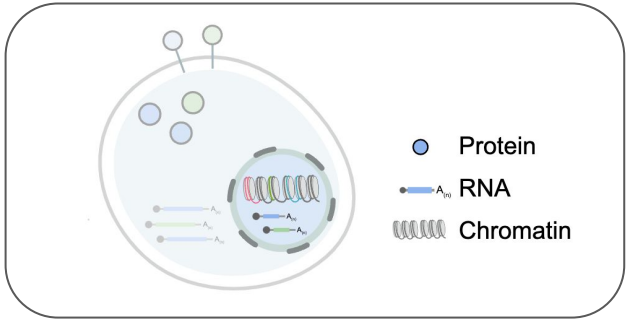
Matching Cells Between Modalities

Xinyao Fan, Ning Shen
UBC Department of Statistics

Biological Background



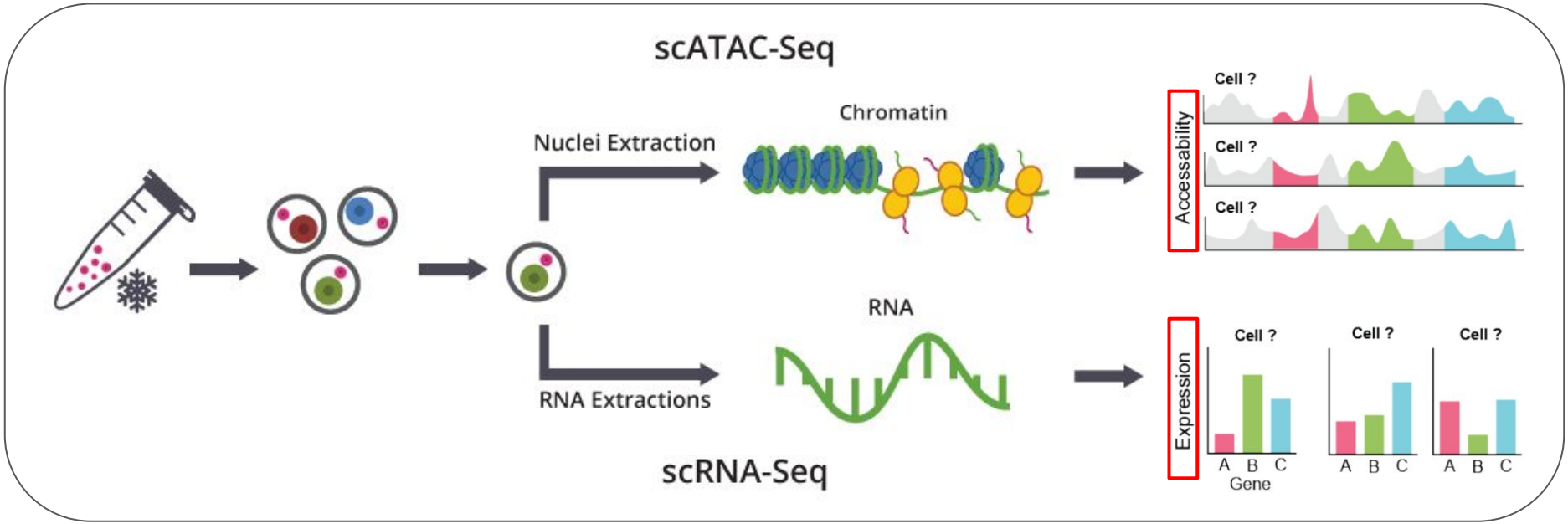
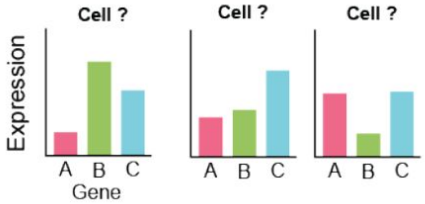
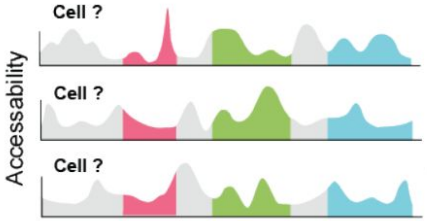


