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

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

Publications - Scopus / CORE Indexed (Sorted by Percentile)

[1] **[CORE A*]** Pengfei Li*, **Naufal Suryanto***, Sicheng Zhang, and Muzammal Naseer. "KaliBench: A Fine-Grained Benchmark for Cybersecurity Tool Use on Kali Linux with Runtime-Free Verifiable Rewards". In: *Advances in Neural Information Processing Systems (NeurIPS)*. *Equal contribution. 2026. URL: https://risys-lab.github.io/KaliBench/.

[2] **[CORE A*]** Pengfei Li, **Naufal Suryanto**, Sicheng Zhang, Mohammad Alsharid, and Muzammal Naseer. "SAFIRE: Safety-Critical Benchmark for Fine-grained Fire and Smoke Understanding in Multimodal LLMs". In: *Findings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2026. URL: https://arxiv.org/abs/2609.07823.

[3] **[CORE A*]** **[Top-1%]** Mohamad Alansari*, **Naufal Suryanto***, Divya Velayudhan, Sajid Javed, Naoufel Werghi, and Muzammal Naseer. "SPARROW: Learning Spatial Precision and Temporal Referential Consistency in Pixel-Grounded Video MLLMs". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. *Equal contribution. June 2026, pp. 17545–17556. URL: https://openaccess.thecvf.com/content/CVPR2026/html/Alansari_SPARROW_Learning_Spatial_Precision_and_Temporal_Referential_Consistency_in_Pixel-Grounded_CVPR_2026_paper.html.

[4] **[CORE A*]** Sicheng Zhang, Muzammal Naseer, Binzhu Xie, **Naufal Suryanto**, Shi Qiu, Jamal Bentahar, Naveed Akhtar, and Mubarak Shah. "ReasonCLIP-58M: Visually Grounded Commonsense Reasoning Supervision for CLIP". In: *European Conference on Computer Vision (ECCV)*. 2026. URL: https://arxiv.org/abs/2606.26794.

[5] **[CORE A*]** **Naufal Suryanto**, Muzammal Naseer, Pengfei Li, Syed Talal Wasim, Jinhui Yi, Juergen Gall, Paolo Ceravolo, and Ernesto Damiani. "RedSage: A Cybersecurity Generalist LLM". In: *The Fourteenth International Conference on Learning Representations (ICLR)*. 2026. URL: https://openreview.net/forum?id=W4FAenIrQ2.

[6] **[CORE A*]** **[Top-1%]** **Naufal Suryanto**, Yongsu Kim, Hyoeun Kang, Harashta Tatimma Larasati, Youngyeo Yun, Thi-Thu-Huong Le, Hunmin Yang, Se-Yoon Oh, and Howon Kim. "DTA: Physical Camouflage Attacks Using Differentiable Transformation Network". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2022, pp. 15305–15314. URL: https://openaccess.thecvf.com/content/CVPR2022/html/Suryanto_DTA_Physical_Camouflage_Attacks_Using_Differentiable_Transformation_Network_CVPR_2022_paper.html.

[7] **[CORE A*]** **[Top-6%]** **Naufal Suryanto**, Yongsu Kim, Harashta Tatimma Larasati, Hyoeun Kang, Thi-Thu-Huong Le, Yoonyoung Hong, Hunmin Yang, Se-Yoon Oh, and Howon Kim. "ACTIVE: Towards Highly Transferable 3D Physical Camouflage for Universal and Robust Vehicle Evasion". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. Oct. 2023, pp. 4305–4314. URL: https://openaccess.thecvf.com/content/ICCV2023/html/Suryanto_ACTIVE_Towards_Highly_Transferable_3D_Physical_Camouflage_for_Universal_and_ICCV_2023_paper.html.

