

YOUNG-GEUN KIM

Younggeun.Kim@nyspi.columbia.edu

<https://kyg0910.github.io/>

<https://scholar.google.com/citations?user=HVqiptEAAAAJ>

RESEARCH INTERESTS

My research interests revolve around developing innovative data science tools and promoting their dissemination on biomedical data. Research topics include, but are not limited to:

- Deep generative models for multi-modal biomedical data (e.g., neuroimaging and multi-omics)
- Deep learning for identifying biomarkers associated with mental illness
- Reinforcement learning-based health care

PROFESSIONAL APPOINTMENTS

Assistant Professor - Hannan Endowed Visiting Scholar
Department of Statistics and Probability, Michigan State University

Aug. 2024 - Present

EDUCATION

Adjunct Associate Research Scientist
Department of Biostatistics, Columbia University
Mentor: Ying Liu, Ph.D.

Jul. 2021 - Aug. 2024

Postdoctoral Researcher
Department of Psychiatry, Columbia University
Mental Health Data Science, New York State Psychiatric Institute
Mentor: Ying Liu, Ph.D.

Jul. 2021 - Aug. 2024

Postdoctoral Researcher
Department of Statistics, Seoul National University
Mentor: Myunghee Cho Paik, Ph.D.

Mar. 2021 - Jun. 2021

Seoul National University
Ph.D. in Statistics
Advisor: Myunghee Cho Paik, Ph.D.

Mar. 2015 - Feb. 2021
Graduated with the Best Dissertation Award

Dissertation: Statistical distance of conditional distributions and its applications

Seoul National University
Triple Major
B.S. in Industrial Engineering
B.S. in Statistics
B.S. in Mathematical Sciences

Mar. 2010 - Feb. 2015
Graduated with Honors (Cum Laude)

HONORS & AWARDS

Career Development Award
Korean International Statistical Society

Dec. 2023

Outstanding Reviewer Award
Thirty-ninth International Conference on Machine Learning

Jul. 2022

Best Dissertation Award

Feb. 2021

College of Natural Sciences, Seoul National University

Seoul National University Innovation Program Scholarship

Mar. 2017 - Feb. 2018

Seoul National University

* Awarded to the Ph.D. student with the highest GPA in the department.

1st Prize, Student Paper Competition

June 2017

Korean Statistical Society

PUBLICATIONS & PREPRINTS

*: First author; ‡: Corresponding author

Journal

1. **Kim, Y.-G.***, Ravid, O.*, Zheng, X., Kim, Y., Neria, Y., Lee, S., He, X.‡, and Zhu, X.‡ (2024). Explaining deep learning-based representations of resting state functional connectivity data: focusing on interpreting nonlinear patterns in autism spectrum disorder. *Frontiers in Psychiatry, section Computational Psychiatry*. [\[Paper\]](#) [\[Code\]](#)
- **Top 86** Psychiatry and Mental Health journal (H-index: 114; upper 14.9%)
2. **Kim, Y.-G.***, Lee, K., and Paik, M.C.‡ (2022). Conditional Wasserstein generator. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. [\[Paper\]](#) [\[Code\]](#)
- **Top 1** Applied Mathematics journal (H-index: 397; upper 0.2%)
3. **Kim, Y.-G.***, Kwon, Y., and Paik, M.C.‡ (2019). Valid oversampling schemes to handle imbalance. *Pattern Recognition Letters*, 125 (1): 661-667. [\[Paper\]](#) [\[Code\]](#)
- **Top 13** AI journal (H-index: 170; upper 4.6%)

