

Gitaek Kwon

AI Research Scientist — Computer Vision, Deep Learning

kwongitack@gmail.com · anears.github.io · [Google Scholar](#) · [GitHub](#) · [LinkedIn](#)

RESEARCH INTERESTS

Computer vision and deep learning, focused on training and deploying models under data and compute constraints.

- **Learning under Data Constraints** — class imbalance, scarce labels, semi-supervised learning, label noise
- **OOD Detection & Robustness** — out-of-distribution detection, robustness to input distribution shift, evaluation metric design
- **Generative Data Augmentation** — text-to-image and mask-to-image synthetic data generation
- **Representation Learning** — self-supervised learning, domain generalization, deep hashing
- **Efficient Inference & Deployment** — model consolidation, CPU inference optimization, MLOps

Record. 10 international papers (ICLR ×2, MICCAI ×2, BMVC Oral, ISBI ×2, 3 SCIE journals) · 4 first-author papers (1 sole, 3 co-first) · 2 first-place finishes in international challenges · 3 patents.

EDUCATION

M.S. in Computer Software — Hanyang University, Seoul · Mar 2019 – Feb 2021

- Computer Vision Lab, advised by Prof. Jongwoo Lim · GPA 4.25 / 4.5
- Thesis: *Deep Hashing Specialized for Multi-Label Image Retrieval*

B.S. in Software — Hanyang University, Seoul · Mar 2015 – Feb 2019

- GPA 3.71 / 4.5

EXPERIENCE

VUNO Inc. — AI Research Scientist, R&D Center / Fundus Team · Mar 2021 – Aug 2026 · Seoul, Korea

Concurrently served as Industrial Technical Research Personnel (alternative military service), Mar 2021 – Mar 2024.

- **Product model development.** Developed and improved multi-label classification and multi-class segmentation models for commercial imaging software, addressing extreme class imbalance (positive rate below 1% for some classes), partially missing labels, and discriminative patterns confined to very small regions of large images. Migrated a legacy training stack to PyTorch and restructured it as multi-task learning, which reduced training from three stages to one.
- **Metric-driven redesign.** Found that an input-validity (OOD) filter selected on AUROC alone was rejecting large numbers of valid inputs. Redefined the operating metric to AUPRC and absolute false-rejection count, collected the invalid-input distribution, and rebuilt the model — **reducing false rejections from 1,359 to 21 out of 70,202 public normal images.**
- **Inference efficiency.** Consolidated a two-model serial pipeline into a single multi-output model with no loss in performance, removing one model's load and inference cost. Resolved a resolution-versus-compute trade-off architecturally with a lightweight backbone (~4.9M parameters) and a high-resolution decoder with sliding-window inference.
- **Serving and release engineering (CPU-only deployment).** Integrated OpenVINO/ONNX runtimes and profiled per-stage latency; built an asynchronous inference service (Celery + Redis, Docker) with timeout and fault isolation; built a one-click release pipeline (weight encryption → image build → smoke test → tagging) and enforced lint and unit tests on every pull request.

- **Evaluation infrastructure.** Built a **23-dataset benchmark harness (~13k images)** that scores every release candidate and blocks performance regressions automatically. Built the evaluation tool and inference backend for a multi-center reader study comparing human performance with and without model assistance.
- **Research automation.** Sole developer of a multi-agent research-automation harness (LangGraph, 8-node StateGraph) that iterates understand → plan → experiment → analyze autonomously, with a least-privilege tool matrix, git branch isolation, and three-tier structured-output fallback. Applied to an active model-development project.

PUBLICATIONS

★ sole first author · ☆ co-first author (equal contribution)

Journals (SCIE)

[J3] J. Son, J.Y. Shin, S.T. Kong, J. Park, **G. Kwon**, H.D. Kim, K.H. Park, K.H. Jung, S.J. Park. "An Interpretable and Interactive Deep Learning Algorithm for a Clinically Applicable Retinal Fundus Diagnosis System by Modelling Finding-Disease Relationship." *Scientific Reports*, 2023.

[J2] ☆ M.S. Yoon, **G. Kwon**, J. Oh, J. Ryu, J. Lim, B. Kang, J. Lee, D.K. Han. "Effect of Contrast Level and Image Format on a Deep Learning Algorithm for the Detection of Pneumothorax with Chest Radiography." *Journal of Digital Imaging*, 2023.

[J1] ☆ **G. Kwon**, J. Ryu, J. Oh, J. Lim, B. Kang, C. Ahn, J. Bae, D.K. Lee. "Deep Learning Algorithms for Detecting and Visualising Intussusception on Plain Abdominal Radiography in Children: A Retrospective Multicenter Study." *Scientific Reports*, 2020.

Conferences

[C7] S. Lee, S. Kang, I. Park, **G. Kwon**, J. Baek, D. Park. "Frequency-Balanced Retinal Representation Learning with Mutual Information Regularization." *ICLR 2026*.

