SHASHANK KATIYAR

github.com/shashkat linkedin.com/in/shashkat shashank_katiyar@dfci.harvard.edu

EDUCATION

Carnegie Mellon University Pittsburgh, PA
Master of Science in Computational Biology *December 2024*
GPA: 4.10/4.33

Indian Institute of Technology Kanpur Kanpur, India
Bachelor of Technology in BioSciences and Bioengineering *June 2023*
GPA: 9.1/10 – Department Rank: 1

SKILLS

Programming: Python, R, C/C++, Golang, Shell Scripting (UNIX/Linux)
ML Techniques: Random Forest, Gradient Boosting, CNN, Reinforcement Learning, Transformers
Bioinformatics: Single Cell Data Analysis (Scanpy, Seurat), Spatial Data Analysis (spatialdata, squidpy)
Toolkit: Pytorch, Pandas, Numpy, Matplotlib, Seaborn, Git/Github, Docker, AWS, SLURM, HPC

EXPERIENCE

Computational Biologist Boston, MA
Filbin Lab, Dana Farber Cancer Institute *Jan 2025 – Present*

- Built a reproducible spatial transcriptomics (10x Xenium) analysis pipeline to support multiple projects, improving analysis efficiency and scalability across projects
- Led analysis and optimization of spatial dataset for a Nature-published study, implementing a novel ”coherence” metric to quantify tumor organization and reveal recurrence-associated spatial patterns
- Developed custom computational workflows integrating spatial transcriptomics with immunofluorescence (image) and calcium imaging data (video) to uncover celltype–specific molecular and functional patterns

Computational Biology Intern Pittsburgh, PA
Novasenta, Inc. *May 2024 – August 2024*

- Identified celltype abundances in Lung Cancer Visium samples using cell2location spot deconvolution, improving company's understanding of localization of target ligand-receptor (L-R) pairs in tumor tissue
- Quantified L-R colocalization in VisiumHD data with binomial test; got significant colocalization, $p<0.01$
- Implemented Nuclei Segmentation on VisiumHD data to get spatial-single cell data for L-R analysis and found multiple celltype-pair interactions with pvalue <0.05 according to custom-made permutation test

Machine Learning Intern Cambridge, MA
TenSixty BioSciences, Inc. *Sep 2022 – July 2023*

- Developed a scalable pipeline for rare cell marker identification in single cell data, including containerized implementation on AWS-parallel for high-throughput isoform-level alignment
- Identified a differentially expressed isoform in small cell lung cancer for targeted drug development and designed a high-accuracy protein binder using RFdiffusion

PROJECTS

Developing a bot for slither.io game using Reinforcement Learning Boston, MA
Self Project *January 2025 – Present*

- Developing a reinforcement learning bot for slither.io game, using PyTorch as framework
- Studied core deep learning and reinforcement learning concepts in depth through Dive into Deep Learning book to build the theoretical background needed for the project

PUBLICATIONS

- Single-cell multidimensional profiling of tumor cell heterogeneity in supratentorial ependymomas. **Nature (2025)**
- The fourth annual Carnegie Mellon Libraries hackathon for biomedical data management, knowledge graphs, and deep learning. **BioRxiv (2023)**