

## CURRICULUM VITAE

**Il-Youp Kwak****Work**

Department of Applied Statistics  
 College of Business & Economics  
 Chung-Ang University  
 84, Heukseok-ro, Dongjak-gu  
 Seoul, 06974, Republic of Korea

Phone: 82-10-5652-1534  
 Email: ikwak2@cau.ac.kr  
 Web: <https://ikwak2.github.io/>

**EDUCATION**

|      |  |
|------|--|
| 2014 | PhD, Statistics, University of Wisconsin-Madison (Advisor: Karl W. Broman) |
| 2008 | MS, Statistics, Korea University, Seoul, Korea (Advisor: Myung-Hoe Huh)    |
| 2006 | BS, Mathematics, Korea University, Seoul, Korea                            |

**PROFESSIONAL POSITIONS**

|           |  |
|-----------|--|
| 2023-     | Associate Professor, Department of Applied Statistics, College of Business & Economics, Chung-Ang University, Seoul, Korea   |
| 2019-2023 | Assistant Professor, Department of Applied Statistics, College of Business & Economics, Chung-Ang University, Seoul, Korea   |
| 2017-2019 | Staff Engineer, Samsung Research, Seoul R&D Campus, Seoul, Korea   |
| 2016-2017 | Research Associate, Lillehei Heart Institute, University of Minnesota, US (Supervisor: Daniel Garry)                         |
| 2014-2016 | Postdoctoral Associate, Division of Biostatistics School of Public Health, University of Minnesota, US (Supervisor: Wei Pan) |

**RESEARCH INTERESTS**

- Deep learning applications (focus on voice data and biological data)
- Biostatistics: statistical methodology and software development for bio-data analysis
- Real-world big data analysis: Develop new service that combines enterprise big data
- Development of Statistical Software tools

**DATA COMPETITIONS**

|        |  |
|--------|--|
| 2023.5 | [3 <sup>rd</sup> ] Our CAU_KU team (So-Yul Han, Taein Kang, Sunmook Choi, Jaejin Seo, Sanghyeok Chung, Sumi Lee, Seungsang Oh, Il-Youp Kwak) ranked 3rd place at the second Audio Deepfake Detection challenge 2023, Track 1.2 (fake audio detection track).   |
| 2022.9 | [3 <sup>rd</sup> ] Our Unlock_DNA team (Wuming Gong, Byeong-Chan Kim, Juhyun Lee, and Il-Youp Kwak) ranked 3rd place at the DREAM challenge "Predicting gene expression using millions of random promoter sequences."  |
| 2022.9 | [3 <sup>rd</sup> and 5 <sup>th</sup> ] Our CAU_UMN team (Jungguk Lee, Taein Kang, Narin Kim, Soyul Han, Hyejin Won, Wuming Gong, and Il-Youp Kwak) ranked 3rd place on Heart Murmur Detection, and 5th place on clinical outcome prediction from Phonocardiogram Recordings at the George B. Moody PhysioNet Challenges. |

- 2022.1 [3<sup>rd</sup>] Our CAU\_KU team (Il-Youp Kwak, Jonghoon Yang, Yerin Lee, Sunmook Choi and SeungSang Oh) ranked 3rd place at the ICASSP 2022 Grand Challenge on Audio Deepfake Detection Track 1 (Low quality fake audio detection).
- 2021.7 [6<sup>th</sup>] Our CAU team (Hyejin Won, Baekseung Kim, Il-Youp Kwak and Changwon Lim) ranked 6th place at the IEEE AASP challenge on Detection and Classification of Acoustic Scenes and Event (DCASE) Task 6 on Audio Captioning.
- 2020.7 [7<sup>th</sup> and 9<sup>th</sup>] Our CAU-ET team (Yerin Lee, Soyoung Lim and Il-Youp Kwak) ranked 7th and 9th place at the IEEE AASP challenge on Detection and Classification of Acoustic Scenes and Event (DCASE) Task 1 subtask B and A
- 2020.3 [Winner of a Biomedical Data Competition] Our team (Il-Youp Kwak and Wuming Gong) won both Subchallenges 2 and 3 of the Allen Institute Cell Lineage Reconstruction DREAM Challenge

