SPARTHA: Enhancing Spatial Gene Expression Prediction with Artifact Disentanglement from Histology Images

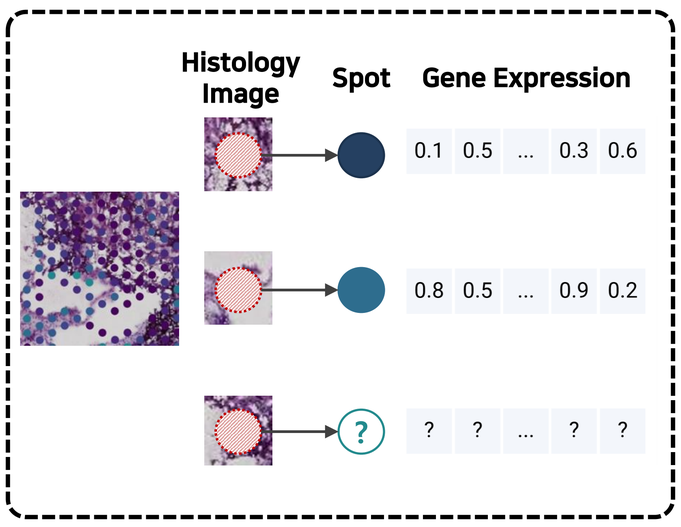
Soyon Park[1] and Jaewoo Kang[1,2] †

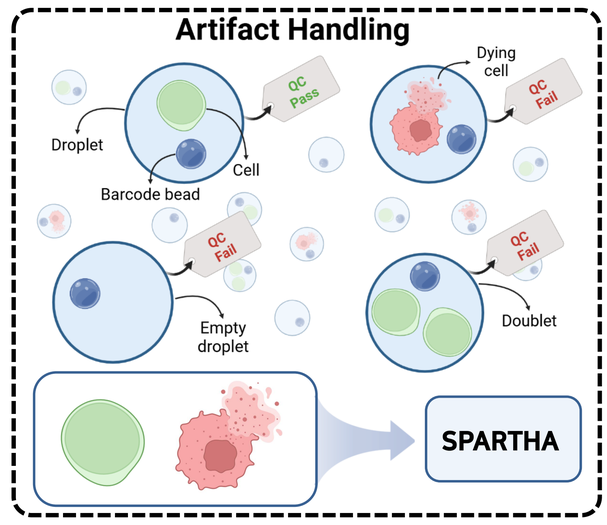
[1]Department of Computer Science and Engineering, Korea University, Seoul, Korea [2]AIGEN Sciences, Seoul., Korea

ISMB/ECCB LIVERPOOL JULY 20-24
33rd CONFERENCE ON Intelligent Systems for Molecular Biology
AND THE 24th European Conference on Computational Biology
DATES: July 20 – 24

