



# Seongyong Park

POSTDOCTORAL RESEARCH FELLOW · CANCER DATA SCIENCE LAB (CDSL)

Building 41, NIH, Bethesda, MD, 20894, USA

(+1) 640-204-6763 | seongyong.park@nih.gov | https://psychemistz.github.io

*“Things which we see are not by themselves what we see... It remains completely unknown to us what the objects may be by themselves and apart from the receptivity of our senses. We know nothing but our manner of perceiving them.” Immanuel Kant*

## Summary

I am a postdoctoral research fellow at the Cancer Data Science Laboratory (CDSL), NCI/CCR, NIH. I develop computational methods to infer cell-to-cell communication signals from gene expression data. My research spans a fourteen-year arc from experimental microfluidic systems for synthetic bacterial signaling (2011–2012) to large-scale computational inference of cytokine and secreted-protein activity from bulk, single-cell, and spatial transcriptomics.

Current contributions include IC2Bert for immune-checkpoint-blockade response prediction (*Scientific Reports*, 2025), a co-authored analysis of AI in precision oncology (*Nature Cancer*, 2025), authorship of **secactpy** (Python, Jiang Lab) and **spatial-gpu** (GPU-accelerated spatial activity inference), main contributions to **ridger**, and contributions to the Jiang Lab activity-inference R stack (**secact**, **spacet**). Further work at NCI/CDSL is ongoing.

My methodological interests include prior-informed machine learning for clinical biomarker discovery (disease gene prioritization, ensemble subspace learning), sparse representation learning for biomolecular classification, overlap-based effect-size measures (GSSMD, GMDM) for data and model quality evaluation, and self-supervised learning for biosignal reproducibility (SERS, phonocardiograms). I have authored or co-authored peer-reviewed work across *Nature Cancer*, *Scientific Reports*, *Analyst*, *Biomaterials*, *Lab on a Chip*, *Cancers*, and *Digital Signal Processing*, along with three issued patents.

## Work Experience

### Cancer Data Science Laboratory (CDSL), NCI/CCR, NIH

Bethesda, MD, USA

POSTDOCTORAL RESEARCH FELLOW

(Jan. 2023 – present)

- Develop computational methods to infer cell-to-cell communication and cytokine/secreted-protein activity from bulk, single-cell, and spatial transcriptomics.
- Author of **secactpy** (Python, Jiang Lab) and **spatial-gpu** (GPU-accelerated spatial kernels) for secreted-protein and cytokine-activity inference; main contributor to **ridger** (R ridge-regression activity inference; initiated by B. Ru); contributor to the Jiang Lab R packages **secact** and **spacet**.
- Develop prior-informed and self-supervised models for clinical biomarker discovery, including **IC2Bert** for immune-checkpoint-blockade response prediction (*Scientific Reports* 2025).
- Co-authored *Nature Cancer* 2025 review on AI contributions to precision oncology.
- Additional resources and manuscripts ongoing.

### BAPOR Lab, AMC.

Seoul, S.Korea

POSTDOCTORAL RESEARCHER

(Aug. 2022 - Dec. 2022)

- research, design and implement algorithms in deep learning for development of biosignal based diagnostic & prognostic models.
- apply theoretical expertise and innovation to generate patents and publications in top-ranked journals.
- using wide range of tools to acquire information and interpret data, writing up reports and presenting finding.

### Synergistic Bioinformatics Lab, KAIST.

Daejeon, S.Korea

PH.D. RESEARCH SCHOLAR

(Aug. 2015 - Aug. 2022)

- research, design and implement algorithms in graph theory, machine learning and deep learning for development of gene expression based diagnostic & prognostic models and drug candidate prediction models.
- apply theoretical expertise and innovation to generate patents and publications in top-ranked journals.
- using wide range of tools to acquire information and interpret data, writing up reports and presenting finding.

### Synergistic Bioinformatics Lab, KAIST.

Daejeon, S.Korea

RESEARCHER

Aug. 2014 - Aug, 2015

- research, design and implement algorithms in graph theory, machine learning and deep learning for development of gene expression based diagnostic & prognostic models and drug candidate prediction models.
- apply theoretical expertise and innovation to generate patents and publications in top-ranked journals.
- using wide range of tools to acquire information and interpret data, writing up reports and presenting finding.

### Biobrain Inc.

Daejeon, S.Korea

RESEARCHER

Aug. 2012 - Aug, 2014

- research, design and implement biosignal processing algorithms for EEG applications.
- using wide range of tools to acquire information and interpret data, writing up reports and presenting finding.

- Evaluation of research project associated with renewable energy.
- writing research trend reports.
- attending staff meetings and general administration.