[8] **[Top-8%]** **Naufal Suryanto**, Andro Aprila Adiputra, Ahmada Yusril Kadiptya, Thi-Thu-Huong Le, Derry Pratama, Yongsu Kim, and Howon Kim. "Cityscape-Adverse: Benchmarking Robustness of Semantic Segmentation With Realistic Scene Modifications via Diffusion-Based Image Editing". In: *IEEE Access* 13 (2025), pp. 69921–69940. URL: https://ieeexplore.ieee.org/document/10870179.

[9] **[Top-17%]** **Naufal Suryanto**, Hyoeun Kang, Yongsu Kim, Youngyeo Yun, Harashta Tatimma Larasati, and Howon Kim. "A Distributed Black-Box Adversarial Attack Based on Multi-Group Particle Swarm Optimization". In: *Sensors* 20.24 (2020). URL: https://www.mdpi.com/1424-8220/20/24/7158.

[10] **[Top-17%]** Derry Pratama, **Naufal Suryanto**, Andro Aprila Adiputra, Thi-Thu-Huong Le, Ahmada Yusril Kadiptya, Muhammad Iqbal, and Howon Kim. "CIPHER: Cybersecurity Intelligent Penetration-Testing Helper for Ethical Researcher". In: *Sensors* 24.21 (2024). URL: https://www.mdpi.com/1424-8220/24/21/6878.

[11] **[Top-17%]** Yongsu Kim, Hyoeun Kang, **Naufal Suryanto**, Harashta Tatimma Larasati, Afifatul Mukaroh, and Howon Kim. "Extended Spatially Localized Perturbation GAN (eSLP-GAN) for Robust Adversarial Camouflage Patches". In: *Sensors* 21.16 (2021). URL: https://www.mdpi.com/1424-8220/21/16/5323.

[12] **[Top-21%]** Aji Teguh Prihatno, **Naufal Suryanto**, Sangbong Oh, Thi-Thu-Huong Le, and Howon Kim. "NFT Image Plagiarism Check Using EfficientNet-Based Deep Neural Network with Triplet Semi-Hard Loss". In: *Applied Sciences* 13.5 (2023). URL: https://www.mdpi.com/2076-3417/13/5/3072.

[13] **Naufal Suryanto**, Andro Aprila Adiputra, Ahmada Yusril Kadiptya, Yongsu Kim, and Howon Kim. "Adversarial Manhole: Challenging Monocular Depth Estimation and Semantic Segmentation Models with Patch Attack". In: *Information Security Applications*. Ed. by Jong-Hyouk Lee, Keita Emura, and Sokjoon Lee. Singapore: Springer Nature Singapore, 2025, pp. 299–310. URL: https://link.springer.com/chapter/10.1007/978-981-96-1624-4_23.

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

[14] Aji Teguh Prihatno, **Naufal Suryanto**, Harashta Tatimma Larasati, Yustus Eko Oktian, Thi-Thu-Huong Le, and Howon Kim. "A New Frontier in Digital Security: Verification for NFT Image Using Deep Learning-Based ConvNeXt Model in Quantum Blockchain". In: *Information Security Applications*. Ed. by Howon Kim and Jonghee Youn. Singapore: Springer Nature Singapore, 2024, pp. 79–90. URL: https://link.springer.com/chapter/10.1007/978-981-99-8024-6_7.

[15] Thi-Thu-Huong Le, **Naufal Suryanto**, Howon Kim, Janghyun Ji, and Shinwook Heo. "Enhancing Intrusion Detection and Explanations for Imbalanced Vehicle CAN Network Data". In: *Proceedings of the 12th International Symposium on Information and Communication Technology*. SOICT '23. Ho Chi Minh, Vietnam: Association for Computing Machinery, Dec. 2023, pp. 777–784. URL: https://dl.acm.org/doi/abs/10.1145/3628797.3628994.

[16] Yongsu Kim, Hyoeun Kang, Afifatul Mukaroh, **Naufal Suryanto**, Harashta Tatimma Larasati, and Howon Kim. "Spatially Localized Perturbation GAN (SLP-GAN) for Generating Invisible Adversarial Patches". In: *Information Security Applications*. Ed. by Ilsun You. Cham: Springer International Publishing, 2020, pp. 3–15. URL: https://link.springer.com/chapter/10.1007/978-3-030-65299-9_1.

