Taehyo Kim

Research Focus **Interpretable ML / DL, High-Dimensional Statistics, Multi-Modal Representation Learning**

Contact Information
Vancouver, Canada V4A9X3 taehyo97@gmail.com
Astoria, New York 11106

Education

New York University, New York, NY
Ph.D. in Biostatistics — Expected January 2027
- Dissertation: *Modern Statistical Machine Learning Approaches for Alzheimer's Disease Research*
- Advisor: Dr. Hai Shu
- Dissertation Committee: Dr. Yang Feng, Dr. Wen Zhou

M.S. in Computer Science — June 2022

University of Toronto, Toronto, ON
B.A.Sc in Computer Engineering, minor in Biomedical Engineering — June 2020

Awards

Student Paper Award (Runner-up), ASA Statistics in Imaging	2024
DataFest Finalist, ENAR	2024
Student Travel Award, American Statistical Association	2024
PhD Fellowship Award, New York University	2022
Certificate for Artificial Intelligence Engineering, University of Toronto	2020
International Summer Research Award, University of Toronto	2019
Gold Award, The Duke of Edinburgh's Award	2015

Skills

Programming Languages: Python, R, C++, C, C#, Java, MATLAB, Shell
Frameworks: PyTorch, TensorFlow, Keras, OpenCV, Git, Spark, Hadoop, Numba, JAX
Machine Learning: GLM, PCA, CCA, Random Forest, MLP, CNN, U-Net, W-Net, Transformers, Masked Autoencoders, Video Vision Transformers, Diffusion Models, Self-Supervised Learning

Professional Experience

Graduate Student Researcher — 2020 - present
Hai Shu Lab, NYU GPH — New York, NY
- Developed spatial false discovery rate control methods for neuroimaging, adapting unsupervised W-Net architecture and fully connected Hidden Markov Random Field.
- Designed L0 penalized sparse canonical correlation analysis for interpretable feature fusion.
- Proposed LLM-guided survival modeling with knowledge-augmented penalization for improved feature selection.

Research Assistant — 2024 - present
Biofeedback Intervention Technology for Speech Lab, NYU Steinhardt — New York, NY
- Applied self-supervised learning (ViViT-based BYOL, VideoMAE) to discover unique tongue shapes in clinically inaccurate /r/ pronunciations in children with speech sound disorders.

Statistical Fellow — Summer 2024
Biostatistical Collaboration and Consultation Core, NYU GPH — New York, NY
- Prepared statistical analysis plans for clients and conducted statistical data analyses to support manuscript development.

Research Assistant — Summer 2020
Multimedia Laboratory, UofT — Toronto, ON

- Applied deep learning to classify histological tissue types and co-authored a large-scale survey on computational pathology analyzing over 800 papers.

Research Assistant Summer 2019
N.1 Institute for Health, National University of Singapore Kent Ridge, Singapore
- Wrote MATLAB functions to automate cleaning and quality checks for a 76TB neural dataset.

Software Developer Intern May 2018 - May 2019
Epson Markham, ON
- Released the Android and Windows Software Development Kit (SDK) designed for augmented reality smart-glasses, following A/B and unit testing standards.

Publications

Kim, T., Jia, Q., de Leon, M. J., Shu, H. (2025). A False Discovery Rate Control Method Using a Fully Connected Hidden Markov Random Field for Neuroimaging Data. arXiv preprint arXiv:2505.20688. Under Review at *Medical Image Analysis*. [Paper] [Code]

Lai, A., **Kim, T.**, Dahlen, A., Lomas, T. (2025). A Global Understanding of Work Enjoyment and Human Wellbeing. Under Review at *Nature Human Behavior*.