Peer-reviewed Conference

1. **Kim, Y.-G.***, Liu, Y.‡, and Wei, X. (2023). Covariate-informed representation learning to prevent posterior collapse of iVAE. *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics (AISTATS 2023)*. [\[Paper\]](#) [\[Code\]](#)
- **Top 6** AI conference (H5-index: 85)
2. Kim, M.*, **Kim, Y.-G.**, Kim, D., Kim, Y., and Paik, M.C.‡ (2021). Kernel-convoluted deep neural networks with data augmentation. *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI 2021)*. [\[Paper\]](#) [\[Code\]](#)
- **Top 4** AI conference (H5-index: 180)
3. **Kim, Y.-G.***, Kwon, Y., Chang, H., and Paik, M.C.‡ (2020). Lipschitz continuous autoencoders in application to anomaly detection. *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics (AISTATS 2020)*. [\[Paper\]](#) [\[Code\]](#)
- **Top 6** AI conference (H5-index: 85)

Patents

1. Paik, M.C.‡, **Kim, Y.-G.**, and Lee, K., Method and apparatus for conditional data generation using conditional Wasserstein generator. Republic of Korea Patent. [\[Info\]](#)
2. Paik, M.C.‡, **Kim, Y.-G.**, and Chang, H., Learning method and learning device for high-dimension unsupervised anomaly detection using kernalized Wasserstein autoencoder to lessen too many computations of Christophel function, and testing method and testing device using the same. Republic of Korea Patent. [\[Info\]](#)

Preprints

1. **Kim, Y.-G.***, Lee, K., Choi, Y., Won, J.-H., and Paik, M.C.[‡] (2023). Wasserstein geodesic generator for conditional distributions (under *Major Revision* at Journal of Machine Learning Research). [\[ArXiv\]](#)[\[Code\]](#)
2. Kim, S.*, **Kim, Y.-G.**, and Wang, Y.[‡] (2023). Temporal generative models for learning heterogeneous group dynamics of ecological momentary data (under *Revision Invited* at Biometrics). [\[BioRxiv\]](#)
3. **Kim, Y.-G.***, Brandt, L., Cheung, K., Nunes, E. V., Roll, J., Luo, S. X.[‡], and Liu, Y.[‡] (2024). Optimizing contingency management with reinforcement learning. [\[MedRxiv\]](#)[\[Code\]](#)

GRANTS & FUNDING

I participated the following projects as a **research scientist**.

- **A data science framework for empirically evaluating and deriving reproducible and transferrable RDoC constructs in youth (R01)** *Jul. 2021 - Present*
Funded by NIH/NIMH
- **Deep learning with incomplete and sequential data: Application to biomedical data** *Mar. 2020 - Jun. 2021*
Funded by National Research Foundation of Korea
- **Development of low-yield trackers via causal inference** *May 2019 - Nov. 2019*
Funded by SK Telecom
- **Statistical approaches to deep learning: New methods for convolutional neural networks in application to medical imaging data** *Mar. 2017 - Feb. 2020*
Funded by National Research Foundation of Korea
- **Deep Learning for the CT based Acute Cerebral Infarction Classification and Lesion Segmentation** *July 2016 - May 2019*
Collaborated with Seoul National University Bundang Hospital
Funded by National Research Foundation of Korea
- **New Robust Methods for Missing or Censored Covariates** *Mar. 2016 - Nov. 2016*
Funded by National Research Foundation of Korea

SELECTED TALKS

Invited Talks

- **Kim, Y.-G.**, Luo, S. X., Brandt, L., Cheung, K., Nunes, E. V., Roll, J., and Liu, Y. (2024). Optimizing contingency management interventions in substance use disorder treatment with reinforcement learning. *The Joint Statistical Meetings (JSM), Portland, OR.*
- **Kim, Y.-G.** and Liu, Y. (2024). Deep Identifiable Generative Models for Multi-Modal Data Analysis. *The 2024 International Chinese Statistical Association (ICSA) Applied Statistics Symposium, Nashville, TN.*
- **Kim, Y.-G.**, Liu, Y., and Wei, X. (2023). Covariate-informed representation learning to prevent posterior collapse of iVAE. *The Twenty Third International Conference on Artificial Intelligence and Statistics (AISTATS 2023), Palau de Congressos, Valencia, Spain.*[†]
- **Kim, Y.-G.**, Kwon, Y., Chang, H., and Paik, M.C. (2020). Lipschitz continuous autoencoders in application to anomaly detection. *The 23rd International Conference on Artificial Intelligence and Statistics (AISTATS 2020), Virtual conference due to COVID-19.*

- **Kim, Y.-G.**, Kwon, Y., Chang, H., and Paik, M.C. (2019). Lipschitz continuous autoencoders in application to anomaly detection. *IMS-China International Conference on Statistics and Probability, Dalian, China*.