[C6] J. Baek, S. Lee, **G. Kwon**, D. Park. "A Comparative Study of Machine Learning and Deep Learning for Out-of-Distribution Detection." *ISBI 2026*.

[C5] E. Kim, **G. Kwon**, J. Kim, H. Park. "Semi-supervised Segmentation through Rival Networks Collaboration with Saliency Map in Diabetic Retinopathy." *MICCAI 2024*. [code](#)

[C4] E. Kim, S. Lee, **G. Kwon**, H. Kim, J. Kim. "Multi-Stage Region-Based Neural Networks for Fine-Grained Glaucoma Screening." *ISBI 2024* — challenge paper, **1st place overall**. [code](#)

[C3] ☆ **G. Kwon**, J. Kim, H. Choi, B. Yoon, S. Choi, K.H. Jung. "Improving Out-of-Distribution Detection Performance using Synthetic Outlier Exposure Generated by Visual Foundation Models." *BMVC 2023* — **Oral**. [code](#)

[C2] ★ **G. Kwon**, E. Kim, S. Kim, S. Bak, M. Kim, J. Kim. "Bag of Tricks for Developing Diabetic Retinopathy Analysis Framework to Overcome Data Scarcity." *MICCAI 2022* — challenge paper, **1st place in all three tasks**.

[C1] J. Ryu, **G. Kwon**, M.H. Yang, J. Lim. "Generalized Convolutional Forest Networks for Domain Generalization and Visual Recognition." *ICLR 2020*.

Domestic Conferences (Korean)

[D3] ★ **G. Kwon**, J. Bae, J. Lim. "Object Counting Using Object Detection and Re-Identification in Video." *IPIU 2020*.

[D2] ★ **G. Kwon** et al. "Object Tracking Using Detection-Based Hard Negative Mining." *IPIU 2019*.

[D1] D. Yoon, **G. Kwon**, J. Lim. "Effective Data Augmentation Methods and Input Image Size for Training Deep Object Recognition Networks." *KSC 2018*.

PATENTS

- **Method and Apparatus for Providing Lesion Information** — KR 10-2023-0184784 (filed Dec 2023)

- **Method and Apparatus for Transferring the Style of an Image** — KR 10-2592666 (granted Oct 2023)
- **Object Tracking Method Using Hard Negative Mining** — KR 10-2224101 (granted Mar 2021)

AWARDS

- **1st Place, Overall** — JustRAIGS: AI Glaucoma Screening Challenge, *ISBI 2024*. 322 participants from 37 countries; 20 teams submitted in the test phase. Contributed as co-author of the winning solution [C4].
- **1st Place, All Three Tasks** — DRAC22: Diabetic Retinopathy Analysis Challenge, *MICCAI 2022*. First author of the challenge paper [C2]; 11–13 teams per task.
- **Bronze Medal** — ACM-ICPC Asia Regional Contest, Korea, 2016.

PERSONAL PROJECTS

Vision-Language Models

- **ko-chart-vlm** — Diagnosed why open VLMs fail on Korean charts, decomposed the failures into three modes, built targeted synthetic data, fine-tuned with QLoRA, and served with vLLM/AWQ. Validation accuracy **70.2% → 96.6%**; also identified a preprocessing mismatch between the evaluation and serving paths that offline benchmarks had hidden. github.com/Anears/ko-chart-vlm
- **fundus-vlm** — Fine-tuned open VLMs on public data only for grading, quantitative estimation, and structured report generation. MedGemma-4B + QLoRA reached **QWK 0.874** on DR grading, above an EfficientNet-B0 baseline at 0.861, and was the only model of the two compared that learned continuous ratio estimation stably (Pearson r 0.368 → 0.737). github.com/anears/fundus-vlm

Developer Tools

- **Claude Tower** — Terminal dashboard unifying coding-agent sessions scattered across servers, via read-only SSH scanning and tmux integration; designed to add zero API cost. TypeScript, Ink. github.com/Anears/claude-tower
- **Artview** — Cross-platform desktop gallery that indexes and full-text-searches hundreds of agent artifacts, with a custom `remote://` protocol that makes remote folders behave like local ones. Rust, Tauri 2, SQLite FTS5. github.com/Anears/Artview

SKILLS

- **Programming** — Python (primary), C/C++, TypeScript, Shell
- **Deep Learning** — PyTorch, PyTorch Lightning, TensorFlow/Keras (legacy maintenance), scikit-learn
- **Topics** — image classification and segmentation, OOD detection and robustness analysis, semi-supervised learning, class imbalance, metric and representation learning, generative augmentation (latent diffusion, LoRA), VLM fine-tuning (QLoRA)
- **Deployment** — ONNX, OpenVINO, Celery, Redis, Docker, GitHub Actions, vLLM, AWQ
- **Experiment Operations** — WandB, MLflow, Optuna, LangGraph
- **Languages** — Korean (native), English (professional working), Japanese (JLPT N3, 2024)