[17] **Naufal Suryanto**, Chihiro Ikuta, and Dadet Pramadihanto. "Multi-group particle swarm optimization with random redistribution". In: *2017 International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC)*. Sept. 2017, pp. 1–5. URL: https://ieeexplore.ieee.org/abstract/document/8228445.

Publications - Domestic Conferences and Journals

[18] Rini Wisnu Wardhani, Dedy Septono Catur Putranto, Thi-Thu-Huong Le, Yustus-Eko Oktian, Uk Jo, Aji Teguh Prihatno, **Naufal Suryanto**, and Howon Kim. "Trend of Paradigm for integrating Blockchain, Artificial Intelligence, Quantum Computing, and Internet of Things". In: *Smart Media Journal* 12.2 (Mar. 2023), pp. 42–55. URL: https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE11597589.

[19] **Naufal Suryanto**, Harashta Tatimma Larasati, Yongsu Kim, and Howon Kim. "Adversarial Wall: Physical Adversarial Attack on Cityscape Pretrained Segmentation Model". In: *Korea Institute of Information Security and Cryptology*. Korea Institute of Information Security and Cryptology. Nov. 2022, VOL 29 NO. 02 PP. 0402 0404 (2022. 11). URL: https://www.cseric.or.kr/literature/ser_view.php?SnxIndxNum=223993.

[20] **Naufal Suryanto**, Ismail, Harashta Tatimma Larasati, Shinwook Heo, and Howon Kim. "OCF Gateway Implementation for IoT Platforms Interoperability". In: *Conference on Information Security and Cryptography-Winter 2019 (CISC-W'19)*. Nov. 2019.

[21] Ismail, **Naufal Suryanto**, Hyeongon Kim, Shinwook Heo, and Howon Kim. "A Study on Federated Access Control in Interoperability of Heterogeneous IoT Platforms". In: *Korea Communications Society Summer Conference* (June 2019), pp. 140–141. URL: https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE09234246.

[22] **Naufal Suryanto**, Fendi Nur Wahyu, and Ahmad Subhan Khailullah. "Software Framework on EROS-6 (EEPIS ROBOT SOCCER-6) Platform". In: *4th Indonesian Symposium on Robot Soccer Competition (ISRSC)*. 2016, pp. 137–140.

Publications - Others (Books and Technical Reports)

[24] **Naufal Suryanto**. "Synthesizing Robust Physical Camouflage for Universal 3D Evasion Attacks". PhD thesis. Busan, South Korea: Pusan National University, Feb. 2024.

[25] **Naufal Suryanto**. *Design and Implementation of ROS Based Software Framework on Humanoid Robot Soccer*. Bachelor Thesis. Surabaya, Indonesia: Electronics Engineering Polytechnic Institute of Surabaya, 2017.

[26] Dadet Pramadihanto, Eko Henfri Binugroho, Ahmad Subhan Khalilullah, Dimas Pristovani, Ardiansyah Al-Farouq, Rimsyah Anjar lestiawan, Muhammad Rindo Wardhana, Anas Abdulloh Zaini, **Naufal Suryanto**, Ajir, and Yasinta Musa Dewi. *TFlow TEAM - Team Description for Humanoid Teen League of RoboCup 2017*. Tech. rep. EEPIS Robotic Research Center (ER2C), 2017. URL: https://humanoid.robocup.org/qualification/2017/b3b640d7c0ccb8808547a87bbb88b0e6b2247b18/T_FLOW_Humanoid_TeenSize_regularanddrop_in_2017_TDP.pdf.