## Skills

<b>Research</b>	Cell-to-cell communication, cytokine / secreted-protein activity inference, single-cell & spatial transcriptomics, biological-prior-informed ML, sparse representation learning, overlap statistics, biosignal processing.
<b>Programming</b>	Python, R, MATLAB, LaTeX, Shell/SLURM; PyTorch, TensorFlow, Keras, scikit-learn, Bioconductor, Seurat, Scanpy, CuPy/JAX for GPU
<b>Open-source software</b>	Author: <b>secactpy</b> (Python, Jiang Lab), <b>spatial-gpu</b> (GPU spatial kernels). Main contributor: <b>ridger</b> (R; initiated by B. Ru). Contributor: <b>secact</b> , <b>spacet</b> (Jiang Lab R stack).
<b>Infrastructure</b>	HPC (SLURM on NIH Biowulf), GPU/CPU cluster management, containerization (Singularity/Apptainer), reproducible pipelines
<b>Languages</b>	Korean (native), English (professional).

## Publications

### Peer-Reviewed Journal Articles

1. Jaeman Shin, **Seongyong Park**, Kyuchool Shin, Wooyeong Seo, Hyunseok Kim, Doo-Kwang Kim, Bongjin Moon, Sang Gu Cha, et al. "Temporal convolutional neural network-based feature extraction and asynchronous channel information fusion method for heart abnormality detection in phonocardiograms." *Computer Methods and Programs in Biomedicine* 269, 108871 (2025). [code]
2. **Seongyong Park**, Sunho Kim, Peng Jiang. "IC2Bert: masked gene expression pretraining and supervised fine tuning for robust immune checkpoint blockade (ICB) response prediction." *Scientific Reports* 15, 28044 (2025). [code]
3. Tian-Gen Chang, **Seongyong Park**, Alejandro A. Schäffer, Peng Jiang, Eytan Ruppín. "Hallmarks of artificial intelligence contributions to precision oncology." *Nature Cancer* 6(3), 417–431 (2025).
4. Sang-Wook Lee, **Seongyong Park**, Ji-Yeon Kim, Bongjin Moon, Dongwoo Lee, Jaewon Jang, Wooyeong Seo, Hyunseok Kim, Sung-Hoon Kim, et al. "Impact of preanesthetic blood pressure deviations on 30-day postoperative mortality in non-cardiac surgery patients." *Journal of Korean Medical Science* 39(35) (2024).
5. **Seongyong Park**, Abdul Wahab, Muhammad Usman, Imran Naseem, Shujaat Khan. "Artificial intelligence in bioimaging and signal processing." *Frontiers in Physiology* 14, 1267632 (2023).
6. **Seongyong Park**, Mohammad Sohail Ibrahim, Abdul Wahab, Shujaat Khan. "GMDM: A generalized multi-dimensional distribution overlap metric for data and model quality evaluation." *Digital Signal Processing* 134, 103930 (2023).
7. Ehtisham Fazal, Mohammad Sohail Ibrahim, **Seongyong Park**, Imran Naseem, Abdul Wahab. "Anticancer peptides classification using kernel sparse representation classifier." *IEEE Access* 11, 17626–17637 (2023).
8. **Seongyong Park**, Abdul Wahab, Minseok Kim, Shujaat Khan. "Self-supervised learning for inter-laboratory variation minimization in surface-enhanced Raman scattering spectroscopy." *Analyst* 148(7), 1473–1482 (2023).
9. **Seongyong Park**, Gwansu Yi. "Development of gene expression-based random forest model for predicting neoadjuvant chemotherapy response in triple-negative breast cancer." *Cancers* 14(4), 881 (2022).
10. Muhammad Usman, Shujaat Khan, **Seongyong Park**, Abdul Wahab. "AFP-SRC: identification of antifreeze proteins using sparse representation classifier." *Neural Computing and Applications* 34(3), 2275–2285 (2022).
11. **Seongyong Park**, Jaeseok Lee, Shujaat Khan, Abdul Wahab, Minseok Kim. "Machine learning-based heavy metal ion detection using surface-enhanced Raman spectroscopy." *Sensors* 22(2), 596 (2022).
12. **Seongyong Park**, Jaeseok Lee, Shujaat Khan, Abdul Wahab, Minseok Kim. "SERSNet: Surface-enhanced Raman spectroscopy based biomolecule detection using deep neural network." *Biosensors* 11(12), 490 (2021).
13. Muhammad Usman, Shujaat Khan, **Seongyong Park**, Jeong-A Lee. "AoP-LSE: antioxidant proteins classification using deep latent space encoding of sequence features." *Current Issues in Molecular Biology* 43(3), 1489–1501 (2021).
14. Yoon Hyeok Lee, Hojae Choi, **Seongyong Park**, Boah Lee, Gwansu Yi. "Drug repositioning for enzyme modulator based on human metabolite-likeness." *BMC Bioinformatics* 18(Suppl 7), 226 (2017).
15. **Seongyong Park**, Xiaoqiang Hong, Woon Sun Choi, Taesung Kim. "Microfabricated ratchet structure integrated concentrator arrays for synthetic bacterial cell-to-cell communication assays." *Lab on a Chip* 12(20), 3914–3922 (2012).
16. Vinuselvi Parisutham, **Seongyong Park**, Minseok Kim, Jung Min Park, Taesung Kim, Sung Kuk Lee. "Microfluidic technologies for synthetic biology." *International Journal of Molecular Sciences* 12(6), 3576–3593 (2011).
17. Woon Sun Choi, Dogyeong Ha, **Seongyong Park**, Taesung Kim. "Synthetic multicellular cell-to-cell communication in inkjet printed bacterial cell systems." *Biomaterials* 32(10), 2500–2507 (2011).
18. **Seongyong Park**, Dasol Kim, Robert J. Mitchell, Taesung Kim. "A microfluidic concentrator array for quantitative predation assays of predatory microbes." *Lab on a Chip* 11(17), 2916–2923 (2011).