## COURSES TAUGHT

|                            |  |
|----------------------------|--|
| Deep Learning for AI       | 2021 Spring, 2022 Spring, 2023 Spring, 2024 Spring |
| Machine Learning for AI    | 2021 Fall, 2022 Fall, 2023 Fall                    |
| Deep Learning with Python  | 2021 Spring, 2022 Spring, 2023 Spring, 2024 Spring |
| Mathematical Statistics II | 2020 Fall  |
| Mathematical Statistics I  | 2020 Spring  |
| Biostatistics              | 2019 Fall  |

## PUBLICATIONS

### Working Papers

- Han S, Seo J, Choi S, Kang T, Chung S, Lee S, Park S, Oh S, **Kwak IY\***, (2024+) Bell-shaped Frequency Feature Masking Augmentation for Voice Spoofing Detection, Submitted.
- Kang T, Han S, Seo J, Choi S, Lee S, Park S, Chung S, Oh S, **Kwak IY\***, (2024+) Experimental Study: Enhancing Spoofing Detection with Fine-tuned Large Voice Models, Submitted.
- Kim H, Kim B, Lee S, Kang M, Nam H, Park S, **Kwak IY\***, Lee J (2024+) RAPID: Robust multi-patch masker using channel-wise Pooled variance with two-stage patch Detection, *Submitted*.
- Kim H, Kim B, Kang M, Nam H, Park S, **Kwak IY\***, Lee J (2024+) OD-SHIELD: Convolutional Autoencoder-based Defense against Adversarial Patch Attacks in Object Detection, *In preparation*.

### International

- 2024 Choi S, Chung S, Lee S, Han S, Kang T, Seo J, **Kwak IY**, Oh S (2024) TB-ResNet: Bridging the Gap from TDNN to ResNet in Automatic Speaker Verification with Temporal-Bottleneck Enhancement, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2024)*, COEX, Seoul, Korea.
- Kwak IY\***, Kim KS, Min HJ (2024) Association of gustatory dysfunction and Alzheimer's disease: a systematic review and meta-analysis, *Rhinology*, 62(2), 130-142.
- Kwak IY\***, Kim B, Lee J, Kang T, Garry JD, Zhang J, Gong W (2024) Proformer: a hybrid macaron transformer model predicts expression values from promoter sequences, *BMC Bioinformatics*, 25, 81.
- Rafi AM, Nogina D, Penzar D, Lee D, Lee D, Kim N, Kim S, Kim D, Shin Y, **Kwak IY**, Meshcheryakov G, Lando A, Zinkevich A, Kim B, Lee J, Kang T, Vaishnav ED, Yadollahpour P, Kim S, Albrecht J, Regev A, Gong W, Kulakovskiy I, Mayer P, (2024) Evaluation and optimization of sequence-based gene regulatory deep learning models, *Nature Biotechnology*, Accepted.
- Han S, Kang T, Lee J, Kim N, Won H, Kim YH, Gong W, **Kwak IY\***, (2024) A Deep Neural Network Approach to Heart Murmur Detection Using Spectrogram and Peak Interval Features, *Engineering Applications of Artificial Intelligence*, Accepted.