[27] Ahmad Subhan Khalilullah, **Naufal Suryanto**, Adi Sucipto, Imam Fajar Fauzi, Fendiq Nur Wahyu, Muhammad Lutfi Santoso, Krisna Adji Syahputra, Aulia Khilmi Rizgi, Ilham Fakhrul Arifin, and Ryan Satria Wijaya. *EROS TEAM - Team Description for Humanoid KidSize League of RoboCup 2017*. Tech. rep. EEPIS Robotic Research Center (ER2C), 2017. URL: https://humanoid.robocup.org/qualification/2017/b47bd43451db4ac63b991dd757561d693dcf7edb/EROS_Team_Humanoid_KidSize_regularanddrop_in_2017_TDP.pdf.

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

[28] Ahmad Subhan Khalilullah, Iwan Kurniato W, **Naufal Suryanto**, Achmad Syaifuddin Zuhri, Luthfi Aminulloh, Hasbulloh Qohar, Fendiq Nur Wahyu, Imam Fajar Fauzi, Adi Sucipto, Dimas Alfiansyah, and Muhammad Lutfi Santoso. *EROS TEAM - Team Description for Humanoid KidSize League of RoboCup 2016*. Tech. rep. EEPIS Robotic Research Center (ER2C), 2016. URL: https://humanoid.robocup.org/qualification/2016/9ea3be9ccad446f182b92a5f181ddd704453301c/EROS_Humanoid_KidSize_2016_TDP.pdf.

[29] Ahmad Subhan Khalilullah, Iwan Kurniato W, Dimas Pristovani, Pratomo Adhi Nugroho, Saifuddin Zuhri, **Naufal Suryanto**, Rizki Taufik, Luthfi Aminulloh, and Hasbulloh Qohar. *EROS TEAM - Team Description for Humanoid KidSize League of RoboCup 2015*. Tech. rep. EEPIS Robotic Research Center (ER2C), 2015. URL: https://humanoid.robocup.org/qualification/2015/ecbb6338f398dc75a053e0afbea1fa6dfb0d0e46/EROS_Humanoid_KidSize_2015_TDP.pdf.

Projects - LLM, Multimodal LLM, and AI Safety

1. RedSage: A Cybersecurity Generalist LLM **ICLR 2026**

Associated with Khalifa University | University of Bonn | University of Milan

- **Project Overview:** Built an open 8B cybersecurity generalist LLM and its evaluation benchmark, covering the full pipeline from domain data curation and continual pretraining to agentic instruction tuning and preference alignment.
- **Key Contributions:**
  - Curated 11.8B tokens of cybersecurity data (CyberFineWeb and 28K seed documents from MITRE, OWASP, NIST, and offensive-security write-ups) for continual pretraining of Qwen3-8B.
  - Designed an agentic augmentation pipeline that generated 266K multi-turn cybersecurity SFT samples, followed by DPO alignment.
  - Created RedSage-Bench with 30K multiple-choice and 240 open-ended questions.
  - Outperformed baselines by up to **+5.59 points** on cybersecurity benchmarks and **+5.05 points** on Open LLM Leaderboard tasks.
- **Notable Outputs: First-author** paper at **ICLR 2026**; base, instruct, and DPO models, data, and code released (project page).
- **Skills:** Large Language Model, Continual Pretraining, SFT, DPO, Data Curation, Agentic Data Generation, Benchmarking, PyTorch, HuggingFace, Slurm, etc.

2. KaliBench: Cybersecurity Tool-Use Benchmark with Verifiable Rewards **NeurIPS 2026**

Associated with Khalifa University | University of Western Australia

- **Project Overview:** Built a fine-grained benchmark for translating natural-language security tasks into Kali Linux commands, with a runtime-free verifiable reward that scores commands without executing them, enabling reinforcement learning on tool use.
- **Key Contributions:**
  - Constructed 8,504 query-command pairs spanning 1,642 tools, 23 capability dimensions, and 5 security phases.
  - Evaluated 24 open-weight model configurations and showed that no model exceeds 42% exact-command accuracy in the unrestricted setting.
  - Trained RedSage-K (8B) with SFT and GRPO, raising the average score from 71.7 to **79.2**, within 1.0 point of DeepSeek-V3.2.
- **Notable Outputs: Joint first-author** paper at **NeurIPS 2026** (Evaluations and Datasets Track); data, models, and code released (project page).
- **Skills:** Benchmark Design, Reinforcement Learning, GRPO, Verifiable Rewards, SFT, LLM Tool Use, Penetration Testing Tools, etc.

