Bioinformatics Summer School Report

Name:
Waleed Alsarhani

Internship Organization:
University of Toronto, Sawh Lab

Supervisor's Name:
Dr. Ahilya Sawh

Project Title:
Development of Deep Learning Solutions for 4D Image Segmentation of Spatial Omics Data

1. Introduction
During my internship at the Sawh Lab, I focused on enhancing genomic architecture analysis through the development of advanced deep learning models for 4D image segmentation. The primary objective was to improve the accuracy and efficiency of segmenting high-resolution 3D images of C. elegans embryos. My work involved deploying and refining deep learning models like Micro-SAM to offer detailed insights into chromosomal dynamics during cell development. The project aimed to surpass traditional 3D segmentation techniques by integrating modern neural networks, thereby providing a deeper understanding of cellular and nuclear structures over developmental periods.

2. Materials and Methodology
To achieve the project's objectives, I utilized several advanced tools and methodologies. The key technologies included:
- Micro-SAM: An advanced segmentation model designed for high-precision 3D

image analysis.
- CNNs and Transformers: Implemented to enhance segmentation accuracy and computational efficiency.
- TigerVNC: Used to access a robust cluster environment for managing large datasets and intensive computations.
- Datasets: High-resolution 3D Z-stack images of C. elegans embryos, crucial for analyzing nuclear dynamics during development.
The process involved deploying Micro-SAM for segmentation, refining the model's parameters to optimize performance, and documenting all procedures on GitHub to maintain transparency and collaboration.

3. Results and Discussion

The deployment of Micro-SAM led to significant improvements in 3D automatic segmentation accuracy, outperforming traditional methods. The use of this model allowed for more efficient processing and analysis of complex cellular structures. Our results demonstrated that Micro-SAM could handle the intricacies of chromosomal dynamics with greater precision, offering valuable insights into genetic organization. These findings suggest that deep learning can significantly enhance the accuracy of bioinformatics tools, particularly in the context of spatial omics data analysis.

4. Skills and Knowledge Gained

Throughout the internship, I gained substantial technical skills, including:
- Deep Learning Models: Proficiency in implementing and optimizing CNNs, transformers, and segmentation models like Micro-SAM.
- Data Management: Experience with handling large bioinformatics datasets in a high-performance computing environment.
- Software Proficiency: Enhanced skills in using tools like TigerVNC for remote cluster access and GitHub for documentation.
In addition to technical skills, I developed soft skills such as improved teamwork, problem-solving abilities, and effective communication within a research team.

5. Challenges and Solutions

One of the significant challenges I faced was the initial difficulty in using Med-SAM, which struggled with processing 3D data. This challenge was addressed by

shifting focus to Micro-SAM, which proved to be more effective for our specific needs. Another challenge was managing the computational demands of processing large datasets, which was mitigated by utilizing a robust cluster environment and optimizing the code for better efficiency.

6. Conclusion

This internship has been a transformative experience, providing me with hands-on experience in bioinformatics and deep learning. It has solidified my interest in pursuing a career in AI-driven biological research. The key takeaway from this experience is the importance of continuous learning and adaptation, particularly in a rapidly evolving field like bioinformatics. My future aspirations include further developing my skills in deep learning and contributing to advancements in genomic research.

7. Acknowledgments

I would like to extend my sincere gratitude to Dr. Ahilya Sawh for her invaluable guidance and support throughout this internship. I also thank Natalie Kwong and the entire Sawh Lab team for their collaboration and encouragement. Additionally, I appreciate the support from King Abdullah University of Science and Technology (KAUST) and the Data Science Institute for making this internship possible.

8. References

All references and additional resources are accessible through the QR code provided on the poster.