Dear Reviewers,

This supplement provides additional details supporting the paper "TrajLens: Visual Analysis for Constructing Cell Developmental Trajectories in Cross-Sample Exploration", including dataset descriptions and extended results.

We thank the reviewers and area chairs for their valuable feedback, which has helped us improve the clarity and completeness of the presented materials.

Yours sincerely,

All authors of VIS 2025 Full Papers – 1247: "TrajLens: Visual Analysis for Constructing Cell Developmental Trajectories in Cross-Sample Exploration".

1 Dataset Description

We utilized four multi-sample scRNA-seq datasets for model prediction and case studies, alongside two established benchmark datasets for GNN in Table 1. CiteSeerand CoraFull are both citation network of scientific articles, commonly used for evalation and testing in GNN.

Table 1: Summary of four multi-sample scRNA-seq datasets used for model prediction and case studies, and two benchmark datasets for GNN evaluation.

Dataset	Nodes	Stages	Publish Venue
Mouse Embryonic Spatiotemporal Development Dataset (Mouse-Embryo) [1]	1922	8	Cell
Mouse Embryonic Dorsal Midbrain Syndrome Dataset (Mouse-Midbrain) [1]	55	3	Cell
Zebrafish-Embryo [2]	38731	12	Science
Schiebinger2019 [4]	236285	19	Cell
CiteSeer [3]	3327	-	Public Dataset
CoraFull [5]	19793	-	Public Dataset

2 Extended Model Evaluation Details

2.1 Model hyperparameter choice

We investigated the impact of hyperparameter settings, including batch size and negative sampling rate (the number of negative samples generated for each positive sample), on model performance, such as accuracy, recall, and F1 score. As shown in Figure 1, the results indicate that a negative sampling rate of 5 yields optimal performance. Batch sizes from 1 to 8 showed similar performance, but larger sizes reduced training time, though excessive sizes lead to insufficient GPU memory. Consequently, we selected a batch size of 8.
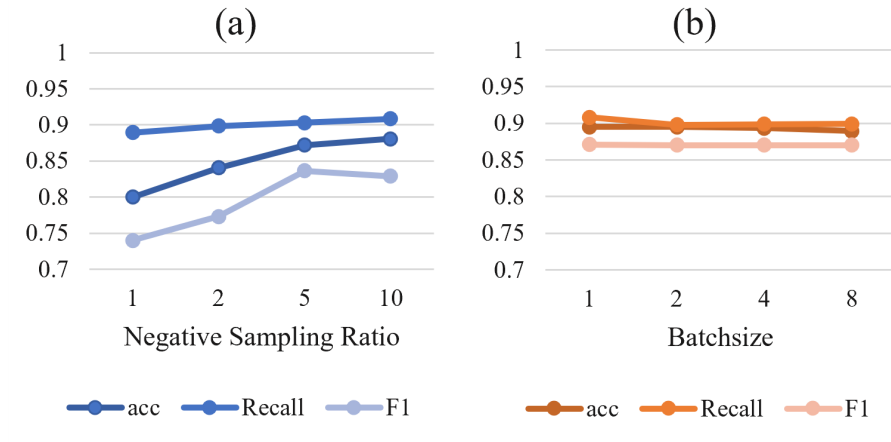
Figure 1: (A) Comparison of model performance with different negative sampling ratios. (B) Comparison of model performance with different batch sizes.

2.2 Model Runtime Evaluation

We conducted two standard experiments widely used in GNN research, total training time and prediction time on 500 edges, on three representative datasets: Mouse-Embryo, CiteSeer and CoraFull. As shown in Table 2, our model not only demonstrates good training efficiency but also provides fast predictions on new data.

Table 2: Model evaluation on training time for 30 epochs (s) and prediction time for 500 edges (ms).

Task	Mouse-Embryo	CiteSeer	CoraFull
Training (s)	187.56 ± 1.63	85.32 ± 0.09	424.55 ± 0.69
Predict 500 edges (ms)	5.85 ± 0.02	6.10 ± 0.03	71.58 ± 0.31

References

[1] A. Chen, S. Liao, M. Cheng, K. Ma, L. Wu, Y. Lai, X. Qiu, J. Yang, J. Xu, S. Hao, et al. Spatiotemporal transcriptomic atlas of mouse organogenesis using dna nanoball-patterned arrays. *Cell*, 185(10):1777–1792.e21, 2022. doi: 10.1016/j.cell.2022.04.003

[2] J. A. Farrell, Y. Wang, S. J. Riesenfeld, K. Shekhar, A. Regev, and A. F. Schier. Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science*, 360(6392):eaar3131, 2018. doi: 10.1126/science.aar3131

[3] C. L. Giles, K. D. Bollacker, and S. Lawrence. Citeseer: an automatic citation indexing system. In *Proceedings of the Third ACM Conference on Digital Libraries*, p. 89–98. Association for Computing Machinery, 1998. doi: 10.1145/276675.276685

[4] G. Schiebinger, J. Shu, M. Tabaka, B. Cleary, V. Subramanian, A. Solomon, J. Gould, S. Liu, S. Lin, P. Berube, et al. Optimal-transport analysis of single-cell gene expression identifies developmental trajectories in reprogramming. *Cell*, 176(4):928–943.e22, 2019. doi: 10.1016/j.cell.2019.01.006

[5] P. Sen, G. Namata, M. Bilgic, L. Getoor, B. Gallagher, and T. Eliassi Rad. Collective classification in network data. *AI Magazine*, 29(3):93–106, 2008. doi: 10.1609/aimag.v29i3.2157