3. SPARROW: Pixel-Grounded Video MLLM **CVPR 2026**

Associated with Khalifa University

- **Project Overview:** Developed a training framework that improves spatial precision and temporal referential consistency of video MLLMs that ground objects with boxes and segmentation masks.
- **Key Contributions:**
  - Curated a referential video dataset of 30,646 videos and 45,231 Q&A pairs from 7 source datasets.
  - Introduced Target-Specific Tracked Features and a dual-prompt [BOX]/[SEG] design with a class-agnostic SAM2-based proposer, without external detectors at inference.
  - Integrated into three open video MLLMs (UniPixel, GLUS, VideoGLaMM) across six benchmarks, gaining up to **+8.9 J&F** on RVOS, **+5 mIoU** on visual grounding, and **+5.4 CLAIR** on GCG.
- **Notable Outputs: Joint first-author** paper at **CVPR 2026**; code, datasets, and models released (project page).
- **Skills:** Multimodal LLM, Video Understanding, Referring Segmentation, Visual Grounding, LoRA, SAM2, PyTorch, etc.

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

4. SAFIRE: Fire and Smoke Safety Benchmark for Multimodal LLMs **EMNLP 2026**
Associated with Khalifa University | University of Western Australia

- **Project Overview:** Co-developed a safety-critical benchmark of 82,842 captioned images across 20 fire-smoke scenarios and 192,904 question-answer pairs over 10 evaluation dimensions. Ten open MLLMs (8B to 38B) averaged only 61.9% accuracy, while adapting vision encoders with 7% of the data raised fire-scene classification from 20.1% to **64.5%**.
- **Notable Outputs:** Paper at **EMNLP 2026 Findings**; code and data released (project page).
- **Skills:** Multimodal LLM, Benchmark Design, Data Annotation Pipelines, Vision Encoders, AI Safety, etc.

5. ReasonCLIP-58M: Reasoning Supervision for CLIP **ECCV 2026**
Associated with Khalifa University

- **Project Overview:** Co-developed a continual-pretraining framework that adds visually grounded commonsense reasoning supervision to CLIP-style encoders without architectural changes, using 58M reasoning captions (ReasonLite-42M and ReasonPro-16M) and the RCLIP-Bench diagnostic benchmark. As a drop-in visual encoder for LLaVA-NeXT, it gives consistent gains at no extra inference cost.
- **Notable Outputs:** Paper at **ECCV 2026**; code, datasets, and models released (GitHub).
- **Skills:** Vision-Language Models, CLIP, Continual Pretraining, Large-Scale Data Curation, etc.

Projects - AI Security, Computer Vision, and Industrial AI Applications

1. Automobile Security **Sep 2024 — May 2025**
Associated with Pusan National University | SmartM2M. Co., Ltd.

- **Project Overview:** Contributed to applying AI in automotive cybersecurity by developing a CAN BUS Intrusion Detection System (IDS) through gathering and processing clean and simulated attack data, curating cybersecurity resources, and creating an AI Agent that can interact with external vulnerability sources for enhanced security assessments.
- **Key Contributions:**
  - Researched AI techniques tailored for automotive security.
  - Implemented data collection, model training, and deployment for a CAN BUS IDS.
  - Curated cybersecurity resources and developed an AI Agent for external vulnerability integration.
  - Developed a proof-of-concept for an AI-driven security assessment framework.
- **Skills:** Automotive Cybersecurity, AI Research, CAN BUS, Machine Learning, Intrusion Detection, Vulnerability Assessment, Python, Deep Learning, Large Language model, Generative AI, AI Agents.

