ORCID: 0000-0002-0396-5938
Abu Dhabi, United Arab Emirates
naufalsuryanto@gmail.com

Naufal Suryanto

Postdoctoral Researcher @ KU

naufalso.github.io
github.com/naufalso
linkedin.com/in/naufal-suryanto

Postdoctoral researcher in **LLMs, multimodal models, and AI safety and security**, with a background in computer vision, adversarial machine learning, and cybersecurity. First author at **ICLR 2026, ICCV 2023, and CVPR 2022**; joint first author at **CVPR 2026 and NeurIPS 2026**. **Area Chair for CVPR 2027 and ECCV 2026** and **Outstanding Reviewer at CVPR 2026 and ICCV 2023**. Experienced across the full model development cycle, from large-scale data curation and continual pretraining to SFT, preference optimization, and RL on HPC clusters, with a consistent record of releasing open-source models, datasets, and benchmarks to the community.

Experience

Postdoctoral Researcher (LLM / Multimodal / Foundation Model / AI Safety) [Full-time] **May 2025 — Now**
Robust Intelligent Systems Lab (RISys-Lab), Khalifa University (Supervisor: Prof. Muzammal Naseer) *Abu Dhabi, UAE*

- ⋆ **5 papers at top-tier AI venues in 2026** (ICLR, CVPR, NeurIPS, ECCV, EMNLP), including **1 first-author** (ICLR) and **2 joint first-author** papers (CVPR, NeurIPS), all with open code, data, and models.
- **Cybersecurity LLMs:** Built **RedSage**, an open 8B cybersecurity LLM (11.8B-token continual pretraining, 266K-sample agentic SFT, DPO), and **KaliBench**, a Kali Linux tool-use benchmark with runtime-free verifiable rewards for RL.
- **Multimodal LLMs:** Co-developed **SPARROW** (pixel-grounded video MLLM, up to +8.9 J&F), **ReasonCLIP-58M** (reasoning supervision for CLIP), and **SAFIRE** (fire and smoke safety benchmark for MLLMs).

Postdoctoral Researcher (AI Security) [Full-time] **Mar 2024 — May 2025**
Information Security and AIoT Lab, Pusan National University *Busan, South Korea*

- Conducted research on **practical adversarial machine learning** and **AI security**, including an LLM assistant for penetration testing and a CAN bus intrusion detection system.

Graduate Researcher (AI, Computer Vision, Cybersecurity, and IoT) [Full-time] **Dec 2018 — Feb 2024**
Information Security and AIoT Lab, Pusan National University *Busan, South Korea*

- **AI Security (2020 - 2024):** Researched **physical adversarial attack & defense simulators**, **LLMs for pentesting**, and **computer vision applications**; published **1st-authored** papers at **CVPR 2022** and **ICCV 2023**, advancing SOTA physical adversarial attacks.
- **Cybersecurity and IoT (2018 - 2019):** Built **Smart City**, **Key Management Server**, **IoT DNS**, and **SSO** systems.

Software Developer (Java) [Full-time], *camLine Pte Ltd, Surabaya, Indonesia* **Oct 2017 — Dec 2018**

- Integrated Java software for manufacturing industries, **pioneering a new product** utilizing embedded devices.

AI and Big Data Scholar and Intern (Computer Vision) [Part-time], *Daewoong Foundation, Remote* **Feb 2021 — Sep 2022**

- **Scholarship** for **three periods** (22% acceptance); built few-shot classification, prescription OCR, and drug-information extraction.

Co-founder and Software Developer (Mobile App) [Part-time], *MUFKO, Remote* **From Feb 2020**

- Developed a cross-platform app for halal food and prayer spaces in Korea, with **100,000+ Google Play downloads**.

Robotic Researcher [Part-time], *EEPIS Robotic Research Center, Indonesia* **Mar 2014 — Oct 2017**

- Led EROS, a humanoid robot soccer team (RoboCup), pioneering **Robot Operating System (ROS)** adoption in the center.

Education

Ph.D. in Computer Engineering, *Pusan National University, South Korea* (GPA: 4.29 / 4.5) Mar 2019 — Feb 2024
- Supervisor: Prof. Howon Kim | Thesis: *Synthesizing Robust Physical Camouflage for Universal 3D Evasion Attacks*
Bachelor of Applied Science in Computer Engineering, *EEPIS, Indonesia* (Best graduate with GPA: 3.77 / 4.0) Aug 2013 — Sep 2017
Exchange Student (Machine Learning), *Anan National College of Technology, Japan* Mar 2017 — Jul 2017

Awards and Service

[Service] Area Chair, CVPR 2027 and ECCV 2026	2026 — 2027
[Service] Reviewer / PC Member: CVPR (2023–2026), ICCV (2023, 2025), ICLR 2026, AAAI 2026, ACL ARR 2026, IJCV	2022 — Now
[International] Outstanding Reviewer at CVPR 2026 (Top 5% of 17K reviewers) and ICCV 2023 (1.8% of reviewers)	2023, 2026
[National] Minister's Award for 'ICT Colloquium 2021' Student Creativity Project, Korea	2021
[National] Excellence Award at the 7th Korea SW Convergence Hackathon Competition, Korea	2020
[International] RoboCup Humanoid KidSize, Germany: Quarter Finalist	2016
[International] RoboCup Humanoid KidSize, China: 2nd Runner Up and 3rd Place in Technical Challenge	2015
[National] Humanoid Robot Soccer Competition, Indonesia: 1st Place and Best Innovation (2015), 2nd Place (2016)	2015 — 2016

Skills

Languages and Tools	Python, PyTorch, HuggingFace, TensorFlow, Git, Docker, Wandb, Slurm, OpenCV, etc.
Model Training	Continual Pretraining, SFT, DPO, GRPO (RLVR), LoRA, Synthetic and Agentic Data Pipelines
Research Interests	Trustworthy AI, AI Security, Cybersecurity LLMs, Multimodal LLMs, Adversarial ML, Computer Vision