Contributed Talks

- **Kim, Y.-G.** and Liu, Y. (2024). Explaining Nonlinear Patterns in Children’s Structural MRI with Multi-modal Identifiable VAE. *The ABCD Insights & Innovations Meeting, MD*.[†]
- **Kim, Y.-G.**, Liu, Y., Brandt, L., Cheung, K., Nunes, E. V., Roll, J., and Luo, S. X. (2023). Optimizing contingency management in substance use disorder treatment using off-policy policy evaluation. *Eastern North American Region (ENAR) 2023 Spring meeting*.

[†] indicates a poster presentation.

TEACHING EXPERIENCE

Statistical Practice and Research for Interdisciplinary Sciences (SPRIS) *Spring 2024*
Columbia University (role: **Guest Lecturer**)

- Graduate-level course on interdisciplinary research topics in Biostatistics.
- Gave the lecture “Recent Topics on Conditional Generative Models”

Statistics Lab *Fall 2015*

Seoul National University (role: **Instructor**)

- Freshman course to introduce R programming.
- Taught and oversaw progress for 17 students, including providing the whole 13 lectures, writing exam problems, and giving final grades.

Deep Learning: A Statistical Perspective *Fall 2021*

Seoul National University (role: **Guest Lecturer**)

- Graduate-level course on deep learning.
- Gave the lecture “Conditional Image Synthesis and Its Applications” in English.

MENTORSHIP EXPERIENCE

Co-mentoring Graduate Students at Columbia University

- Bin Yang, Ph.D. Candidate, Department of Biostatistics April. 2024 - Present
 - Conducted regular weekly meetings with Dr. Yuanjia Wang.
 - Provided mentorship on variational autoencoders for EEG data.
- Soohyun Kim, Ph.D., Department of Biostatistics Mar. 2022 - Present
 - Conducted regular weekly meetings with Dr. Yuanjia Wang.
 - Provided mentorship on the doctoral dissertation and the following paper:
*Kim, S., **Kim, Y.-G.**, and Wang, Y. (2023). Temporal generative models for learning heterogeneous group dynamics of ecological momentary data (under Revision Invited at Biometrics).*
- Zekai Jin, Master Student, Department of Biostatistics Dec. 2022 - Oct. 2023
 - Conducted regular bi-weekly meetings with Dr. Seonjoo Lee.
 - Provided mentorship on deep learning-based EEG denoising methods.

OTHER PROFESSIONAL ACTIVITIES

Conference Organizer

- Invited Session at JSM 2024 (role: **Organizer & Speaker**); Title: Reliable and Cost-effective Mental Health Care with Reinforcement Learning
- Invited Session at 2024 ICSCA (role: **Speaker**); Title: Recent Advances in Precision Medicine and Adaptive Experiments
- Invited Session at ENAR 2023 (role: **Chair**); Title: Advanced Methods for Analyzing Large-Scale Neuroimaging Data from Nationwide Consortia for Mental Health Research [\[Info\]](#)
- Oral Presentation Session at ICML 2022 (role: **Chair**); Title: Theory [\[Info\]](#)

Reviewer

- JAMA Psychiatry
- Expert Systems with Applications
- Pattern Recognition Letters
- International Journal of Computer Assisted Radiology and Surgery
- International Conference on Machine Learning
- International Conference on Artificial Intelligence and Statistics