Portfolio — Hyeongju Lee

Data & Backend Engineer with a 3D computer-vision research background. Master's completion Mar. 2027.
zoocasso1590@gmail.com | zoocasso.github.io

1. Data & Backend Engineering — Vision Inside (Jul. 2022 – Nov. 2024)

KISTI-commissioned Public R&D Value-Creation Platform (later SMART K2C) *Aug. 2022 – Jun. 2024*

- **Context:** an outsourcing contract from KISTI to build the data foundation of a platform that evaluates products' market competitiveness. University partners designed some of the models; Vision Inside built and ran the data and serving layer.
- **Role:** the project's data engineer on the Vision Inside side, **owning collection, loading and serving**.
- **Collection:** Selenium crawlers for Amazon and Danawa and public-data-portal APIs for the Public Procurement Service, run on a distributed crawling farm of laptops feeding a central MySQL. Collected Amazon and Danawa category by category and de-duplicated products by ID (ASIN).
- **Data:** procurement, Danawa and Amazon data in one comparable dataset.
- **Models I built (3–5), in production:** project–company keyword matching in SQL; price-band K-means clustering with Dunn-index k selection (competitor benchmarking); a daily KeyBERT + Sentence-Transformers + Mecab batch extracting customer-required quality keywords from reviews; procurement price/attribute/delivery statistics in SQL.
- **Serving:** FastAPI services with Jinja2 + Tabulator admin dashboards (filtering, sequential batch runs, unified search, Excel export). In the Java Spring KISTI prototype web app, owned the DB schema, backend integration and frontend wiring (with the in-house Java team), and moved it with a partner's models 1–2 to company servers on a Docker runtime.
- **Delivery:** wrote the audit deliverables (operator manual, initial data-build report); presented crawling and model results in technical meetings with KISTI researchers and handled their data and code requests directly. The project later led to a follow-on main contract.
- **Stack:** Python, Selenium, BeautifulSoup, pandas, SQLAlchemy, MySQL, scikit-learn, KeyBERT, Sentence-Transformers, FastAPI, Jinja2, Tabulator, Java Spring. Details: zoocasso.github.io/en/work/kisti

Genomic Variant Annotation & Risk-Report System (Kyung Hee Univ. × Canadian bio lab) *Aug. 2022 – Sep. 2023*

- **Problem:** in an international collaboration between Kyung Hee University's College of Korean Medicine and a Canadian bio research lab, turn customer VCF or chip (genotyping-array) input into readable disease/trait reports using large public genomic references.
- **Role: backend, pipeline and frontend developer** (algorithm designed with the team lead).
- **Variant calling:** FASTQ → BWA → samtools → Picard → GATK → VCF; PyVCF + regex normalization of heterogeneous VCF/array inputs.
- **Reference data:** dbSNP GRCh37/38, 1000 Genomes, GWAS Catalog, UniProt variants; loaded into MySQL via chromosome-wise splitting (sed/split) and chunked pandas/SQLAlchemy ingestion. Annotation is a 4-stage SQL join keyed on position and alleles or rsID.
- **Report logic:** GWAS odds-ratio/beta-based risk scores per disease/trait, allele-frequency comparison across 26 1000 Genomes populations, and drug/herbal-medicine lookups (MedlinePlus, MeSH, ATC, PharmDB-K, OASIS), exposed as a FastAPI JSON API. LDA and 1D-CNN disease-to-herb models were prototyped.
- **Deployment:** Docker images through a Harbor private registry I built on a Korea IDC Ubuntu server, with domain and HTTPS (Let's Encrypt); access restricted to internal IPs (not publicly launched).
- **Improved:** web screens rendered inside the server → a separate JSON API for the frontend.
- **Collaboration:** technical bridge between the Korean and Canadian teams, working with the lab's designer.
- **Stack:** Python, BWA, samtools, Picard, GATK, PyVCF, pandas, SQLAlchemy, MySQL, FastAPI, Docker, Harbor. Code: github.com/zoocasso/bio_fastapi. Details: zoocasso.github.io/en/work/genomics
- **Life-insurance disclosure tool** (team of three): loaded insurance disclosure workbooks into MySQL and served them through a FastAPI web viewer.

2. Research & Applied Vision — Kyushu Institute of Technology (Apr. 2025 – present)

Training-free Light-Field / Integral-Imaging Reconstruction (advisor: Prof. Min-Chul Lee) *4 first-author papers*

- **Ray-based Inverse Mapping (RIM):** redistributes angular samples to spatial positions as a bijection with zero quantization error under orthographic conditions. On 7 Stanford Lytro scenes: PSNR 23.31 → 25.70 dB (**+2.39 dB**), SSIM 0.741 → 0.877; downstream YOLO detection mAP50 0.665 → 0.776 (**+0.11**). Variant iRIM synthesizes virtual lens arrays from multi-view images. Details: zoocasso.github.io/en/research/rim
- **Accelerated origin-point optimization:** replaced the iterative search with direct camera-parameter computation — **81 reconstructions → 1** on 9×9 elemental images (paper: 0.31 s on CPU vs. an estimated 25.45 s = 81 × 0.3142 s; 0.0057 s on GPU with CuPy, RTX 3090); same origin and SSIM in the paper's experiments. Details: zoocasso.github.io/en/research/origin-point
- **Real-time pipeline (master's thesis, in preparation):** detect once on a lens-array image, then refocus only inside each box on the GPU — **16.4 ms per real Lytro frame (61 fps, RTX 3090)**.

