Soyon Park

Email: soyon_park@korea.ac.kr
Phone Number: +82-10-2047-4762
Google Scholar (Link) **Linkedin** (Link)

Education

2024.03 – Korea University, Seoul, South Korea
Master of Science in Engineering:
- Major in Computer Science & Engineering
- GPA: 4.5 /4.5 (4.0/4.0)
- Data Mining & Information Systems Laboratory (advisor Prof. Jaewoo Kang)

2018.03 – 2023.02 Ewha Womans University, Seoul, South Korea
Bachelor of Science:
- Major in Life Sciences
- Double Major in Bioinformatics
- Minor in Computer Science & Engineering
- GPA: 3.99 /4.3 (3.849/4.0)
- Graduated Magna Cum Laude

Research Publications

1. **[AAAI 2025]** (Link)
   CRADLE-VAE: Enhancing Single-Cell Gene Perturbation Modeling with Counterfactual Reasoning-based Artifact Disentanglement.
   Seungheun Baek*, **Soyon Park***, Chok Yan Ting, Junhyun Lee, Jueon Park, Mogan Gim, Jaewoo Kang. (2025).

2. [**ISMB/ECCB 2025 / Bioinformatics]** (Link)
   GPO-VAE: Modeling Explainable Gene Perturbation Responses utilizing GRN-Aligned Parameter Optimization.
   Seungheun Baek*, **Soyon Park***, Chok Yan Ting, Mogan Gim, Jaewoo Kang. (2025).

3. **[Bioinformatics Advances]** (Link)
   BADGER: biologically-aware interpretable differential gene expression ranking model.
   Hajung Kim*, Mogan Gim*, Seungheun Baek, **Soyon Park**, Sunkyu Kim, Jaewoo Kang. (2025).

4. [**BIBM 2025**] (Link)
   CoTox: Chain-of-Thought based Molecular Toxicity Reasoning and Prediction.
   Jueon Park, Yein Park, Minju Song, **Soyon Park**, Donghyeon Lee, Seungheun Baek, Jaewoo Kang. (2025).

5. [Under Review]
   Transductive Learning for Out-of-Distribution Molecular Property Prediction.
   Kiwoong Yoo, Hajung Kim, **Soyon Park**, Junseok Choe, Sunkyu Kim, Jaewoo Kang. (2025).

6. [Under Review] (Link)
   HiRef: Leveraging Hierarchical Ontology and Network Refinement for Robust Medication Recommendation.
   Yan Ting Chok, **Soyon Park**, Seungheun Baek, Hajung Kim, Junhyun Lee, Jaewoo Kang. (2025).

7. **[ISMB 2024 / Bioinformatics]** (Link)
   MolPLA: a molecular pretraining framework for learning cores, R-groups and their linker joints.
   Mogan Gim*, Jueon Park*, **Soyon Park**, Sanghoon Lee, Seungheun Baek, Junhyun Lee, Ngoc-Quang Nguyen, Jaewoo Kang. (2024).

   Co-first authors*

Ongoing Projects

Publications

1. **[ISMB/ECCB 2025 poster, Ongoing Project]**
   SPARTHA: Enhancing Spatial Gene Expression Prediction with Artifact Disentanglement from Histology Images. (Link)
   - developed a multi-modal conditional VAE that integrates histology images and spatial transcriptomics to generate spatial gene expression from tissue morphology while disentangling batch effects

2. **Development of Multi-Cancer Biomarkers via AI-Driven Proteomic Analysis of Plasma MS/MS Data.**
   - proposed a transformer-based, ID-free representation learning approach on peptide DIA MS/MS spectra to identify cancer-specific embeddings and classify ten cancer types

Research Interests

Bioinformatics	Single cell omics, Spatial Transcriptomics, Multi-omics, Perturbation, Protein
Medical Imaging	PET/CT Image Denoising, Pathology Imaging, 3D Image Reconstruction
Cheminformatics	Molecule Generation, Molecule Optimization, Drug Response
Deep Learning	Computer Vision, Generative Models, Graphs
Causal Learning	Causal Inference, Counterfactual Reasoning, Disentangled Representation Learning

Work Experience

Internship

2025.09 – 2026.02 **Research internship**
- École polytechnique fédérale de Lausanne (EPFL), Lausanne, Switzerland
- Machine Learning for Biomedicine (MLBio) Lab (Prof. Maria Brbić)

2023.03 – 2023.07 **Research internship** (Paper) (PPT)
- Seoul National University Hospital, Seoul, South Korea
- Bio-Medical Informatics (BMI) Lab

: developed methods for understanding human disease via deep learning
- Multi-omics drug response prediction via Graph based Neural Network
- Spatial Transcriptomic data analysis via deep learning

2022.07 – 2022.09
2021.12 – 2022.02 **Research internship** (Paper) (PPT)
- Seoul National University College of Medicine, Seoul, South Korea
- Department of Biomedical Sciences
- Functional & Molecular Imaging System Lab

: Brain PET image Denoising
- Supervised/Semi-supervised(Noise2Noise) Learning
- 2D/3D U-Net, 3D image reconstruction from 2D image

2021.07 – 2021.09 **Research internship**
- Ewha Womans University, Seoul, South Korea
- Department of Life Sciences

- Cancer Biology and Genomics Lab

: Relationship between senescence and cancer

  - Western Blot, Immunofluorescence, β-gal staining

Project

2021.06 – 2021.12 **COVID-19 Drug Reposition via Deep Learning** (Poster)

- Poster presented in KSMCB 2021(The Korean Society for Molecular and Cellular Biology) Conference

: developed a model to predict COVID-19 drug candidates using CNN model trained by COVID-19 related compounds according to SMILES notation

Awards

2025.05 **BioASQ 13B Phase B: Biomedical Semantic QA (Winner)** (Link)

- Organizer: CLEF 2025 BioASQ Workshop
- developed an ensemble of large language models (LLMs) for large-scale biomedical semantic indexing and question answering competition (Team DMIS-KU)

2023.09 **2023 Artificial intelligence competition using herbal medicine experimental information data Best award (Winner)** (Link)

- Organizer: National Institute for Korean Medicine Development & Ministry of health and welfare
- Analyzing herbal medicines using network pharmacology and artificial intelligence models (Team 2Park)
  - Drug-Target Interaction (DTI) Prediction

Scholarship

2024.03 – 2026.03 **Presidential Science Scholarship** (Link) **($25,000)**

- Organizer: Ministry of Science and ICT

: scholarship for outstanding science and engineering students

2024.03 – 2024.12 **BK(Brain Korea)21FOUR Scholarship ($4,000)**

Programming skills

Languages	C/C++, Python
Frameworks & Libraries	Matplotlib, Numpy, Pandas, Scikit-learn, RDKit, PyTorch, PyTorch Geometric, Scanpy, Networkx, PyTorch Lightning