Yiming Jia

yiming.jia@mail.utoronto.ca | +1-416-828-2739 | jymmmmm.github.io | linkedin.com/in/jiayiming

Education

University of Toronto, M.Sc. in Applied Computing (MScAC), GPA: 4.0/4.0 Sep. 2023 – Jun. 2025

- Core Courses: Neural Networks and Deep Learning, Computational Imaging, Topics in Storage Systems

Beijing University of Posts and Telecommunications, B.E. in Software Engineering, Sep. 2019 – Jun. 2023
GPA: 92.04/100 (3.83/4.0), Ranking: 1/178

- Core Courses: Principles of Operating Systems, Algorithms and Data Structures, Principles of Database Systems, Compiler Principle and Technology, Computer Networks

Publication

scGeneRythm: Using Neural Networks and Fourier Transformation to Cluster Genes by Time-Frequency Patterns in Single-Cell Data Nov. 2023
Jia Y., Wu H., & Ding J.
https://doi.org/10.1101/2023.11.26.568761

Research Experience

Research Intern, Meakins-Christie Laboratories, McGill University May 2022 – Nov. 2023

- **Single Cell Data Analysis and Model Development**
  - Analyzed and preprocessed time-series single-cell data using PCA, UMAP, Clustering, and FFT.
  - Proposed a unique method to represent genes using frequency domain information.
  - Developed and trained an integrated GCN+VAE model pipeline to derive gene embeddings.
  - Demonstrated the biological significance of gene embeddings and the superiority of our method through GO Enrichment Analysis.
  - Conducted an ablation study to highlight the importance of frequency domain representation via FFT.

Work Experience

Machine Learning Engineer Intern, Kore Geosystems, Toronto, Canada May 2024 – Jan. 2025

- **Geotech Model Development and Optimization**
  - Optimized and retrained a lightweight segmentation model based on LinkNet, Achieving a 20% increase in Dice Score through iterative training improvements and in-depth analysis.
- **Conditional Denoising Diffusion Model for Image Generation**
  - Developed, trained, and optimized a conditional diffusion model using 13k labeled and over 30k unlabeled images with a classifier-free approach, enabling the generation of high-quality geological images with mask guidance.
  - Resolved color shift issues in generated images using V-score and exponential moving average.
  - Utilized a color guidance generation algorithm to fine-tune the color quality of the images.
  - Applied the diffusion model to generate a large volume of corner-case data, improving the balance and diversity of the manually labeled dataset.
- **ViT Pretraining using Dino**
  - Utilized Dino's training framework to conduct unsupervised pretraining on ViT-base and ViT-large models with over 50,000 high-resolution mineral images, effectively capturing the visual features of the mineral images.
  - To evaluate the model, I simulated image labels based on the majority class, utilizing linear probing and 3D visualization through t-SNE dimensionality reduction. This approach demonstrated superior performance compared to the backbone pretrained with VAE-based self-supervision.

Edge AI Engineer Intern, Sony, Beijing, China Jul. 2022 – Feb. 2023

- **Fine Clothing Segmentation Based on Google Object Detection API**
  - Reproduced and trained CenterMask based on Google object detection API (MobileNetV2) with the Fashionpedia dataset to achieve fine instance segmentation of clothing.
  - Optimized segmentation performance of the model in PC simulations.
  - Simplified, quantified, and transplanted the model to Sony IMX-500 chips (Float32 to Int8).
  - Evaluated the ultimate performance of the model on IMX-500 chips under real-world scenarios.