# Ganghui Yi

https://hughyi.github.io

✉ ganghuiy@usc.edu
in linkedin.com/in/hughyi  ⌂ github.com/hughyi

## SUMMARY

USC Computer Science master's student with a strong interdisciplinary foundation in machine learning and life sciences. Experienced in building reproducible PyTorch-based pipelines for biomedical ML, including WSI foundation-model feature integration with multi-omics and clinical metadata, memory-efficient MIL for digital pathology, and few-shot multi-omics modeling on TCGA/GEO. Focused on representation learning, robust generalization across cohorts, and interpretable modeling for therapeutic discovery.

## EDUCATION

**University of Southern California** — Aug 2025 – May 2027
*M.S. in Computer Science* — Los Angeles, California

**Sogang University** — Mar 2017 – Feb 2024
*B.S. in Computer Science and Engineering & Life Science (CGPA 3.81/4.3, **Magna Cum Laude**)* — Seoul, Korea

## SKILLS

**Languages**: Python, C/C++, Java, SQL, JavaScript, TypeScript, HTML/CSS
**Frameworks & Libraries**: PyTorch, TensorFlow, Pandas, NumPy, scikit-learn, CUDA, Hugging Face, RDKit, Matplotlib
**Technologies**: ML/DL, Multi-Omics, Digital Pathology, Biomedical Imaging, Multimodal AI, HPC, XAI, Docker, Git/GitHub, LLMs, Linux/Unix

## PROFESSIONAL EXPERIENCE

**Melady Lab**, University of Southern California — Aug 2025 – Present
*Graduate Student Researcher* — Los Angeles, California

- Built a memory-efficient WSI MIL pipeline in PyTorch using foundation-model embeddings, integrating multimodal learning with clinical metadata and compact multi-omics representations, and applying unlearning to improve cross-cohort robustness under protocol shifts.
- Developed a TCR-guided de novo binder design pipeline that initializes from natural TCR sequences and applies constraints to reduce diffusion-style off-target cross-reactivity while optimizing pMHC binding affinity.
- Engineered an LLM-based job recommender system that de-confounds candidate preference vs. employer qualification via dual-perspective modeling, and applied constraint-aware ranking with controllable policy alignment using LoRA fine-tuning.
- Implemented a reproducible molecular GNN/MPNN workflow using Chemprop and RDKit for virtual screening and hERG cardiotoxicity prediction, emphasizing scalable training, reproducible evaluation, and interpretable analysis.

**Bioinformatics & Machine Learning Lab**, Chungbuk National University — Aug 2024 – Jun 2025
*Research Intern* — Cheongju, Korea

- Developed a MAML-based few-shot meta-learning framework in PyTorch for predictive modeling on multi-omics datasets, improving cross-domain generalization in small data regimes.
- Built an end-to-end preprocessing pipeline across TCGA and GEO and constructed a Reactome-based gene/miRNA–pathway mapping to enable pathway-level modeling and attribution using XAI.
- Achieved 0.94 mean PR-AUC in the 5-shot setting on non-cancer test cohorts and nominated high-effect candidate biomarker pathways via explainable pathway attributions with cohort-level statistical filtering.

**LikeLion**, LikeLion Sogang — Mar 2021 – Dec 2021
*Full-stack Developer* — Seoul, Korea

- Architected a full-stack URL repository service, developing a responsive and scalable UI with React and building RESTful APIs using Django REST Framework for core services and FastAPI for high-performance endpoints.
- Implemented core features including URL saving, sharing, and a dynamic keyword search to streamline information retrieval and enhance user collaboration, driving a 25% increase in user engagement and monthly active collaborations.

## PUBLICATIONS

**Journal / Conference Publications**

Ju, J.*, **Yi, G.***, & Jo, K.[#] PathwayMAML: pathway-aware meta-learning for cancer-to-non-cancer transfer in few-shot disease classification.
*International Conference on Intelligent Systems for Molecular Biology (ISMB) (Under review)* — **2026**

Kan, B.*, Yang, W.*, Nguyen, E., **Yi, G.**, Yi, B., Yu, C., & Liu, Y.[#] De-conflating Preference and Qualification: Constrained Dual-Perspective Reasoning for Job Recommendation with Large Language Models.
*ACL Rolling Review (ARR) (Under review)* — **2026**

**Yi, G.***, Ju, J.*, & Jo, K.[#] Few-Shot Learning for Multi-Omics Disease Classification with a MAML-Based Model.
*KIPS Transactions on Software and Data Engineering (KTSDE)* — **2025**
* indicates equal contribution, [#] indicates corresponding author.

**Oral Presentations**

Few-Shot Learning for Multi-Omics Disease Classification with a MAML-Based Model.
*9th International Conference on Medical and Health Informatics (ICMHI)*, Kyoto, Japan — **2025**

## Teaching

**9307261 Current Topics on Machine Learning**, Chungbuk National University — Spring 2025
*Teaching Assistant* (with Prof. Kyuri Jo)

**5110134 Artificial Intelligence**, Chungbuk National University — Fall 2024
*Teaching Assistant* (with Prof. Kyuri Jo)

## Honors & Awards

**Outstanding Startup Idea Award**, Sogang University — Fall 2021
**1st Place**, LikeLion Hackathon — Fall 2021
**1st Place**, LikeLion Ideathon — Summer 2021

## Standardized Test Scores

**iBT TOEFL**: 111 (Reading: 30/30, Listening: 29/30, Speaking: 28/30, Writing: 24/30)
**GRE General Test**: Verbal 161/170 (86%), Quant 170/170 (92%), Analytical Writing 4.5/6.0 (83%)