- 2023 Lee Y, Kim N, Jeong J, **Kwak IY\*** (2023) Experimental Case Study of Self-supervised Learning for Voice Spoofing Detection, *IEEE Access*, 11, 24216-24226
- Park YB, Kim H, Lee HJ, Baek SH, **Kwak IY**, Kim SH (2023) The Clinical Application of Machine Learning Models for Risk Analysis of Ramp Lesions in Anterior Cruciate Ligament Injuries, *The American Journal of Sports Medicine*, 51(1), 107-118
- Kwak IY\***, Kim KS, Min HJ (2023) Gustatory dysfunction is related to Parkinson's disease: A systematic review and meta-analysis, *International Forum of Allergy & Rhinology*, 13(10), 1949-1957.
- Kwak IY\***, Kwag S, Lee J, Jeon Y, Hwang J, Choi HJ, Yang JH, Han SY, Huh JH, Lee CH, Yoon JW (2023) Voice Spoofing Detection through Residual Network, Max Feature Map, and Depthwise Separable Convolution, *IEEE Access*, 11, 49140-49152
- Namgung M, Ahn C, Park Y, **Kwak IY**, J Lee, M Won (2023) Mortality of adult patients due to sepsis and septic shock in South Korea: a systematic review and meta-analysis, *Clinical and Experimental Emergency Medicine*, 10(2), 157-171.
- Won H, Kim B, **Kwak IY\***, Lim C (2023) Using various pre-trained models for audio feature extraction in automated audio captioning, *Expert Systems with Applications*, 231(30), 120664.
- Han SY, Kang T, Choi S, Seo J, Chung S, Lee S, Oh S, **Kwak IY\*** (2023) CAU KU deep fake detection system for ADD 2023 challenge, *The 32nd International Joint Conference on Artificial Intelligence (IJCAI 2023, Workshop on Deepfake Audio Detection and Analysis)*, Macao SAR, China.
- Han SY, Jeon W, Gong W, **Kwak IY\*** (2023) MCHear: Multi-Channel based Heart Signal Processing Scheme for Heart Noise Detection using Deep Learning, *Biology*, 12(10), 1291.
- 2022 Gong W, Kim H, Garry D, **Kwak IY\*** (2022) Single cell lineage reconstruction using distance-based algorithms and the R package, DCLEAR, *BMC Bioinformatics*, 23(1), 1-14
- Yang J, Choi S, Lee Y, Oh S, **Kwak IY\*** (2022) Light-Weight Frequency Information Aware Neural Network Architecture for Voice Spoofing Detection, *26th International Conference on Pattern Recognition (ICPR 2022)*, Montréal Québec, Canada.
- Choi S, **Kwak IY**, Oh S (2022) Overlapped Frequency-Distributed Network: Frequency-Aware Voice Spoofing Countermeasure, *Interspeech 2022*, Incheon, Korea.
- Kwak IY\***, Choi S, Yang J, Lee Y, Oh S (2022) Low-quality Fake Audio Detection through Frequency Feature Masking, *30th ACM International Conference on Multimedia (ACM MM 2022, Workshop on Deepfake Detection for Audio Multimedia)*, Lisbon, Portugal.
- Lee J, Kang T, Kim N, Han S, Won H, Gong W, **Kwak IY\*** (2022) Deep Learning Based Heart Murmur Detection using Frequency-time Domain Features of Heartbeat Sounds, *The 49th computing in cardiology conference (CinC 2022)*, Tampere, Finland.
- Yoshihara T, Miller ND, Rabanal FA, Myles H, **Kwak IY**, Karl WB, Sadkhin B, Dilkes BP, Hudson ME, Spalding EP (2022) Leveraging orthology within maize and Arabidopsis QTL to identify genes affecting natural variation in gravitropism, *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 119 (40) e2212199119.
- Choi H, Jung C, Kang T, Kim H, **Kwak IY\*** (2022) Explainable Time-series Prediction using a Residual Network and Gradient-based Methods, *IEEE ACCESS*, 10, 108469 - 108482.
- Kwak IY\***, Kim KS, Min HJ (2022) Differential effect of meteorological factors and particulate matter with  $\leq 10\text{-}\mu\text{m}$  diameter on epistaxis in younger and older children, *Scientific Reports*, 12, 21029.
- 2021 **Kwak IY\***, Kwag S, Lee J, Huh J, Jeon Y, Hwang J, Yoon J, Lee C. (2021) RESMAX: Detecting Voice Spoofing Attacks with Residual Network and Max Filter Map, *25th International Conference on Pattern Recognition (ICPR2020)*, Milan, Italy
- Lee Y, Lim S, **Kwak IY\*** (2021) CNN-Based Acoustic Scene Classification System, *Electronics*, 10(4), 371
- Kim K, **Kwak IY**, Min H (2021) Particulate Matter 10 (PM10) Is Associated with Epistaxis in Children and Adults, *International Journal of Environmental Research and Public Health*, 18(9), 4809,
- Gong W, Granados A, Hu J, Jones M, Raz O, Martinez IS, Zhang H, Chow KK, **Kwak IY**, ... , Meyer P. (2021) Benchmarked approaches for cell lineage reconstructions of in vitro dividing cells and in silico models of *Caenorhabditis elegans* and *Mus musculus* developmental trees, *Cell Systems*, 12, 1-17.