2. AI Assistant for Penetration Testing **Mar 2023 — May 2025**
Associated with Pusan National University | SmartM2M. Co., Ltd.

- **Project Overview:** Co-led the development of an AI-powered assistant for penetration testing. Collaborated with the Security team to provide advanced assistance and insights to novice pentesters, guiding them through various cybersecurity tools and penetration testing techniques.
- **Key Achievements:**
  - Co-led the project, overseeing the AI team's contributions and collaborating with the Security team.
  - Utilized a strong open-source base Language Model (LLM) for developing conversation capabilities.
  - Actively processed high-quality datasets for basic cybersecurity and penetration testing scenarios.
  - Implemented UI for the chat interface and facilitated continuous model improvements through SFT, RLHF, and DPO.
  - Conducted continuous research and development to enhance the model capabilities.
  - Directed the development of a Korean version of the model for broader accessibility.
- **Skills:** Deep Learning, Large Language Model, Generative AI, Project Management, Research, Software Development, WandB, HuggingFace, LangChain, Pytorch, Gradio, StreamLit, FastAPI, Retrieval-Augmented Generation, VectorDB, Data Engineering, etc.

3. Physical Adversarial Attack and Defense Simulator **Dec 2020 — Sep 2024**
Associated with Pusan National University | SmartM2M. Co., Ltd. | Agency of Defense Developments, South Korea

- **Project Overview:** Perform the development of Adversarial AI technology by creating a simulator for performing and evaluating physical adversarial attacks/defenses in diverse situations. Conducted research and development of adversarial attack/defense techniques, playing a key role in building a robust physical simulation environment.
- **Key Achievements:**
  - Designed and implemented scalable adversarial attack simulation modules, employing various attack techniques.
  - Developed a physical attack simulation environment and engineered robust deep learning models.
  - Integrated attack models based on 3D modeling, graphics engines, neural rendering, etc., for the attack pattern generation.
  - Established a synthetic data production environment for generating diverse attack patterns.

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

  - Developed an evaluation environment to assess the effectiveness of provided attack/defense techniques.
  - Designed scalable defense model frameworks for the Adversarial Defense Simulation Module.
- **Notable Outputs:** Published first-author papers in **CVPR 2022** and **ICCV 2023**, establishing the current **state-of-the-art** in Physical Camouflage Attack. Awarded the **Minister's Award** in the student creative autonomous task category at the **ICT Colloquium 2021**.
- **Skills:** AI Security, Deep Learning, Computer Vision, Object Detection, Object Segmentation, Software Development, Research, TensorFlow, PyTorch, OpenCV, MMdetection, Wandb, Unreal Engine, Carla Simulator, Neural Rendering, NeRF, etc.

4. NFT Digital Content Similarity Verification Technology **Jun 2023 — Dec 2023**

Associated with Pusan National University | SmartM2M. Co., Ltd.

- **Project Overview:** Conducted comprehensive research on similarity verification technology for NFT digital images. Explored image hashing techniques and AI-based similarity measurement methods. Developed and demonstrated image similarity verification concept for the NFT market, incorporating a dataset survey and experimentation with forged NFT test images.
- **Key Achievements:**
  - Engineered a robust deep-learning-based image similarity model using a contrastive learning approach.
  - Seamlessly integrated the image similarity model into the backend server, ensuring efficient model inference during deployment.
  - Implemented image similarity verification using advanced AI techniques and image hashing.
- **Skills:** Deep Learning, Computer Vision, Software Development, Research, TensorFlow, Keras, Contrastive Learning, FastAPI, etc.