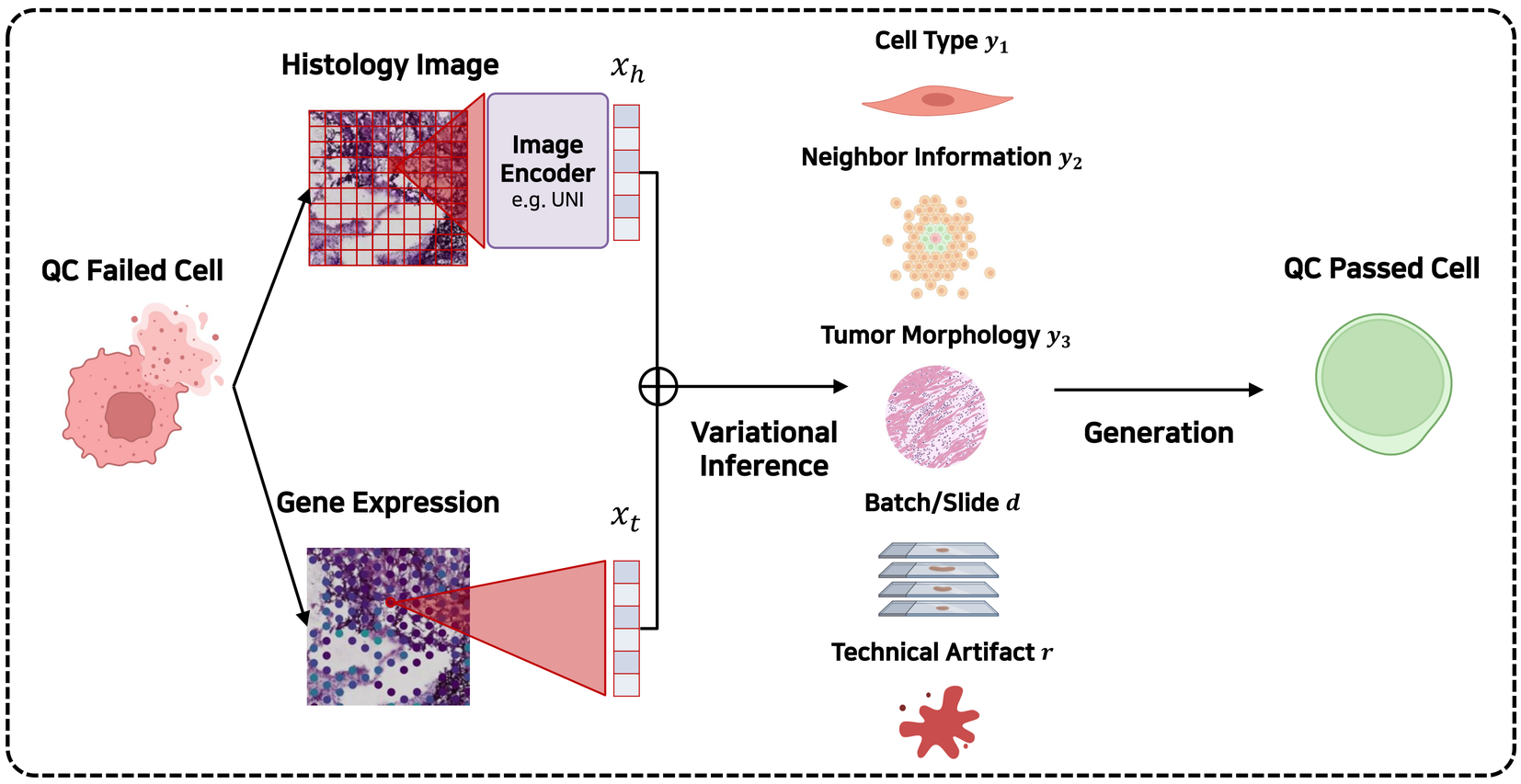
INTRODUCTION

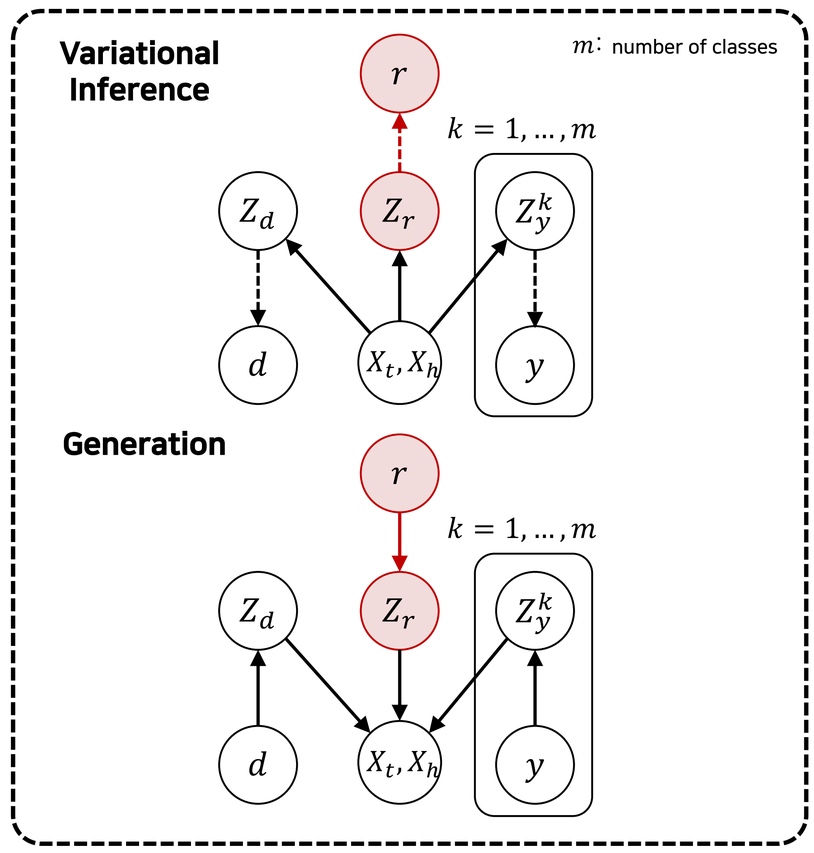
- **Task:** **Predicting RNA expression from corresponding histopathology image. Due to the high cost of spatial transcriptomics, many datasets contain abundant histology images but only a small subset with paired RNA expression.**
- **Limitations of Previous Approaches:**
  - **Strict preprocessing may result limited data availability and discard useful signals.**
  - **Filtering for highly variable genes reduces throughput.**
- **Motivation:** **Maximizing the use of all available data and throughput (number of samples and genes) while preserving potentially valuable information from QC failed data that are filtered out.**





METHODS





- **Data Preprocessing**: Instead of filtering out low-quality data, we annotate the data based on widely used quality criteria, classifying it as "pass" or "fail".
- **Modeling Class Specific Features** : Our goal is to model the underlying biological variation across spatial transcriptomics and histology modalities, while removing confounding domain effects. We achieve this through causal latent variable modeling, disentangling biological signals (e.g., cell-type, tissue morphology) from confounders (e.g., batch, artifacts).
- **Modeling Domain Specific Feature**: We explicitly model confounders (e.g., batch, slide) in a separate latent space, ensuring that biological signals remain invariant to such factors.