- Park J, **Kwak IY**, Lim C (2021) A Deep Learning Model with Self-Supervised Learning and Attention Mechanism for COVID-19 Diagnosis Using Chest X-ray Images, *Electronics*, 10(16), 1996.
- Won H, Kim BS, **Kwak IY** and Lim C (2021) Transformer Followed by Transfer Learning for Audio Captioning, *The 6th Workshop on Detection and Classification of Acoustic Scenes and Events (DCASE2021)*, Online.
- 2020 Ahmed E, **Kwak IY**, Huh J, Kim I, Oh T, Kim H (2020) Void: A fast and light voice liveness detection system, *USENIX Security Symposium*, Boston, USA.
- Yoon HY, **Kwak IY\*** (2020) Association Modeling on Keyword and Abstract Data in Korean Port Research, *Journal of Korea Trade*, 25 (5), 71-86
- 2019 **Kwak IY\***, Huh J, Kim I, Han S, Yoon J. (2019) Voice Presentation Attack Detection through Text-Converted Voice Command Analysis, *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, Glasgow, UK.
- Magli A, Baik J, Mills LJ, **Kwak IY**, Dillon BS, Gonzalez RM, Stafford DA, Swanson SA, Stewart R, Thomson JA, Garry DJ, Dynlacht BD, Perlingeiro R (2019) Time-dependent Pax3-mediated chromatin remodeling and cooperation with Six4 and Tead2 specify the skeletal myogenic lineage in developing mesoderm, *PLoS biology*, 17 (2), e3000153
- Magli A, Baik J, Pota P, Cordero CO, **Kwak IY**, Garry DJ, Love PE, Dynlacht BD, Perlingeiro R (2019) Pax3 cooperates with Ldb1 to direct local chromosome architecture during myogenic lineage specification, *Nature Communications*, 10, 2316
- 2018 Gong W, **Kwak IY\***, Pota P, Koyano-Nakagawa N, Garry DJ. (2018) DrImpute: imputing dropout events in single cell RNA sequencing data, *BMC Bioinformatics*, 19 (1), 220
- Gong W, **Kwak IY**, Koyano-Nakagawa N, Pan W, Garry DJ. (2018) TCM visualizes trajectories and cell populations from single cell data, *Nature Communications*, 9 (1), 2749
- 2017 **Kwak IY\***, Pan W. (2017) Gene- and pathway-based association tests for multiple traits with GWAS summary statistics. *Bioinformatics*, 33 (1), 64-71.
- Chen Z, Zhu W, Bender I, Gong W, **Kwak IY**, Yellamilli A, Hodges TJ, Nemoto N, Zhang J, Garry DJ, Berlo JH (2017) Pathologic Stimulus Determines Lineage Commitment of Cardiac C-kit+ Cells. *Circulation*, circulationaha. 117.030137
- 2016 Wei P, Cao Y, Zhang Y, **Kwak IY**, Xu Z, Crosby J, Boerwinkle E, Pan W. (2016) On Robust Association Testing for Quantitative Traits and Rare Variants. *G3*, 116.035485.
- Kwak IY\***, Pan W. (2016) Adaptive Gene- and Pathway-Trait Association Testing with GWAS Summary Statistics. *Bioinformatics*, 32 (8), 1178-1184.
- 2015 Pan W, **Kwak IY**, Wei P. (2015) A Powerful Pathway-Based Adaptive Test for Genetic Association with Common or Rare Variants. *The American Journal of Human Genetics*, 97:86-98
- Kwak IY\***, Moore CR, Spalding EP, Broman KW. (2015) Mapping quantitative trait loci underlying function-valued traits using functional principal component analysis and multi-trait mapping. *G3*, 6 (1), 79-86.
- 2014 **Kwak IY\***, Moore CR, Spalding EP and Broman KW. (2014) A simple regression-based method to map quantitative trait loci underlying function-valued phenotypes. *Genetics*, 197(4):1409-1416
- 2013 Moore CR, Johnson LS, **Kwak IY**, Livny M, Broman KW, Spalding EP (2013) High-throughput computer vision introduces the time axis to a quantitative trait map of a plant growth response. *Genetics* 195:1077–1086
- 2010 Billings T, Sargent EE, Szatkiewicz JP, Leahy N, **Kwak, IY**, Bektassova N, Walker M, Hassold T, Graber JH, Broman KW, Petkov PM (2010) Patterns of recombination activity on mouse chromosome 11 revealed by high resolution mapping. *PLoS One* 5:e15340

