Curriculum Vitae

Personal

Name	**Si Been Kim**
Phone number	**+82-10-2501-8439**
Email	**bioceo78@korea.ac.kr**

Education and Qualifications

Mar 2015 - Feb 2018
Gyeonggi Science High School, Suwon

Mar 2018 - Feb 2025
B.S. Biomedical Engineering (GPA - Major: 3.93 / 4.50, Overall: 3.62 / 4.50)
Korea University, Seoul

Work experience

Jan 2024
Intern
Letsur, Seoul

Ongoing Task 1. Development of a Binary Classification Deep Learning Model for Predicting Progressive Vertebral Bone Collapse after Osteoporotic Vertebral Fracture (OVF), with the Aim of Forecasting the Need for Surgery within the Next 5 Years
Ongoing Task 2. Development of a Model to Enhance Classifier Performance in Long-Tail Situations through Data Expansion

Jun 2023 - Nov 2023
Student Researcher
Laboratory of Advanced Neuroimaging Biomarker Research (LANIB), CCIDS, Severance Hospital, Seoul

Park, C.J.*, Choi, S.H.*, Kim, D.I., **Kim, S.B.**, Han, K.H., Ahn, S.S., Lee, W.H., Choi, E.C., Keum, K.C., Kim, J.: MRI radiomics may predict early tumor recurrence in patients with sinonasal squamous cell carcinoma. *European Radiology*, Nov 2023. (*: equal contribution)

1. **3D Brain MRI Preprocessing.** Executed comprehensive preprocessing steps including DICOM to NIfTI conversion, sequence extraction, reorientation, isovoxel resampling, N4 bias field correction, and co-registration. Implemented skull stripping using HD-BET on T1CE sequences.
2. **Glioma Segmentation**. Performed four-class (non-tumor, edema, necrosis, enhancing tumor) glioma segmentation using a nnUNet model, fine-tuned for Severance Hospital data.
3. **Feature Extraction & Observer Stability Assessment.** Conducted feature extraction from masked voxel regions in brain MRI scans. Evaluated inter-observer reliability using Intraclass Correlation Coefficient (ICC) and Concordance Correlation Coefficient (CCC).

Sep 2023 - Oct 2023
Student Researcher
Laboratory of Advanced Neuroimaging Biomarker Research (LANIB), CCIDS, Severance Hospital, Seoul

Project: Comparative Study of IDH-Wildtype and IDH-Mutant Gliomas
Advisor: Prof. Sung Soo Ahn, Yae Won Park
Objective. To compare IDH-wildtype glioma patients, IDH-mutant glioma patients, and control subjects in terms of low ADC peak values in the subventricular zone (SVZ).
Methodology. Conducted lateral ventricle segmentation and defined SVZ by selecting a range of 3 to 5 voxels lateral to each lateral ventricle. Performed segmentation to isolate tumor regions, essential for accurate SVZ estimation. Removed non-brain areas from the SVZ VOI. Employed double Gaussian fitting to ADC values within SVZ VOI to extract lower ADC peak values for comparative analysis.

Apr 2023 - Jun 2023
Student Researcher
Laboratory of Advanced Neuroimaging Biomarker Research (LANIB), CCIDS, Severance Hospital, Seoul

Project: Automated Preprocessing and Segmentation Pipeline
Development of Preprocessing Class. Created for complete automation of MRI preprocessing pipeline.

Included capabilities for generating quality control images and robust error handling.
Development of Segmentation Class. Created to automate MRI segmentation, with initial inspection and sorting of available sequences for model selection. Devised a staged prediction strategy for cases with incomplete MRI sequences.
Achievements. Enhanced MRI processing efficiency and accuracy. Facilitated adaptable model selection for segmentation, handling various MRI datasets effectively.
Inspirational Impact. This project inspired the formulation of the "Advanced Multi-modal Brain Tumor Segmentation in MRI with Novel Handling of Missing Modalities" project.

Sep 2021 - Sep 2021

2022 KAKAO BLIND RECRUITMENT

2nd Code Test - Game Matching Algorithm Design:
- **Objective.** Design a real-time server-interacting process for game matching, ensuring fair skill evaluation and minimizing wait times for users.
- **Challenge.** Address the issue of users encountering opponents with significantly higher skills, leading to dissatisfaction and potential game abandonment.
- **Matching Goal.** Ensure that the matchmaking results, based on users' inherent skill levels (unknown to the participant), closely align with the assigned ratings, while keeping user wait times minimal.
- **Additional Complexity.** Consideration of a 5% probability of encountering abusers who intentionally fake their abilities to match with beginners.
- **Inspiration for Future Work.** This experience inspired the "Neuroimmune-Inspired Q-learning for AI Data Label Accuracy" project, focusing on anomaly detection in a related context.