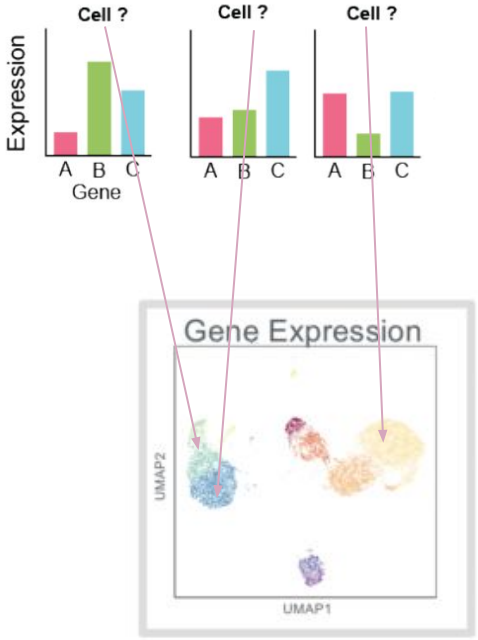
Image credited to: link1, link2

Problem Definition





Problem Definition



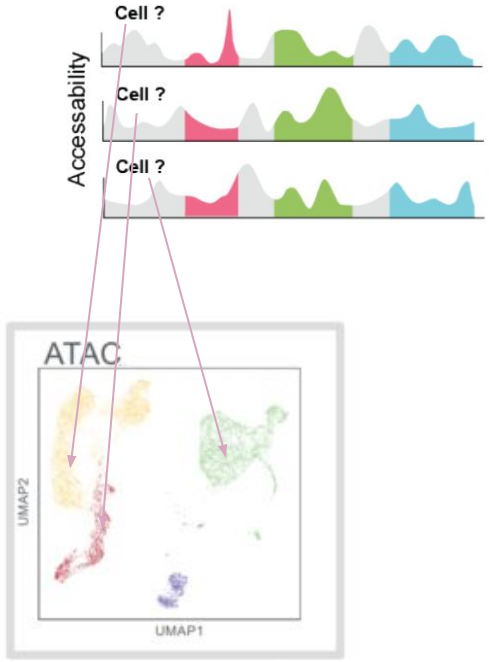
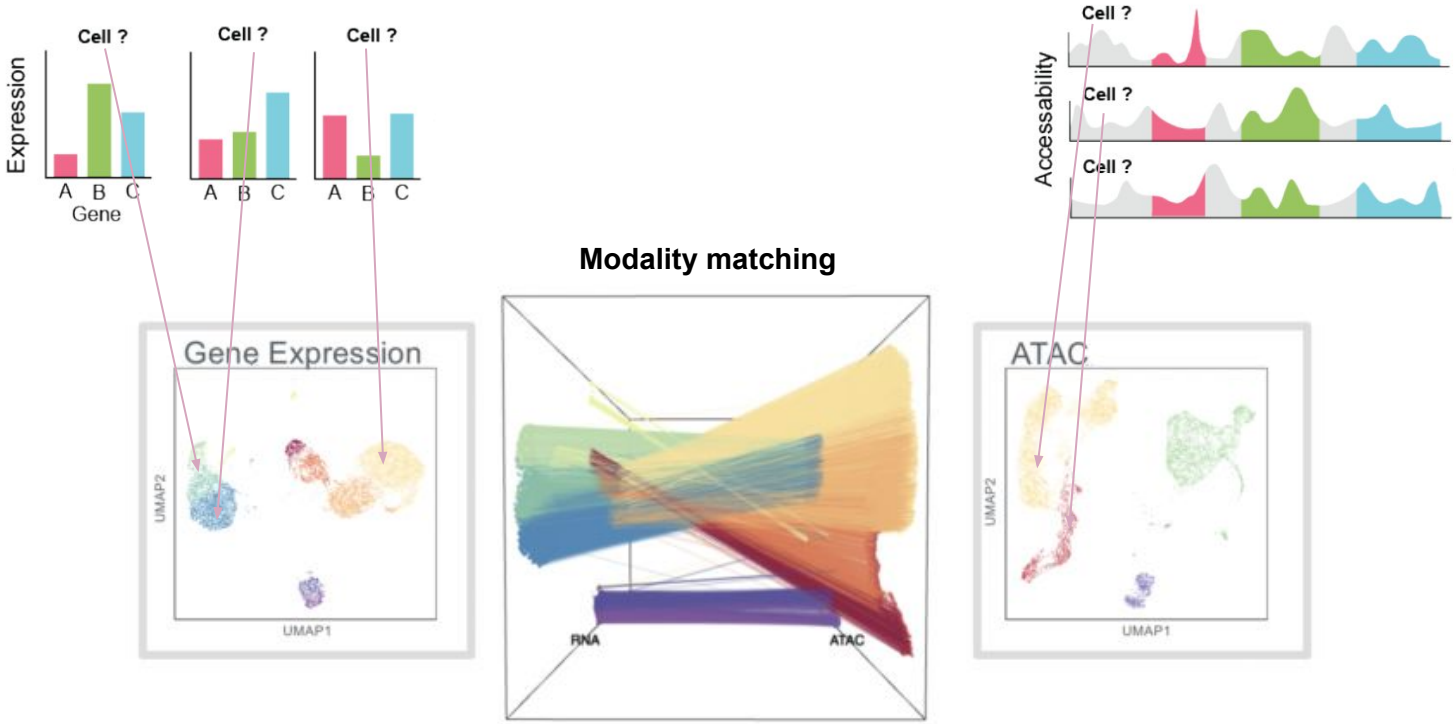