- 2024 Seo J, Kang T, **Kwak IY\*** (2024) Pre-trained models and ensemble technique for speech emotion recognition, *Journal of the Korean Data and Information Science Society*, Accepted
- 2023 Park K, **Kwak IY\*** (2023) Comparative study of data augmentation methods for fake audio detection, *The Korean Journal of Applied Statistics*, 36(2), 101-114  
Kang T, **Kwak IY\*** (2023) A two-stage training approach for voice spoofing detection, *Journal of the Korean Data and Information Science Society*, 34, 203-214
- 2021 Yang JH, **Kwak IY\*** (2021) Research trends in statistics for domestic and international journal using paper abstract data, *The Korean Journal of Applied Statistics*, 34(2), 267—278  
Choi HJ, **Kwak IY\*** (2021) Data augmentation in voice spoofing problem, *The Korean Journal of Applied Statistics*, 34(3), 435—446  
Yoon HY, **Kwak IY\*** (2021) Analysis of Shipping and Logistics News Articles using Topic Modeling, *Korea Trade Review*, 46(4), 61—76  
Lim SY, **Kwak IY\*** (2021) Light weight architecture for acoustic scene classification, *The Korean Journal of Applied Statistics*, 34(6), 979—993
- 2020 Yoon HY, **Kwak IY\*** (2020) Trends in FTA Research of Domestic and International Journal using Paper Abstract Data, *Korea Trade Review*, 45(5), 37—53  
Yoon HY, **Kwak IY\*** (2020) The Association Modeling on Keywords and Documents of Korea International Trade Research using Paper Abstract data, *Korea International Commerce Review*, 35(2), 45—64
- 2019 Yoon HY, **Kwak IY\*** (2019) The Association Modeling on Keywords and Documents of Logistics Research using Paper Abstract data, *Korea International Commerce Review*, 34(3), 147—166
- 2008 **Kwak IY\***, Huh MH (2008) Andrews' Plot for Extended Uses, *The Korean Communications in Statistics*, 15: 87–94  
**Kwak IY\***, Huh MH (2008) Parallel Coordinate Plots of Mixed-Type Data, *The Korean Communications in Statistics*, 15: 587–595  
Park MR, **Kwak IY**, Huh MH (2008) Applications of Parallel Coordinate Plots for Visualizing Gene Expression Data, *The Korean Journal of Applied Statistics*, 21(6): 911–921

## SOFTWARE DEVELOPMENT

- Python/PrixFixe A Python module prepared for Random Promoter Dream Challenge 2022. Sequence-to-expression models trained on expression measurements of random promoter sequences. The models would receive regulatory DNA sequence as input and use it to predict the corresponding gene expression value. The top-performing solutions in the challenge exceeded performance of all previous state-of-the-art models for similar data.
- R/DCLEAR An R package for R package for Distance based Cell LineAge Reconstruction (DCLEAR). These codes are created during the participation of Cell Lineage Reconstruction DREAM challenge. Our team achieved the best performance on subchallenges 2 and 3 of the competition.
- R/Drimpute An R package for imputing dropout events. Many statistical methods in cell type identification, visualization and lineage reconstruction do not account for dropout events ('PCAreduce', 'SC3', 'PCA', 't-SNE', 'Monocle', 'TSCAN', etc). 'Drimpute' can improve the performance of such software by imputing dropout events.
- R/aSPU An R package for the (adaptive) Sum of Powered Score ('SPU' and 'aSPU') tests, inverse variance weighted Sum of Powered score ('SPUw' and 'aSPUw') tests and some gene-based and pathway based association tests (aSPUs, aSPUsPath, MTaSPUsSet, MTaSPUsSetPath, etc). (<https://cran.r-project.org/web/packages/aSPU/index.html>).