## Conference Proceedings

1. **Seongyong Park**, Shujaat Khan, Muhammad Moinuddin, Ubaid M. Al-Saggaf. "GSSMD: A new standardized effect size measure to improve robustness and interpretability in biological applications." *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 1096–1099 (2020).

## Preprints

1. **Seongyong Park**, Shujaat Khan, Abdul Wahab. "E3-targetPred: Prediction of E3-target proteins using deep latent space encoding." *arXiv:2007.12073* (2020).
2. **Seongyong Park**, Shujaat Khan. "GSSMD: New metric for robust and interpretable assay quality assessment and hit selection." *arXiv:2001.06384* (2020).

## Patents

1. **Seongyong Park**, et al. "Method, computer program, and system for analyzing heart valve abnormalities based on heart murmur." *KR Patent App. 10-2023-0014972*, registered 30 December 2025.
2. **Seongyong Park**, et al. "Pulmonary function test apparatus and method using lung sound." *KR Patent App. 10-2023-0022834*, registered 21 February 2023.
3. **Seongyong Park**, et al. "Method for measuring preload of patients undergoing general anesthesia surgery based on acoustic variability index, and electronic device for executing thereof." *KR Patent App. 10-2023-0021404*, registered 17 February 2023.
4. Bumki Min, Gwansu Yi, **Seongyong Park**. "System and method for disease prediction based on group marker consisting of genes having similar function." *KR Patent 10-2236194*, issued 30 March 2021.
5. Taesung Kim, **Seongyong Park**. "Microfluidic concentrator for communication assays of microbes." *KR Patent 10-1330473*, issued 4 February 2013.
6. Taesung Kim, **Seongyong Park**. "Microfluidic concentrator array for observing predation behavior of microbes." *KR Patent 10-1238556*, issued 22 February 2013.

## Software

Author:

1. **secactpy** — Python package for secreted-protein activity inference from bulk and single-cell gene expression (hosted under the Jiang Lab data2intelligence organization). [github.com/data2intelligence/secactpy](https://github.com/data2intelligence/secactpy)
2. **spatial-gpu** — GPU-accelerated spatial kernels for neighborhood activity inference on spatial transcriptomics (Python). [github.com/psychemistz/spatial-gpu](https://github.com/psychemistz/spatial-gpu)

Main contributor (project initiated by B. Ru):

1. **ridger** — R package for ridge-regression-based activity inference. [github.com/beibeiru/ridger](https://github.com/beibeiru/ridger)

Contributor (Jiang Lab, NCI):

1. **secact** — R package for secreted-protein activity inference. [github.com/data2intelligence/secact](https://github.com/data2intelligence/secact)
2. **spacet** — Spatial Cellular Estimator for Tumors (R). [github.com/data2intelligence/spacet](https://github.com/data2intelligence/spacet)

## Education

---

### KAIST(Korea Advanced Institute of Science and Technology)

PH.D. IN BIO & BRAIN ENGINEERING

- Got the KAIST student scholarship which is given to promising students.

*Daejeon, S.Korea*

*Sep. 2015 - Aug. 2022*

### UNIST(Ulsan National Institute of Science and Technology)

M.S. IN MECHANICAL ENGINEERING

- *Thesis*: Microfluidic Concentrator Array for Quantitative Predation Study of Predatory Microbes[[online](#)]
- Got the UNIST student scholarship which is given to promising students.

*Ulsan, S.Korea*

*Mar. 2010 - Feb. 2012*

### PKNU (Pukyong National University)

B.E. IN MECHANICAL ENGINEERING

- Got BK21 scholarships which are given to top-ranked students in each Dept.

*Busan, S.Korea*

*Mar. 2002 - Feb. 2010*