5. Fish Disease Classification and Detection **Sep 2022 — Mar 2023**

Associated with Pusan National University

- **Project Overview:** Developed AI models to classify and detect diseases in fish using real images captured by a fisherman in Jeju. Used few-shot learning for classification due to limited data and trained an object detection model for precise disease location.
- **Key Achievements:**
  - Created and implemented multiple AI models for accurate fish disease classification and detection.
  - Employed few-shot learning techniques to enhance the classification model's performance with a small dataset.
  - Trained an object detection model ensuring precise localization of fish diseases in images.
- **Skills:** Deep Learning, Computer Vision, Object Classification, Object Detection, Few-shot Learning, TensorFlow, etc.

6. Vision Global Web Scraping **Apr 2022 — Sep 2022**

Associated with Daewoong Foundation

- **Project Overview:** Developed a comprehensive solution for web scraping and data extraction from e-commerce websites to analyze drug market trends. Managed the Vision Color Extractor, creating a pipeline for extracting dominant colors from images on various e-commerce sites. The extracted data is utilized by the business team for in-depth analysis.
- **Key Achievements:**
  - Engineered a robust pipeline for color extraction, covering foreground extraction, clustering, classification, and leveraging different color spaces.
  - Conducted continuous evaluation and enhancements to improve the performance of the color extraction pipeline.
- **Skills:** Machine Learning, Computer Vision, Web Scraping, Data Extraction, Image Processing, OpenCV, etc.

7. Medical Prescriptions Data Extraction **Jul 2021 — Oct 2021**

Associated with Daewoong Foundation

- **Project Overview:** Automated the data input process from printed medical prescription forms using Optical Character Recognition (OCR). The sales department receives thousands of these forms monthly, and the project's goal is to extract text and data, reducing manual input time, minimizing errors, and enhancing overall efficiency.
- **Key Achievements:**
  - Researched and tested various open-source OCR software/models, implementing preprocessing techniques to improve accuracy. Adjusted form angles, removed color stamps, and classified forms based on keywords using regex.
  - Implemented grayscale conversion, thresholding, inversion, and dilation for noise reduction and improved OCR results.
  - Developed methods for extracting dates using format-specific regular expressions. Implemented a clustering-based algorithm for table data extraction, grouping data under each column header.
  - Conducted evaluations on OCR accuracy, date extraction, and table extraction, showing promising results with low error rates.
- **Skills:** Machine Learning, Computer Vision, OCR, Image Preprocessing, Data Extraction, Regular Expressions, OpenCV, etc.

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

8. Drug Package Classifier **Feb 2021 — May 2021**

Associated with Daewoong Foundation

- **Project Overview:** Developed drug package classifiers enabling patients to retrieve information by capturing a picture of the package. Originally proposed with OCR-based text detection, I introduced an innovative alternative: a few-shot classifier achieving 100% test accuracy. Designed and implemented the backend for model serving and a cross-platform Web App, enhancing accessibility to the drug information service. Recognized as an excellent project.
- **Key Achievements:**
  - Proposed and implemented an alternative few-shot classifier with 100% test accuracy.
  - Developed a robust backend infrastructure for model deployment and serving.
  - Created a cross-platform Web App, improving user accessibility to the drug information service.
- **Skills:** Deep Learning, Computer Vision, Few-shot Learning, TensorFlow, Back-end Development, Web App Development, etc.

9. Decentralized Intelligent CCTV / Video Surveillance System **Nov 2020 — Dec 2020**

Associated with Pusan National University

- **Project Overview:** Developed a Decentralized Intelligent CCTV system employing RetinaFace for face region detection, FaceNet for facial feature extraction, and DeepSort for face tracking. The system assigns unique IDs upon face detection, storing facial features as embedding vectors in a database. This facilitates efficient face tracking, providing crucial information about target individuals.
- **Key Achievements:**
  - Engineered an Intelligent CCTV system integrating RetinaFace, FaceNet, and DeepSort for effective face detection, feature extraction, and tracking.
  - Implemented a unique ID assignment and database storage mechanism, ensuring accurate individual tracking.
  - Established a querying system using facial embeddings to retrieve face-tracking logs, aiding in the identification of target individuals with detailed time and camera information.
- **Skills:** Deep Learning, Computer Vision, Object Detection, Object Recognition, Object Tracking, Edge AI, OpenCV, TensorFlow, etc.