|          |  |
|----------|--|
| R/funqtl | An R package with add-on functions for the R/qlt package to deal with QTL mapping for function-valued traits. ( <a href="https://github.com/ikwak2/funqtl">github.com/ikwak2/funqtl</a> ). |
| R/xoi    | Tools for analyzing crossover interference. ( <a href="https://github.com/kbroman/xoi">github.com/kbroman/xoi</a> ). Assisted Karl Broman to develop the package using C language.         |
| R/APCP   | An R package that plots Andrews type parallel coordinate plot. ( <a href="https://github.com/ikwak2/APCP">github.com/ikwak2/APCP</a> ).  |

## PRESENTATIONS

- 2022 Low-quality Fake Audio Detection through Frequency Feature Masking, 30th ACM International Conference on Multimedia (ACM MM 2022, Workshop on Deepfake Detection for Audio Multimedia), Lisbon, Portugal
- 2021 Detecting voice spoofing attacks with residual network and max filter map with Grad-CAM activation, 4th International Conference on Econometrics and Statistics (EcoSta 2021), HKUST, Hong Kong
- 2021 ResMax: Detecting Voice Spoofing Attacks with Residual Network and Max Filter Map, 10th World Congress in Probability and Statistics, Seoul, Korea
- 2020 ResMax: Detecting Voice Spoofing Attacks with Residual Network and Max Filter Map, *Spring Conference, The Korean Statistical Society*, Cheonan, Korea
- 2019 A survey on voice spoofing detection methods, *Autumn Conference, The Korean Statistical Society*, Seoul, Korea
- 2019 Voice Presentation Attack Detection through Text-Converted Voice Command Analysis, *Spring Conference, The Korean Statistical Society*, Kangwon, Korea
- 2019 Voice Presentation Attack Detection through Text-Converted Voice Command Analysis, *2019 CHI Conference on Human Factors in Computing Systems*, Glasgow, UK
- 2017 Gene- and pathway-based association tests for multiple traits with GWAS summary statistics, *2017 ENAR spring meeting*, Washington, DC
- 2016 Adaptive Gene- and Pathway-Trait Association Testing with GWAS Summary Statistics, *2016 ENAR spring meeting*, Austin, TX
- 2015 Regression-Based Methods to Map Quantitative Trait Loci Underlying Function-Valued Phenotypes, *2015 ENAR spring meeting*, Miami, FL
- 2015 Adaptive Gene- and Pathway-Trait Association Testing with GWAS Summary Statistics, *2015 3rd Annual BICB Symposium*, Minneapolis, MN
- 2013 QTL analysis with Function-valued Trait Data, *2013 ENAR Spring meeting*, Orlando, FL
- 2008 Parallel Coordinate Plots of Mixed-type Data, *2008 Korea-Japan Student Conference*, Seoul, Korea
- 2007 Andrews' Plot for Extended Uses, *Autumn Conference, The Korean Statistical Society*, Seoul, Korea

## POSTERS

- 2021 RESMAX: Detecting Voice Spoofing Attacks with Residual Network and Max Filter Map, 25th International Conference on Pattern Recognition (ICPR2020), Milan, Italy

- 2021 DCLEAR: Reconstructing Single Cell Lineage Trees from CRSIPR recorders by Distance-based Methods, Keystone symposia on Molecular and Cellular Biology, Single Cell Biology (EK26), Virtual conference.
- , 2016 Gene-based association tests for multiple traits with GWAS summary statistics, TC ASA Fall Research Meeting, Mounds View, MN
- 2015 Adaptive Gene- and Pathway-Trait Association Testing with GWAS Summary Statistics, TC ASA Fall Research Meeting, Rochester, MN
- 2013 QTL analysis with Function-valued Trait Data, Complex Trait Community annual meeting, Madison, WI

#### **JOURNAL MANUSCRIPT REVIEWS**

Molecular Ecology Resources, Scientific Reports, Biometrics, IEEE/ACM Transactions on Computational Biology and Bioinformatics, Bioinformatics, Journal of Korean Statistical Society, Communications for Statistical Applications and Methods, iScience, IEEE Transactions on Mobile Computing, International Journal of Environmental Research and Public Health, Healthcare, Frontiers in Genetics, Journal of king saud university - computer and information sciences, Journal of Digital Imaging, Frontiers in Plant Science