Image credited to: link1, link2

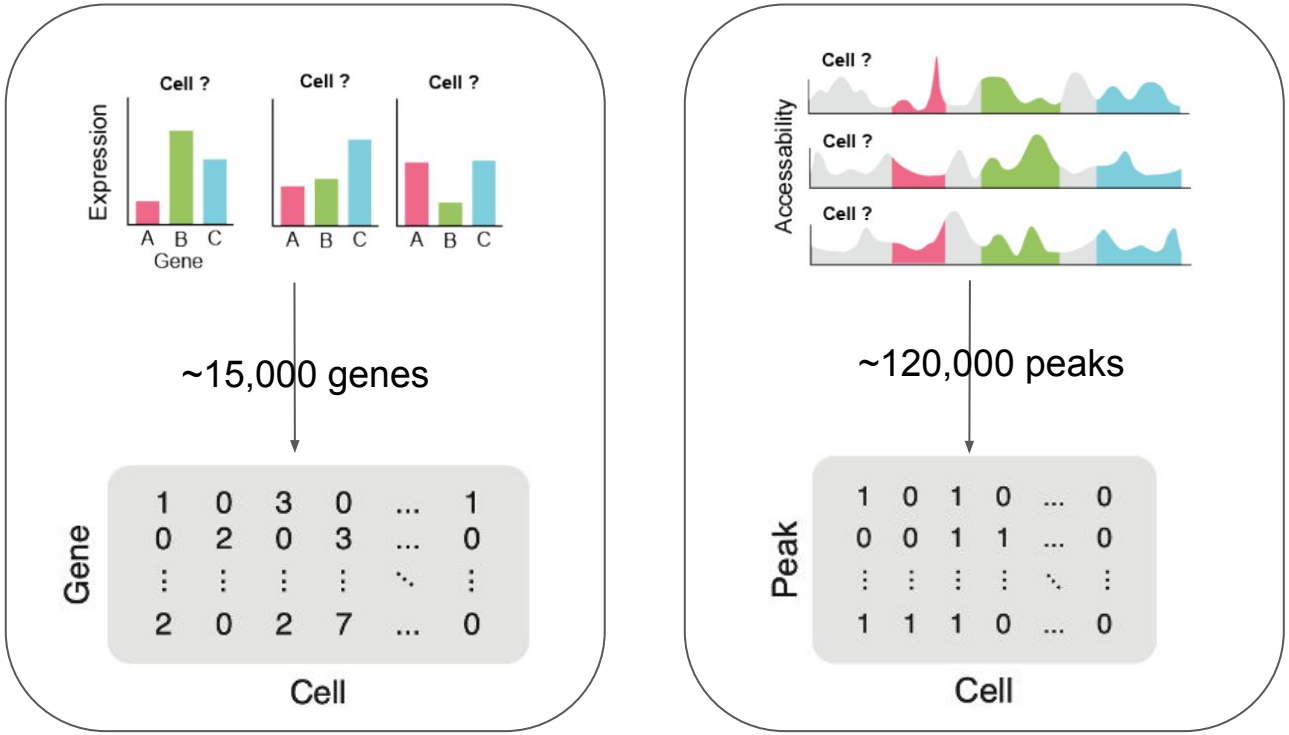
Problem Definition



- One of the tasks in '[Multimodal Single Cell Data Integration Challenge]' - a NeurIPS Competition (2021).

