

SangHyeok Lee

☎ (+82) 10-8869-2796 | ✉ john8538@inha.edu | 🌐 lexxsh | 🌐 lexxsh

Education

Inha University

B.S. IN INFORMATION AND COMMUNICATION ENGINEERING

- Total GPA of 3.75 / 4.5, Major GPA of 3.79 / 4.5

Incheon, S.Korea

Mar. 2020 - Aug. 2026

Internships

LG AI Research, Physical Intelligence Lab

AI ENGINEER INTERN

- Developed a VLM-based (Qwen3-VL; segment-reason-classify) defect-inspection pipeline, raising detection accuracy from 81% to 89% (+8 pp) and reducing false-positive (overkill) rate by 2.6 pp through data-centric refactoring
- Redesigned GT labeling criteria (verdict-blind → structured verdict-aware annotations) and scaled human-verified labels 3× (387 → 1,192), correcting systematic label noise against inspector judgments
- Experimented with diverse input conditions (reference images, defect-type hints, multi-bbox grouping) and image-token variations (depth modality, multi-image fusion), training models via supervised fine-tuning (SFT)

Seoul, South Korea

Feb. 2026 - Present

Generative Computing Lab

UNDERGRADUATE RESEARCHER

- Research on Diffusion Models Applications and Generative AI Across Diverse Domains
- Participated in research paper seminars

Inha Univ

Mar. 2025 - Mar. 2026

CVIP Lab

UNDERGRADUATE RESEARCHER

- CS231n Study and basic computer vision paper seminars

Inha Univ

Jan. 2024 - Aug. 2024

Publications

Compositional Image Synthesis with Inference-Time Scaling

CO-FIRST AUTHOR (EQUAL CONTRIBUTION)

- Minsuk Ji*, Sanghyeok Lee*, and Namhyuck Ahn

ICASSP 2026

Jan. 2026

Awards & Honors

AWARDS

- 2025.11 **1st AI Place, 2nd Overall Award**, Financial AI Challenge
- 2025.08 **3rd Award**, SW-Centered University Digital Competition: AI Division
- 2025.08 **3rd Award**, INHA AI Challenge
- 2025.06 **4th Award**, Bias-A-Thon: The Bias Response Challenge
- 2025.06 **Silver Medal**, Image Matching Challenge
- 2024.07 **1st Award**, INHA SW NET-Zero Collaborative Hackathon
- 2023.11 **3rd Award**, Innothink Makerthon
- 2023.11 **3rd Award**, Advantech AIoT Innworks Project

FSI, Dacon

SW-Univ, Dacon

Inha Univ, Dacon

SKKU CMI, Dacon

Czech Tech Univ, Kaggle

Inha Univ

Inha Univ

Advantech Kr

HONORS

- 2025 **Geumgang Hong Soon-Tak Scholarship Foundation Scholar**, Academic Excellence, Financial Need

Incheon, S.Korea

Extracurricular Activity

Naver Boostcamp(AI Tech 7th)

CORE MEMBER

- Intensive AI bootcamp covering CV, NLP, and Generative AI with hands-on implementation of classification, detection, segmentation.
- Ranked 1st in Hand Bone Segmentation and Recyclable Object Detection, 3rd in Multilingual OCR competitions (among 23-24 teams).
- Served as team moderator leading ideation, project management, and knowledge-sharing initiatives to foster collaborative problem-solving.

Naver Connect Foundation

Aug. 2024 - Feb. 2025

- Led Christopher M. Bishop's Deep Learning textbook study group and organized internal Korean sentence relation classification competition.
- Delivered special lecture on competition strategies and best practices, fostering knowledge-sharing culture within the community.

Project Experience

Financial AI Challenge [1st Place]

Financial Security Institute (FSI)

TEAM LEADER & BASE MODEL DESIGNER

Aug. 2025 - Oct. 2025

- Led team project designing financial security QA system using KT Mi:dm 11.5B LLM with Advanced RAG architecture for domain-specific knowledge retrieval.
- Built preprocessing pipeline using PDFPlumber and regex filtering for financial regulations and ISMS-P guidelines; implemented recursive text splitting (2000 chars, 500 overlap) and FAISS vector database with cosine similarity (k=10 retrieval).
- Designed reranking system selecting top-1 chunk with 0.75 score threshold and iterative self-consistency (5 iterations) for answer validation, eliminating need for Chain-of-Thought reasoning.

SummarAI: RAG-based Paper Chatbot Service

Naver Boostcamp AI Tech 7th

CORE MEMBER

Dec. 2024 - Feb. 2025

- Built end-to-end PDF-to-text pipeline supporting mathematical formulas, tables, and images using 6 specialized models (YOLO for formula detection, PaddleOCR for text, pix2text for formula recognition)
- Designed Advanced RAG pipeline with HyDE, query refinement, sentence chunking with window method, and reranking, achieving 0.75 average score on Allganize Korean RAG benchmark (180/240)
- Fine-tuned YOLO model on custom Korean+formula dataset (100 papers), improving detection accuracy from 0.52 to 0.78 for Korean papers

Recyclable Waste Object Detection [1st/24 teams]

Naver Boostcamp AI Tech 7th

TEAM MODERATOR & PIPELINE DEVELOPER

Dec. 2024

- Built detection pipeline using MMDetection library with task-optimized models (DINO, CO-DETR, Cascade R-CNN).
- Improved mAP50 from 49% to 76% by applying SR-based dataset augmentation, TTA, and WBF ensemble techniques.

Multilingual Receipt OCR [3rd/24 teams]

Naver Boostcamp AI Tech 7th

MODEL DEVELOPER

Jan. 2025

- Optimized bounding boxes using annotation tools and developed specialized models for separator line detection.
- Achieved 23%p performance improvement in DetEval metric through hard voting ensemble strategy.

AI-Generated Text Classification [6th Place]

2025 SW-Centric Univ Council

TEAM LEADER & MODEL DEVELOPER

Jul. 2025 - Aug. 2025

- Led team project to distinguish AI-generated text from human-written content, addressing media literacy challenges in the generative AI era.
- Implemented LLM-based data augmentation using Gemma2-9B and LLaMA3.1, expanding dataset from 97K to 1.13M samples (11.6x increase) with balanced 2:3 class ratio.
- Developed multi-stage training pipeline: sliding window approach for full-text learning, pseudo-labeling strategy, and confidence-based soft ensemble of 5 DeBERTa models.

Bias Mitigation Challenge [6th Place]

2025 Bias-A-Thon

INDIVIDUAL PARTICIPANT & PROMPT ENGINEER

Apr. 2025 - May. 2025

- Developed bias-free response generation system for Llama-3.1-8B-Instruct through systematic prompt engineering, addressing ethical challenges in LLM outputs.
- Designed comprehensive system prompt with persona setting ("ultra-precise analyst"), explicit bias prevention rules, and 5-stage internal analysis framework for consistent reasoning.

Skills

Back-end	Flask, Express, REST API
Front-end	Redux, React
Tech	Python, SQL, C++, Typescript
Languages	Korean (Native), English (OPIc IM2)
Certifications	ADsP (Advanced Data Analytics Semi-Professional)