- **Modeling Artifacts**: The key idea is to model artifacts as residual, orthogonal to biological signals. We treat artifacts as a hidden nuisance variable and learn a dedicated latent space to capture artifact-induced variation. To guide learning, we incorporate an artifact classification objective to encourage separability between QC pass/fail patterns.

- **ELBO Loss**:

$$\mathcal{L}_{\text{ELBO}}(\theta,\phi) = \mathbb{E}_{q_\phi(Z_y,Z_d,Z_r|X_t,X_h)}\left[\log p_\theta(X_t \mid Z_y,Z_d,Z_r) + \log p_\theta(X_h \mid Z_y,Z_d,Z_r)\right]$$

$$-\sum_{k=1}^{m}\text{KL}\left(q_\phi(Z_y^k \mid X_t,X_h)\,\|\,p_\theta(Z_y^k \mid L^k)\right) - \text{KL}\left(q_\phi(Z_d \mid X_t,X_h)\,\|\,p_\theta(Z_d \mid L_d)\right) - \text{KL}\left(q_\phi(Z_r \mid X_t,X_h)\,\|\,p_\theta(Z_r \mid L_r)\right)$$

  - Z_y : Class specific latent (cell type, spatial context, morphology)
  - Z_d: Batch/Slide latent
  - Z_r : Artifact latent
- **Auxiliary Classification Loss**: KL Divergence for artifact distribution

$$\mathcal{L}_{\text{cls}} = \sum_{k=1}^{m}\mathbb{E}_{q_\phi(Z_y^k)}\left[\log q_\psi(L_y^k \mid Z_y^k)\right] + \mathbb{E}_{q_\phi(Z_d)}\left[\log q_\psi(L_d \mid Z_d)\right] + \mathbb{E}_{q_\phi(Z_r)}\left[\log q_\psi(L_r \mid Z_r)\right]$$

  - L_y: Domain-Invariant labels (cell type, spatial context, morphology)
  - L_d: Batch/Slide label
  - $L_r = y_{qc} \in \{0,1\}$: Artifact label

RESULTS

Q: Did the generated data quality improve along with the prediction performance?