Sep 2021 - Sep 2021

2022 KAKAO BLIND RECRUITMENT

1st Code Test - Algorithmic Proficiency:
- **Tasks.** Tackled various algorithmic challenges including hash data structure utilization, base conversion and prime number identification, DFS and bit masking, and recursive function-based exhaustive search problems.
- **Languages Used.** Employed C++, Python, and Swift, choosing each based on the complexity and nature of the problem.
- **Outcome.** Successfully passed the test, recognized as one of the most challenging algorithm code tests among Korean tech companies.
- **Impact on AI Skills.** This experience laid the foundation for efficiently constructing trainer classes, model classes, and data loading/preprocessing classes in AI projects.

Feb 2016 - Aug 2017

R&E Student

B-ICT Lab, Bio-Mechatronic Engineering, SKKU, Suwon

Graduation thesis: "Influence of Abnormal Foot Progression Angle on Adolescent Knee", 2017.
Advisor: Prof. Jung Hwan Moon
Objective. Investigate the impact of abnormal gait, specifically varying foot progression angles, on knee joints in adolescents.
Study Design. Conducted a comparative analysis involving 10 male adolescents performing three types of gait: Out-toeing (O group), Normal (N group), and In-toeing (I group). Measured Knee Adduction Moment (KAM) and Knee Internal Rotation Moment during these gaits.
Key Findings. Observed the smallest KAM in Out-toeing and the largest in In-toeing, suggesting a higher risk of arthritis in In-toeing gait. Found that Knee Internal Rotation Moment decreased in the order of Out-toeing, Normal, In-toeing, indicating compensatory mechanisms in play.
Conclusion. The study provided insights into how abnormal walking patterns in adolescence could lead to knee joint changes and potentially influence the risk of degenerative joint diseases.

Coursework

Jan 2024 - Mar 2024

Basic Study

Data and Visual Analytics (DAVIAN) Lab, KAIST AI

Ongoing Study. Linear Algebra / Machine Learning / Deep Learning

Dec 2023 - Dec 2023

Biomedical AI II

Biomedical Engineering, Korea University

Project: Optimizing Training Procedures for UNet-Based Brain Tumor Segmentation
Objective. To enhance the training efficiency and effectiveness of a UNet family segmentation model using the BraTS18 dataset by implementing strategic cropping techniques based on tumor frequency

and location data.
Preprocessing and Analysis. Implemented bounding box cropping, N4 Bias Field Correction, and brain-masked z-normalization on the dataset. Conducted shape analysis to understand the absolute and relative dimensions of tumor bounding boxes. Performed location analysis including tumor distribution, center coordinate distribution, and density estimation using Gaussian Kernel for tumor centers.
Innovative Training Approach. Hypothesized that cropping more tumor-frequent volumes (guided by Tumor Center Density Estimation) rather than random cropping would enable the model to learn critical tumor features more efficiently. Validated this hypothesis with initial patch heat map comparisons between random and kernel density estimation (KDE) guided cropping. Tested various patch sizes derived from shape analysis and UNet architecture constraints to balance focus on tumor regions and overall context.
Experimental Setup. Executed a 4-fold cross-validation comparing KDE-guided cropping (with patch sizes (80, 96, 80), (96, 112, 96), (112, 128, 112)) against random sampling (patch size (112, 128, 112)).
Results. Observed a trade-off between validation loss and dice loss with varying patch sizes, indicating a need for balancing cross-entropy and dice loss weights. Noted significantly better performance in initial epochs for KDE-guided cropping over random cropping. However, the model trained with random cropping showed continued improvement beyond 200 epochs, eventually matching the performance of the KDE-guided model.
Discussion. Proposed a phased training approach - starting with KDE-guided cropping for early phase to focus on critical features, then transitioning to random cropping for generalization.
Conclusion and Future Work. Suggests a novel, phased approach to training segmentation models, combining targeted learning with broad generalization for improved performance. Future work will focus on optimizing the transition between guided and random cropping phases to maximize model efficiency and accuracy.

Sep 2023 - Nov 2023

Biomedical AI II

Biomedical Engineering, Korea University

Project: Advanced Multi-modal Brain Tumor Segmentation in MRI with Novel Handling of Missing Modalities
Objective. To develop a sophisticated model for segmenting brain tumors into four categories (non-tumor, edema, necrosis, enhancing tumor) in 3D brain MRI images (T1CE, T1, T2, T2-FLAIR), specifically addressing the challenge of missing modalities in real-world scenarios.
Problem Formulation. Focused on robust segmentation capable of adapting to missing MRI modalities, a common issue in clinical settings.
Architecture Development. Developed an architecture combining attention mechanisms (inspired by A2FSeg [2]) and shared encoders (from ShaSpec [3]) within the nnUNet [1] framework. Emphasized the separation of shared and modality-specific features, turning off deep supervision for simplicity and clearer analysis.
Logical Flow.
- **A2FSeg Analysis.** Recognized that A2FSeg's state-of-the-art (SOTA) results might stem more from nnUNet's settings than its novel architecture. Adapted its attention block concept for balancing modal contributions, revising the provided code for a solid baseline.
- **ShaSpec Adaptation.** Modified A2FSeg code for ShaSpec's architecture, focusing on shared and specific feature separation.
- **Core Logic.** Questioned the approach of averaging features for shared representation and the handling of modality-specific features. Proposed a more logical sequence: separate shared and specific features, apply attention to specific features, and then combine them with shared features for final output.
Generator Integration. Inspired by CollaGAN [4], introduced a generator to synthesize specific features for any missing modality, training it to generate these features based on the shared feature average and modality. This approach offers an intuitive solution to the missing modality problem.
Current Status and Future Plans. Initial model performance indicated a need for further optimization. Considering cascade training with binary segmentation for a more logical approach. Dedicated to refining the model, focusing on aspects discussed and planning for extensive testing and optimization during the winter break.
References.
[1] Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. In: Nature methods **18**(2), 203–211 (2021)
[2] Wang, Z., Hong, Y.: A2FSeg: Adaptive Multi-modal Fusion Network for Medical Image Segmentation. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2023, October 2023, pp. 673-681.
[3] Wang, H., Chen, Y., Ma, C. Avery, J., Hull, L., Carneiro, G.: Multi-modal Learning with Missing Modality via Shared-Specific Feature Modelling. In: CVPR 2023.