Image credited to: link1, link2

Data Description

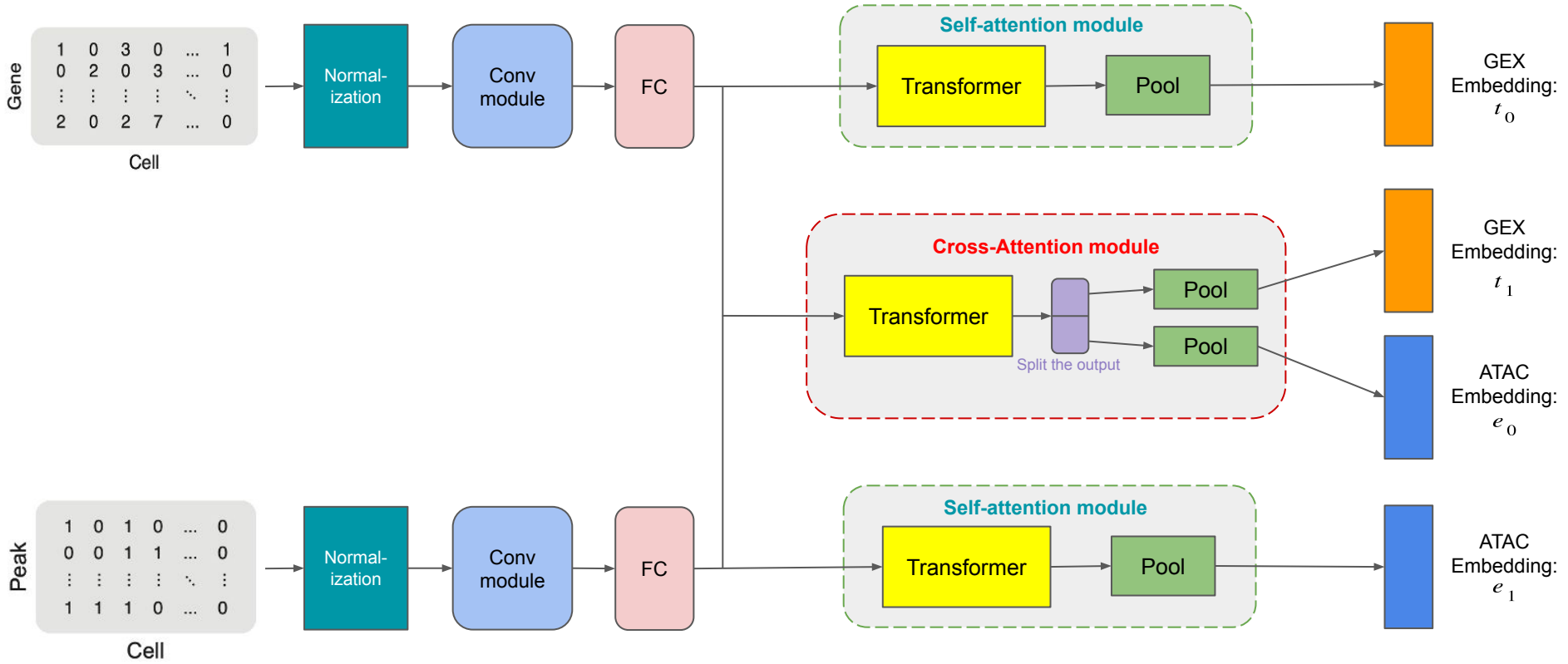


~ 70,000 cells from from 10 donors

Multi-Modality Cross Attention Network for Image and Sentence Matching

Xi Wei[1], Tianzhu Zhang[1,*], Yan Li[2], Yongdong Zhang[1], Feng Wu[1]
[1] University of Science and Technology of China; [2] Kuaishou Technology
wx33921@mail.ustc.edu.cn; {tzzhang,fengwu,zhyd73}@ustc.edu.cn; liyan@kuaishou.com

Model Overview



Loss Function

For pair RNA R and ATAC A, we learn two pairs of embeddings (i_0, c_0) and (i_1, c_1). The similarity scores are defined as

$$S(R, A) = i_0 \cdot c_0 + \alpha(i_1 \cdot c_1)$$

where α is a hyper-parameter

The loss function (bi-directional triplet ranking loss) is defined as

$$\mathcal{L} = \max[0, m - S(R, A) + S(R, \hat{A})] + \max[0, m - S(R, A) + S(\hat{R}, A)]$$

Similarity scores of matched modalities

Similarity scores of mismatched modalities

where m denotes the margins, (R, A) denotes the true matched RNA-ATAC pair, and $\hat{R} = \arg\max_{x \neq R} S(x, A)$ and $\hat{A} = \arg\max_{y \neq R} S(A, y)$.

