

Taehyo Kim

RESEARCH FOCUS	Interpretable Machine Learning, Representation Learning, High-Dimensional Statistics		
CONTACT	New York, NY	taehyo97@gmail.com	website — scholar — github — linkedin
EDUCATION	New York University , New York, NY		
	Ph.D. in Biostatistics		Expected Spring 2027
	M.S. in Computer Science		2022
	University of Toronto , Toronto, ON		
	B.A.Sc in Computer Engineering, Biomedical Engineering minor		2020
PUBLICATIONS	Kim, T. , Shu, H., Jia, Q., de Leon, M. J. (2024). DeepFDR: A Deep Learning-based False Discovery Rate Control Method for Neuroimaging Data. In Proceedings of the 27th International Conference on Artificial Intelligence and Statistics (AISTATS), PMLR 238: 946–954. [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]		
	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 <i>Medical Image Analysis</i> . [Paper] [Code]		
	Lai, A., Kim, T. , Dahlen, A., Lomas, T. (2025). A Global Understanding of Work Enjoyment and Human Wellbeing. Under Review at <i>Nature Human Behavior</i> .		
	Tang, T., Chen, Y., Kim, T. , Shu, H. (2025). 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 <i>BMC Medical Imaging</i> . [Paper] [Code]		
SKILLS	Programming Languages: Python, R, C/C++/C#, SQL, Java, Shell, MATLAB, SAS Frameworks: PyTorch, TensorFlow, JAX, OpenCV, Spark, Hadoop, Numba, Git, Multiprocessing Machine Learning: Transformers, Masked Autoencoders, VideoMAE, CNNs, Video Vision Transformers, U-Nets, Diffusion Models, Graphical Modeling, Large Language Models		
PROFESSIONAL EXPERIENCE	Graduate Student Researcher 2020 - present		
	Hai Shu Lab, NYU GPH New York, NY <ul style="list-style-type: none">Designed computationally efficient, spatial false discovery rate control methods for ultra-high dimensional neuroimaging, adapting unsupervised W-Net and dense graphical models.Developed L0 penalized sparse canonical correlation analysis for interpretable feature fusion with application to imaging-omics data.Proposed LLM-RAG guided survival modeling with knowledge-augmented penalization for improved feature selection.		
	Deep Learning Research Assistant 2024 - present		
	Biofeedback Intervention Technology for Speech Lab, NYU Steinhardt New York, NY <ul style="list-style-type: none">Applied transformer-based self-supervised learning (ViViT, BYOL, VideoMAE) for representation learning on ultrasound video data, leveraging GPU-accelerated PyTorch pipelines for efficient embedding extraction.		

	Statistical Fellow Biostatistical Collaboration and Consultation Core, NYU GPH <ul style="list-style-type: none"> Prepared statistical analysis plans for clients and conducted statistical data analyses to support manuscript development. 	Summer 2024 New York, NY
	Deep Learning Research Assistant Multimedia Laboratory, UofT <ul style="list-style-type: none"> Applied deep learning to classify histological tissue types and co-authored a large-scale survey on computational pathology analyzing over 800 papers. 	Summer 2020 Toronto, ON
	Research Intern N.I Institute for Health, National University of Singapore <ul style="list-style-type: none"> Wrote MATLAB functions to automate cleaning and quality checks for a 76TB neural dataset. 	Summer 2019 Kent Ridge, Singapore
	Software Developer Intern Epson <ul style="list-style-type: none"> Released the Android and Windows Software Development Kit (SDK) designed for augmented reality smart-glasses, following A/B and unit testing standards. 	May 2018 - May 2019 Markham, ON
AWARDS	Student Paper Award (Runner-up), ASA Statistics in Imaging DataFest Finalist, ENAR Student Travel Award, American Statistical Association PhD Fellowship Award, New York University Certificate for Artificial Intelligence Engineering, University of Toronto International Summer Research Award, University of Toronto Gold Award, The Duke of Edinburgh's Award	2024 2024 2024 2022 2020 2019 2015
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 "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 "DeepFDR: A Deep Learning-based False Discovery Rate Control Method for Neuroimaging Data" Oral, Joint Statistical Meetings, Portland, OR	August 2025 March 2025 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