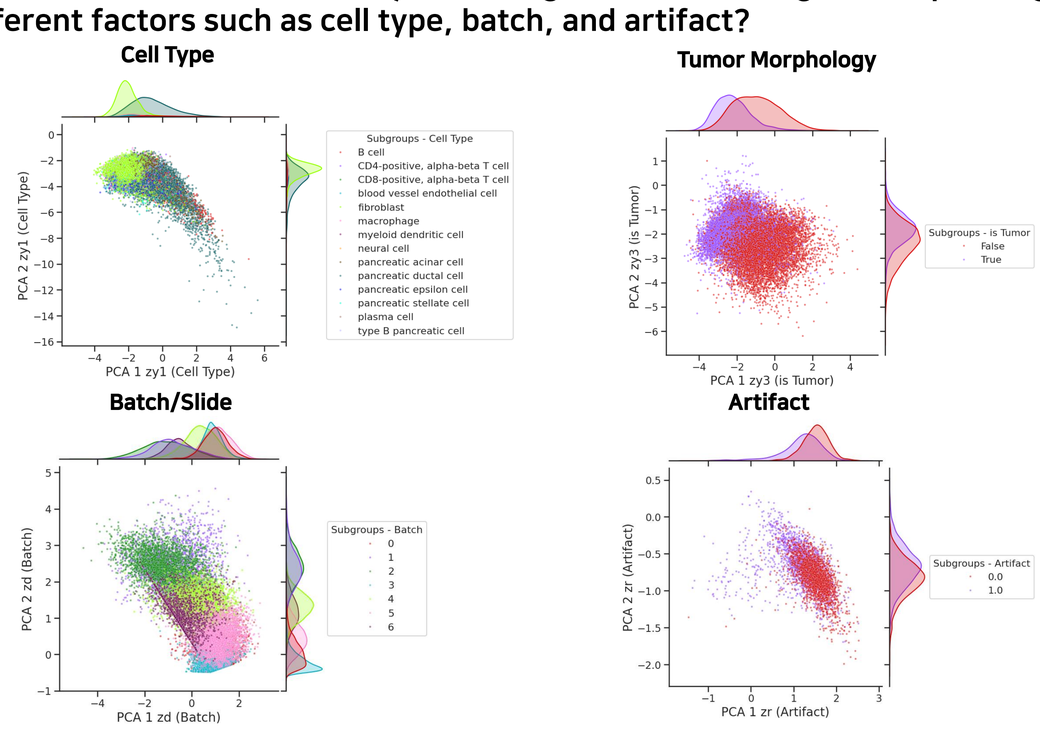
- **PCC, R^2, MAE, MSE : Estimate accuracy of predicted gene expression compared to ground truth**
- **QCPR: Proportion of generated samples that pass standard QC**
- **All: all genes in the gene expression profile (~36K)**
- **HVG: only highly variable genes (top 2500)**

All	PCC ↑	R^2 ↑	MAE ↓	MSE ↓
HisToGene	0.0803	-43.570	0.7914	1.0144
SpatialDIVA	0.7171	0.4479	**0.1302**	0.1387
Ours	**0.7209**	**0.4618**	0.1335	**0.1352**

HVG	PCC ↑	R^2 ↑	MAE ↓	MSE ↓
HisToGene	0.0317	-17.645	0.7989	1.0250
SpatialDIVA	0.6271	0.3815	**0.2587**	0.3752
Ours	**0.6289**	**0.3875**	0.2677	**0.3716**

QCPR	QCPR3 ↑	QCPR4 ↑	QCPR5 ↑
SpatialDIVA	0.6683	0.8415	0.9521
Ours	**0.795**	**0.9306**	**0.9831**

Q: Does the model successfully disentangle the embeddings corresponding to different factors such as cell type, batch, and artifact?



Information

Soyon Park
soyon_park@korea.ac.kr
DMIS Lab
https://dmis.korea.ac.kr/

Poster