TikTok track).

Audio Modality | D'Noise (Capstone Project): Speech Enhancement   [GitHub], 2025

- Trained Google's Guided Speech Enhancement (GSE) Network in **PyTorch** to achieve an 87.2% noise reduction over non-deep learning methods with real-time inference (20ms latency).
- Collected over 20 hours of beamformed audio with IMDA's Singaporean local English to train the GSE Network with **Google Cloud Platform's Compute Engine** and **Cloud Storage**.

Visual Modality | No Pain, Just Gain (50.035 Computer Vision): Pose Estimation   [GitHub], 2024

- Trained Google's BlazePose model in **TensorFlow**, introducing Convolutional Block Attention Modules (CBAM) to achieve near real-time inference (100ms) for estimating **skeleton keypoints** in human poses.

Thermal & Depth Modalities | SmartDrive (UROP): Driver Fatigue Detection   2023

- Implemented a low-light fatigue detection pipeline using Luxonis Depth Cameras and FLIR Thermal Sensors through vision-based algorithms: PERCLOS and Head Pose Estimation.

Additional Information

Technical Skills: Python, PyTorch, TensorFlow, Google Cloud Platform
Awards: SUTD Honours List | Jyoti and Aditya Mathur Student Achievement Award | Google Professional ML Engineer Certification
Community Involvement: Singapore Youth AI Sub-committee | Access Singapore Volunteer (Programs and Operations) | CAT-IG Vice President | SUTD Climbers Secretary | Freshman Orientation Welfare Executive | TA for SUTD Freshmore Math and Computing courses.
Languages: Native English (spoken and written) | Conversational Mandarin (spoken)