- **Image enhancement (patent application in preparation):** enhancement for low-light and hazy scenes, one image at a time.
- **Starting point in vision and ML (2022):** depth estimation with a correlation filter and VCR (course project, MATLAB) became the starting point of this research. The same year: an MNIST CNN in MATLAB with backpropagation and momentum gradient descent written by hand, no deep-learning framework. Interest in vision started with the Smart Mirror graduation project (two-person team: Raspberry Pi, OpenCV LBPH face recognition, Kakao speech API).
- **Implementation:** Python, NumPy, CuPy; evaluation loop with PSNR, SSIM, and detection mAP.

Image-Data Pipeline — joint research with Hitachi *Apr. 2025 – present*

- **Role:** research assistant (part-time) in a Kyushu Institute of Technology lab.
- **Problem:** the inherited code for producing training images could not be maintained, and dataset creation depended on developers.
- **Built:** an end-to-end Python pipeline — machine-vision camera control and capture, cleaning, preprocessing into model-ready datasets — with a PyQt GUI that non-developer researchers operate themselves.
- **Improved:** only the first frame of dark, low-light video, with a home-made C++ camera layer → single high-quality still captures at the right brightness, through the recommended SDK (PySpin).
- **Outcome:** deployed in a lab in the Netherlands: used hands-on by Hitachi staff, demonstrated to the Dutch researchers. Design handed over to two junior researchers. Working languages: Japanese and English.
- **Stack:** Python, OpenCV, Spinnaker SDK (PySpin), PyQt, PyInstaller. Details: zoocasso.github.io/en/work/hitachi

3. Other Projects

- **3DOS — vision enhancement on an XR headset** (2024 XR Device Content Makerthon finals, Nov. 2024; team 3DOS, five lab members registered) — **owned all development and the live demo**. Fourier peplography (fog/smoke removal) and photon counting (low-light recovery), methods from Prof. Myungjin Cho's research, in C++/OpenCV, built as an Android native plugin (arm64) and loaded by a Unity (IL2CPP) app on the Metalens 2 headset, applied to its own camera feed. Slides: github.com/zoocasso/realtime-vision-xr-device
- **WJMax game server** (side project, team of two, Aug. 2023 – Feb. 2024). Owned the server, database and security: FastAPI on Docker (AWS EC2) with RDS MySQL behind a Cloudflare proxy/WAF, and a play-log ingestion API for a community rhythm game (∼300 average, ∼2,000 peak concurrent players during updates). Code: github.com/zoocasso/wjmax-game-server; details: zoocasso.github.io/en/work/wjmax

4. Publications (first author)

- Hyeongju Lee, Myungjin Cho, Min-Chul Lee, "Angular-to-Spatial Information Redistribution via Ray-based Inverse Mapping," *ICMV 2026*, Budapest (poster, Best Poster Award).
- Hyeongju Lee, Myungjin Cho, Min-Chul Lee, "Virtual Lens Array Synthesis from Multi-View Images via iRIM," *ICCAS 2026*, Sapporo (poster).
- Hyeongju Lee, Myungjin Cho, Min-Chul Lee, "Accelerated Origin Point Optimization for Volumetric Computational Reconstruction," *ITC-CSCC 2026*, Bangkok (oral). Code for both ITC-CSCC papers: github.com/zoocasso/origin-point-optimization-vcr
- Hyeongju Lee, Gilsu Yeo, Myungjin Cho, Min-Chul Lee, "Image Enhancement of Volumetric Computational Reconstruction Using Origin Point Optimization," *ITC-CSCC 2025*, Seoul (poster).
- Hyeongju Lee et al., "Image Enhancement of Volumetric Computational Reconstruction Using Origin Point Optimization" (in Korean), *Photonics Conference 2024*, Optical Society of Korea (poster).

5. Skills, Education & Recognition

- **Skills:** Python, SQL, FastAPI, Java Spring, SQLAlchemy, pandas, MySQL, Selenium, Docker, Harbor, Linux, AWS (EC2, RDS), Cloudflare; PyTorch, scikit-learn, OpenCV, CuPy; BWA, samtools, Picard; PyQt, openpyxl, xlwings; C++, Java, JavaScript, C#, MATLAB (basic–intermediate).
- **Education:** Master's program, Computer Science and Systems Engineering (Creative Informatics), Kyushu Institute of Technology, Apr. 2025 – Mar. 2027 (expected), GPA 3.625/4.0. B.S., Electrical, Electronic & Control Engineering, Hankyong National University, Feb. 2025, GPA 4.00/4.50.
- **Recognition & other:** ICMV 2026 Best Poster Award (Oct. 2026); Rotary Yoneyama Memorial Foundation Scholar (2026, Japanese-language interview); SQLD — SQL Developer (Oct. 2023); military service completed as Industrial Technical Personnel (Jul. 2022 – Jul. 2024). Languages: Korean (native), Japanese (business conversational), English (conversational).