Projects - Cybersecurity and IoT

1. Crypto SoC EVM V2 PCB Design **Oct 2019 — Dec 2019**

Associated with Pusan National University

- **Project Overview:** Designed the PCB for the hardware team's Crypto SoC (System on Chip), focusing on encrypting KEPCO's devices at the hardware level. Developed multiple PCB versions based on the hardware team's requirements, incorporating components such as power regulator blocks, serial communication blocks using FTDI, and hardware user interfaces (buttons, dip switches, headers). Additionally, designed PCBs for the FTDI version to aid hardware teams in prototyping.
- **Key Achievements:**
  - Successfully designed multiple PCB versions tailored to the hardware team's specifications.
  - Implemented essential components, including power regulator blocks, FTDI-based serial communication blocks, and hardware user interfaces.
  - Designed PCBs for both the main SoC chips and the FTDI version, facilitating hardware team prototyping activities.
- **Skills:** Cybersecurity, Hardware, Printed Circuit Board (PCB) Design, Eagle PCB, etc.

2. ETRI Smart Glasses Key Management Server **Sep 2019 — Dec 2019**

Associated with Pusan National University

- **Project Overview:** Completed full-stack development for ETRI Smart Glasses Key Management, including backend, frontend, and Android App. Independently designed and implemented systems, utilizing asymmetric key encryption for secure credentials. The backend functions as a REST API server for registration, authentication, etc., while the frontend provides a user-friendly interface using QR codes for Smart Glasses interaction.
- **Key Achievements:**
  - Completed full-stack development, covering backend, frontend, and Android App for ETRI Smart Glasses Key Management.
  - Implemented asymmetric key encryption for secure Smart Glasses credential management.
  - Developed a REST API server for user and device registration, authentication, etc.
  - Integrated QR codes for efficient information and public key exchange with the Smart Glasses app.
- **Skills:** Cybersecurity, Android, Software Development, Front-end Development, Back-End Development, MySQL, NodeJS, ExpressJS, NodeJS, VueJS, etc.

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

3. KETI Smart City Data Hub Platform **May 2019 — Dec 2019**

Associated with Pusan National University

- **Project Overview:** Involved in the development of the KETI Smart City Data Hub Platform, contributing to access control, microservices architecture, and security implementations.
- **Key Achievements and Contributions:**
  - Enhanced the user interface, containerized applications, and upgraded the Access Control Module to Role-Based Access Control (RBAC).
  - Managed database migration, implemented OAuth2 for enhanced security, and contributed to comprehensive documentation.
- **Skills:** Cybersecurity, Software Development, Java, JavaScript, Docker, JCasbin, Java Spring Security, JWT Token, MySQL, PostgreSQL, OAuth2, NodeJS, VueJS, JIRA, Bitbucket, etc.

4. Identification and Interworking System in IoT Platforms using IoT DNS **Dec 2018 — Feb 2019**

Associated with Pusan National University

- **Project Overview:** Developed and implemented an IoT DNS system for identifying and interworking various IoT Platforms, including OCF, OneM2M, GS1, and IBM Watson.
- **Key Achievements:**
  - Successfully implemented IoT Gateway, enabling platform and protocol translation.
  - Implemented AAS Server (Authentication and Authorization Server) with support for Single Sign-On (SSO) functionality.
  - Developed a robust IoT DNS system for registration, discovery, and related functionalities.
  - Seamlessly integrated all system components, ensuring efficient operation and interworking across diverse IoT Platforms.
- **Skills:** Cybersecurity, IoT Interoperability, SSO, DNS, System Architecture, Full Stack Development, Authentication and Authorization, Docker, NodeJS, VueJS, OAuth2, JIRA, Bitbucket, etc.