Performance Evaluation

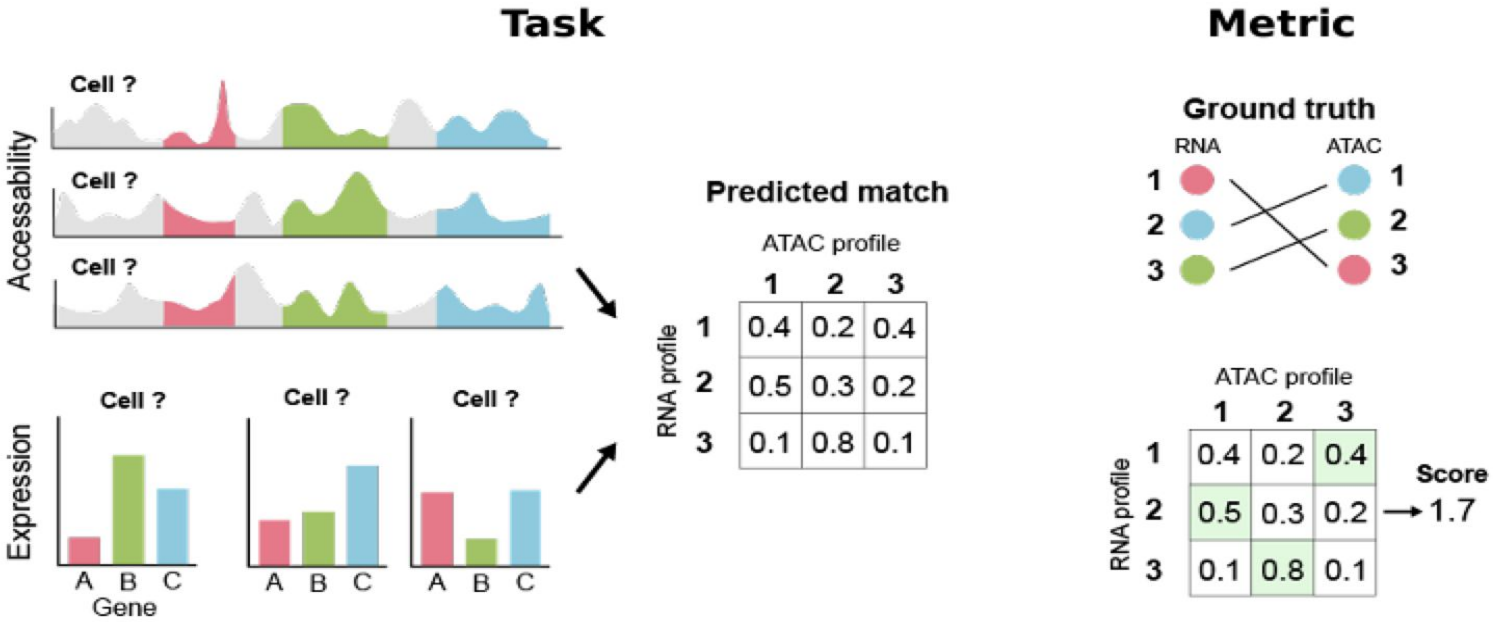


Image credited to: link

Timeline and Plans

- Nov.1 - Nov.15:
  - Data preprocessing
  - Detail finalization of network architecture
  - Proposal writing — Ning and Xinyao
- Nov.15 - Nov.20:
  - Dimension-reduced representation learning — Ning and Xinyao
- Nov.20 - Dec.7 :
  - Model implementation
  - Model modification (if applicable)
  - Hyperparameter tuning — Ning and Xinyao
- Dec.7 - Dec.14 :
  - Model fine-tuning
  - Result summarization and report writing — Ning and Xinyao

References

1. Multi-Modality Cross Attention Network for Image and Sentence Matching CVPR 2020 · Xi Wei, Tianzhu Zhang, Yan Li, Yongdong Zhang, Feng Wu
2. A sandbox for prediction and integration of DNA, RNA and proteins in single cells. (NeurIPS Datasets and Benchmarks 2021) Malte D. Luecken et al.
3. Multimodal single cell data integration challenge: Results and lessons learned (NeurIPS 2021 Competition and Demonstration Track) Christopher Lance et al.
4. Gene transformer: Transformers for the gene expression-based classification of lung cancer subtypes (preprint arXiv:2108.11833, 2021 - arxiv.org) A Khan, B Lee
5. Attention is all you need. In Advances in neural information processing systems, pages 5998–6008, 2017 Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin.
6. Semi-supervised Single-Cell cross-modality Translation Using Polarbear. Recomb2022 Ran Zhang, Laetitia Meng-Papaxanthos, Jean-Philippe Vert,William Stafford Noble