

## EDUCATION

---

- **Graduate School of AI, KAIST** Seoul, Korea  
*PhD candidate advised by Sungsoo Ahn; GPA: 4.21/4.3* Mar. 2024 – Present
- **Korea Advanced Institute of Science and Technology** Daejeon, Korea  
*Major in Computer Science, Electrical Engineering; GPA: 3.78/4.3* Mar. 2018 – Mar. 2024

## SELECTED PUBLICATIONS

---

- [Under Review] **Learnable Tokenization for DNA Foundational Models**  
*Taewon Kim, Hyomin Kim, Jihwan Shin, Youngmok Jung, Jonghoon Lee, Won-Chul Lee, Insu Han, Sungsoo Ahn*  
Motivated by the lack of study in tokenization methods specific for DNA sequences which lack natural separators such as spaces, proposed a learnable tokenization method for DNA sequence.
- [Under Review] **Learning Flexible Forward Trajectories for Masked Molecular Diffusion**  
*Hyunjin Seo\*, Taewon Kim\*, Sungsoo Ahn*  
Introduced a flexible noise scheduling framework for masked diffusion models, significantly enhancing the validity of masked diffusion models in molecules, which have previously shown to have diminishing results.
- [ICLR 25] **REBIND: Enhancing Ground-state Molecular Conformation Prediction via Force-Based Graph Rewiring**  
*Taewon Kim\*, Hyunjin Seo\*, Sungsoo Ahn, Eunho Yang*  
Developed a graph rewiring technique incorporating physical force into attention bias for accurate ground-state molecular conformation prediction.

## PUBLICATIONS IN GENERAL DOMAIN

---

- [WACV 26] **Towards Reliable Test-Time Adaptation: Style Invariance as a Correctness Likelihood**  
*Gilhyun Nam\*, Taewon Kim\*, Eunho Yang*
- [NeurIPS 24 Workshop] **AdapTable: Test-Time Adaptation for Tabular Data via Shift-Aware Uncertainty Calibrator**  
*Changhun Kim\*, Taewon Kim\*, Seungyeon Woo, June Yong Yang, Eunho Yang*
- [NeurIPS 22] **NOTE: Robust Continual Test-time Adaptation Against Temporal Correlation**  
*Taesik Gong, Jongheon Jeong, Taewon Kim, Yewon Kim, Jinwoo Shin, Sung-Ju Lee*

## PROJECTS

---

- **AI-Specialized Foundation Model Development Project (Oct. 2025 - Present)**: Participating in project to surpass AlphaFold, developing a general model capable of predicting both apo and holo structures of proteins. Within it, currently developing a I-JEPA inspired protein representation learning framework
- **Genetic Language Model for Cancer Subtype Classification (Jul. 2025 - Present)**: Based upon the DNA language model developed via learnable tokenization method, incorporating it with real-world patient genomics data to create an AI framework to classify patients' cancer subtypes.

## AWARDS

---

- **22 QualComm Innovative Fellowship Korea Finalist**: Oct. 2022
- **Deans' List for Electrical Engineering**: Spring 2022
- **K-Startup Challenge Finalist**: Fall 2020

## SKILLS

---

- **Programming**: Python, Pytorch, Tensorflow, Tensorflowlite, Android Studio, React
- **Language**: English (Served as translator in Korea Army), Chinese