

# YOUNG-GEUN KIM

Younggeun.Kim@nyspi.columbia.edu ◇ Personal Website ◇ Google Scholar

## BIOGRAPHICAL INFORMATION

---

Date of Birth: September 10, 1991

Citizenship: Republic of Korea

## TRAINING & EDUCATION

---

**Adjunct Associate Research Scientist**

*Jul. 2021 - Present*

Department of Biostatistics, Columbia University

**Research Scientist II**

*Jul. 2021 - Present*

Department of Psychiatry, Columbia University

Research Foundation for Mental Hygiene, New York State Psychiatric Institute

**Postdoctoral Researcher**

*Mar. 2021 - Jun. 2021*

Department of Statistics, Seoul National University

**Seoul National University**

*Mar. 2015 - Feb. 2021*

Ph.D. in Statistics

Graduated with the Best Dissertation Award

**Advisor:** Myunghee Cho Paik, Ph.D.

**Dissertation:** Statistical distance of conditional distributions and its applications

**Seoul National University**

*Mar. 2010 - Feb. 2015*

**Triple Major**

Graduated with Honors (Cum Laude)

B.S. in Industrial Engineering

B.S. in Statistics

B.S. in Mathematical Sciences

## HONORS & AWARDS

---

**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.

**Student Paper Competition 1st Prize**

*June 2017*

Korean Statistical Society

## PUBLICATIONS & PREPRINTS

---

### Journal

- **Kim, Y.-G.\***, Lee, K., and Paik, M.C.<sup>‡</sup> (2022). Conditional Wasserstein generator. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. [\[Paper\]](#) [\[Code\]](#)  
- IF: 23.6; H-index: 397 (**Top 1** Applied Mathematics journal; upper 0.2%)
- **Kim, Y.-G.\***, Kwon, Y., and Paik, M.C.<sup>‡</sup> (2019). Valid oversampling schemes to handle imbalance. *Pattern Recognition Letters*, 125 (1): 661-667. [\[Paper\]](#) [\[Code\]](#)  
- IF: 5.1; H-index: 170 (**Top 13** AI journal; upper 4.6%)

## Conference

- **Kim, Y.-G.\***, Liu, Y.<sup>‡</sup>, 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)*. [\[Paper\]](#) [\[Code\]](#)  
- H5-index: 91 (**Top 6** AI conference)
- Kim, M.\*, **Kim, Y.-G.**, Kim, D., Kim, Y., and Paik, M.C.<sup>‡</sup> (2021). Kernel-convoluted deep neural networks with data augmentation. *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI 2021)*. [\[Paper\]](#) [\[Code\]](#)  
- H5-index: 212 (**Top 4** AI conference)
- **Kim, Y.-G.\***, Kwon, Y., Chang, H., and Paik, M.C.<sup>‡</sup> (2020). Lipschitz continuous autoencoders in application to anomaly detection. *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics (AISTATS 2020)*. [\[Paper\]](#) [\[Code\]](#)  
- H5-index: 91 (**Top 6** AI conference)

## Patent

- Paik, M.C.<sup>‡</sup>, **Kim, Y.-G.**, and Lee, K., Method and apparatus for conditional data generation using conditional Wasserstein generator. Republic of Korea Patent. [\[Info\]](#)
- Paik, M.C.<sup>‡</sup>, **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. [\[Info\]](#)

## Preprint

- Kim, S.\*, **Kim, Y.-G.**, and Wang, Y.<sup>‡</sup> (2023). Temporal generative models for learning heterogeneous group dynamics of ecological momentary data (under *Revision Invited* at Biometrics). [\[BioRxiv\]](#)
- **Kim, Y.-G.\***, ..., and Paik, M.C.<sup>‡</sup> (2023). Wasserstein geodesic generator for conditional distributions (under review at Journal of Machine Learning Research). [\[ArXiv\]](#)[\[Code\]](#)
- **Kim, Y.-G.\***, Ravid, O.\*, ..., and Zhu, X.<sup>‡</sup> (2023). Explaining deep learning-based representations of resting state functional connectivity data: focusing on interpreting nonlinear patterns in autism spectrum disorder (under review at International Journal of Molecular Sciences). [\[BioRxiv\]](#) [\[Code\]](#) \*: Equal contribution of the first authors  
– Special Issue: Artificial Intelligence in Biomarker Discovery 2.0

## GRANT

---

### (Submitted) NIH K99/R00: Pathway to Independence Award (role: PI)

- Title: Development of reinforcement learning-based tools for evaluating contingency management intervention in substance use disorder treatments; Total Grant Amount: \$1,113,066.

## SELECTED INVITED TALKS

---

- **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.<sup>‡</sup>

- Liu, Y, **Kim, Y.-G.**, and Wei, X (2023). Covariate informed identifiable variational autoencoder to learn representations from brain imaging measures. *Eastern North American Region (ENAR), Nashville, TN*.
- Kim, M., **Kim, Y.-G.**, Kim, D., Kim, Y., and Paik, M.C. (2021). Kernel-convoluted deep neural networks with data augmentation. *The 35th AAAI Conference on Artificial Intelligence (AAAI-21), Virtual conference due to COVID-19*.
- **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*.

† indicates a poster presentation.

## OTHER PROFESSIONAL ACTIVITIES

---

### Conference Organizer

- (Under review for the Invited Session) Joint Statistical Meetings 2024 (role: **Organizer & speaker**); Title: Reliable and Cost-effective Mental Health Care with Reinforcement Learning
- Invited Session at Eastern North American Region 2023 (role: **Chair**); Title: Advanced Methods for Analyzing Large-Scale Neuroimaging Data from Nationwide Consortiums for Mental Health Research [\[Info\]](#)
- Oral Presentation Session at International Conference on Machine Learning 2022 (role: **Chair**); Title: Theory [\[Info\]](#)

### Reviewer

- International Conference on Machine Learning: H5-index 254 (**Top 3** AI conference)  
- Awarded **Top 10%** Reviewer [\[List\]](#)
- JAMA Psychiatry: IF 25.8; H-index 394 (**Top 1** in Psychiatry and Mental Health; upper 0.2%)
- Expert Systems with Applications: IF 8.5; H-index 249 (**Top 2** in AI; upper 0.7%)
- Pattern Recognition Letters: IF 5.1; H-index 170 (**Top 12** in AI; upper 4.6%)
- International Conference on Artificial Intelligence and Statistics: H5-index 91 (**Top 6** AI conference)
- International Journal of Computer Assisted Radiology and Surgery: IF: 3.0; H-index: 61 (**Top 77** in Radiology, Nuclear Medicine and Imaging; upper 23.7%)

## TEACHING EXPERIENCE

---

- Deep Learning: A Statistical Perspective (2018-2021) at Seoul National University
  - Graduate-level course on deep learning.
  - Gave the lecture “Conditional Image Synthesis and Its Applications” in English.
  - Gave lectures about deep learning programming languages and deep learning-based object detection algorithms in English.

- Statistics Lab. (2015) at Seoul National University
  - Freshman course to introduce R programming.
  - Gave whole lectures.