[4] Lee, D.W., Kim, J.Y., Moon, W.J., Ye, J.C.: CollaGAN: Collaborative GAN for Missing Image Data Imputation. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR*), 2019.

Dec 2023 - Dec 2023

Introduction to Neural Engineering

Biomedical Engineering, Korea University

Project: Neuroimmune-Inspired Q-learning for AI Data Label Accuracy
Objective. Create a neuroimmune system-inspired AI model using Q-learning and MLP to enhance data label accuracy in machine learning.
Methodology.
- **Inspiration.** Integrated concepts from neuroimmune system functioning, particularly microglia roles and maternal-fetal development, into AI.
- **Mother-Daughter MLP Model.** Designed a dual MLP architecture, where the mother model trains the daughter model, simulating maternal influence in neural development.
- **Microglia Model.** Incorporated a Microglia-like decision-making component using Q-learning to evaluate and adjust to the reliability of training data labels.
Innovative Aspects. Pioneered a unique approach by combining neuroscience principles with AI, particularly in addressing data label infallibility. Demonstrated adaptability and resilience in the model, mirroring neuroimmune system responses.
Results. Achieved enhanced model performance in scenarios with varying degrees of label accuracy. Proved the model's capacity to adaptively learn and improve decision-making in complex data environments.
Significance. Advanced AI training methodologies by introducing neuroscientific insights. Contributed a novel perspective to the field of AI, emphasizing the importance of biological inspiration in machine learning.

Nov 2023 - Nov 2023

Biomedical AI II

Biomedical Engineering, Korea University

Project: MMSE Score Regression Using Cortical Thickness Data
Objective. To predict Mini-Mental State Examination (MMSE) scores from cortical thickness data using various regression models.
Methods Employed. Kernelized Support Vector Regression (KSVR), Regression Tree, Generalized Additive Model (GAM)
Approach. Utilized iterative cross-validation (CV) to choose optimal hyperparameters for each method. Tested two variations for each method: considering 'napoe4' feature as a dummy variable and without it. Developed a total of six models through this process.
Key Findings. The optimal model was identified as KSVR with the 'napoe4' feature considered as a dummy variable. This model demonstrated superior performance on the test set compared to the other five models.

Nov 2023 - Nov 2023

Introduction to Neural Engineering

Biomedical Engineering, Korea University

Project: Multi-Layer Perceptron for MNIST Dataset Classification
Methodology. Implemented a multi-layer perceptron for classifying the MNIST dataset, using self-derived softmax activation in the output layer and self-derived sigmoid activation in other layers.
Challenges and Improvements. Addressed numerical stability issues in sigmoid and softmax functions. Suggested separating regularization from cross-entropy loss for more flexible code.
Optimization Techniques. Implemented the Adam optimizer and Xavier weight initialization for improved model performance from scratch. Conducted extensive hyperparameter tuning, including the selection of an optimal batch size and layer complexity, achieving high test accuracy.
Achievements. Developed a model demonstrating 99% accuracy on the MNIST test dataset. Critically analyzed and proposed improvements to the assignment's given code structure for enhanced efficiency and performance.

Sep 2022 - Dec 2022

Deep Learning

Computer Science and Engineering, Korea University

Project: ResNet50 for CIFAR-10 classification
Methodology. Configured static values for the residual blocks. Integrated these blocks into a 4-layer ResNet50 architecture.
Achievement. Successfully achieved a test accuracy of over 80% on the CIFAR-10 dataset, demonstrating the effectiveness of the ResNet50 configuration.

Sep 2022 - Dec 2022

Deep Learning

Computer Science and Engineering, Korea University

Project: Multi-Layer Perceptron Design for CIFAR-10 classification
Methodology. Implemented backpropagation using self-derived formulas based on linear algebra principles. Systematically tuned hyperparameters through validation processes.
Visualization and Results. Visualized the weights of the best-performing model, enhancing understanding of the model's decision-making process.

English Proficiency

TOEIC 975 / 990 (Valid until 2025/07/09)