Kim, T., Shu, H., Jia, Q., de Leon, M. J. (2024). DeepFDR: A Deep Learning-based False Discovery Rate Control Method for Neuroimaging Data. Proceedings of Machine Learning Research, 238, 946–954. [Paper] [Code]

Tang, T., Chen, Y., **Kim, T.**, Shu, H. (2024). UKAN-EP: Enhancing U-KAN with Efficient Attention and Pyramid Aggregation for 3D Multi-Modal MRI Brain Tumor Segmentation. arXiv preprint arXiv:2408.00273. Under Review at *BMC Medical Imaging*. [Paper] [Code]

Hosseini, M. S., Bejnordi, B. E., Trinh, V. Q., Chan, L., Hasan, D., Li, X., Yang, S., **Kim, T.**, Zhang, H., Wu, T., Chinniah, K., Maghsoudlou, S., Zhang, R., Zhu, J., Khaki, S., Buin, A., Chaji, F., Salehi, A., Nguyen, B. N., Samaras, D., ... , Plataniotis, K. N. (2024). Computational Pathology: A Survey Review and the Way Forward. Journal of Pathology Informatics, 15, 100357. [Paper]

Manuscripts in Preparation

Kim, T., Shu, H. L0-IPLS: An L0 Penalized Sparse Canonical Correlation Analysis Method with Application to High-Dimensional Imaging-Omics Data.

Chen, Y.*, **Kim, T.***, Chen, Z., Patippe, C. Causal Determinants of Blood Pressure Control among US Adults with Hypertension: A Data-Driven Causal Graphical Learning, NHANES 2013 to 2023

Eads, A., Benway, N., **Kim, T.**, McFee, B., Preston, J., Shu, H., McAllister, T. Enhancing AI-based Speech Therapy through Acoustic to Articulatory Mapping.

Chen, Y., **Kim, T.**, Shu, H., Feng Y. Transfer-guided Conditional Score-based Diffusion Network for Replenishment Sampling Imputation.

Kim, H., Cardoso, D. d. M., Kayahara, G. M., **Kim, T.**, Shu, H., Bernabé, D. G., Ye, Y. Pre-treatment pain phenotypes and their association with disease progression and post-treatment pain in head and neck cancer.

Talks and Presentations

"A Global Understanding of Work Enjoyment and Human Wellbeing"
Poster, Joint Statistical Meetings, Nashville, TN August 2025

"Causal Determinants of Blood Pressure Control among US Adults with Hypertension: A Data Driven Causal Graphical Learning, NHANES 2013 to 2023"

Poster, ENAR Spring Meeting, New Orleans, LA — March 2025

"DeepFDR: A Deep Learning-based False Discovery Rate Control Method for Neuroimaging Data"
Oral, Joint Statistical Meetings, Portland, OR — August 2024

"DeepFDR: A Deep Learning-based False Discovery Rate Control Method for Neuroimaging Data"
Poster, International Conference in Artificial Intelligence and Statistics, Valencia, Spain — May 2024

"Enhancing AI-based Speech Therapy through Acoustic to Articulatory Mapping"
Poster, AI Research Symposium: Bridging AI Innovation and Societal Impact
New York, NY — April 2024

"Machine Learning-driven Risk Factor Identification on Post-2013 Blood Pressure Control Decline in Hypertensive Populations"
Oral, ENAR Spring Meeting, Baltimore, MD — March 2024

Teaching Experience

Teaching Assistant, New York University, New York, NY

Applied Bayesian Analysis in Public Health (GPH-GU 2272/3372)	Fall 2024
Applied Survival Analysis (GPH-GU 2368/3368)	Spring 2024
Statistical Inference (GPH-GU 3225)	Fall 2023

Graduate Student Mentor, New York University, New York, NY

Pathways into Quantitative Aging Research Summer Program	Summer 2024
Pathways into Quantitative Aging Research Summer Program	Summer 2022

Membership and Affiliations

Member, American Statistical Association	2022 - present
Member, Eastern North American Region of the International Biometric Society	2022 - present
Member, Korean International Statistical Society	2024 - present
Journal Reviewer, The American Statistician	Aug 2025