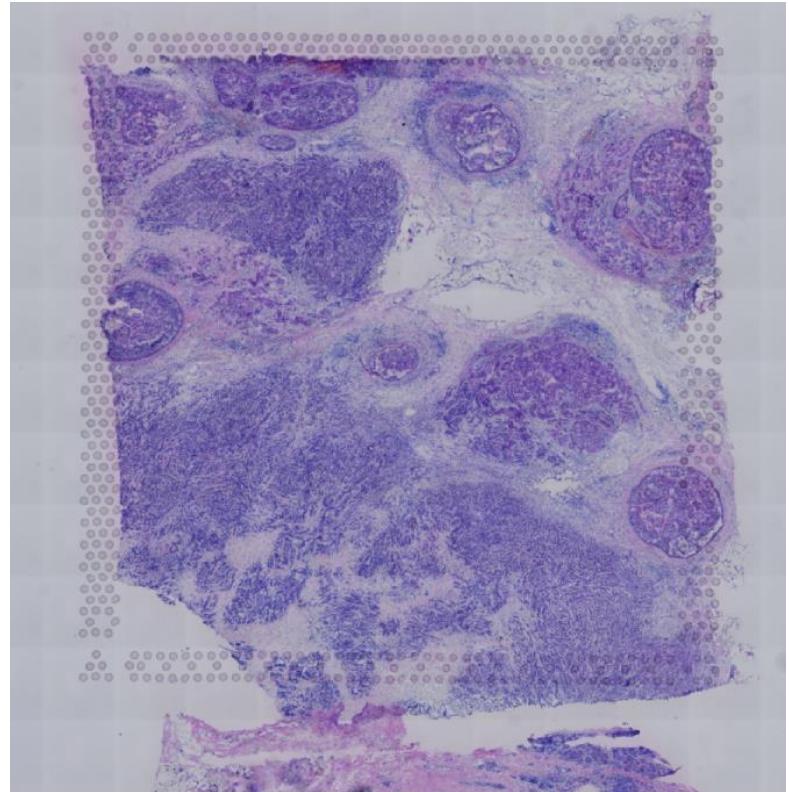


Contrastive learning methodologies on spatial transcriptomics data

Spatial transcriptomics

Transcriptomics is a branch of molecular biology that focuses on the **study of the set of information** present in a **cell, tissue, or organism** at a specific point in time.

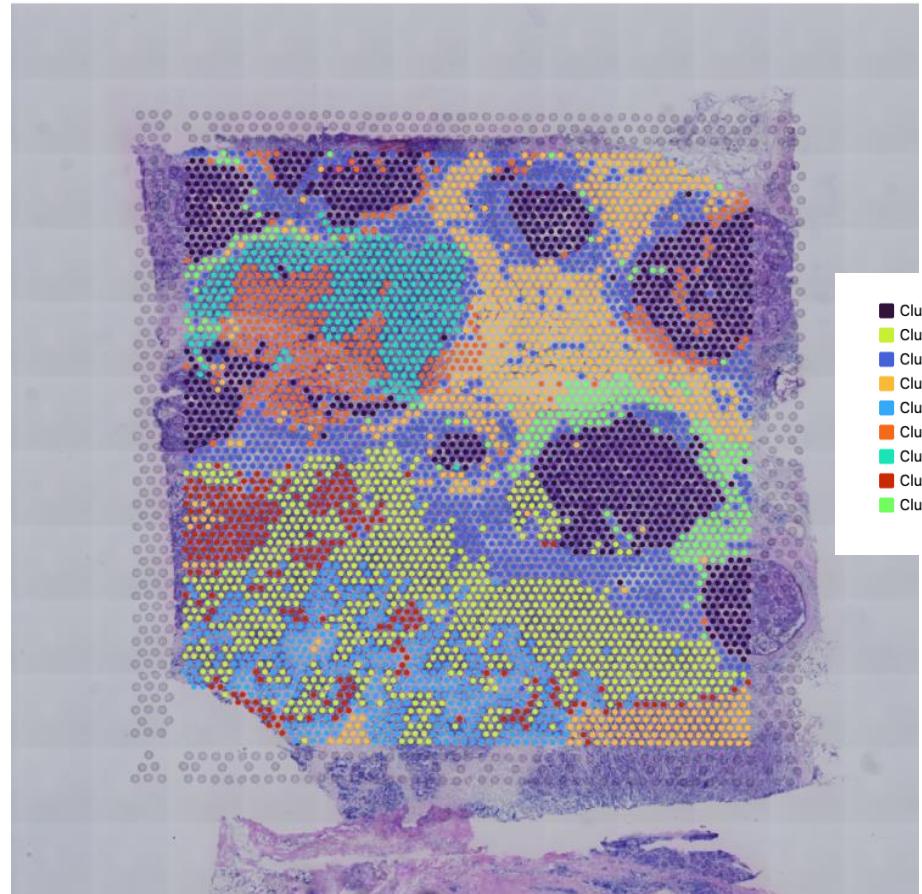
It provides a **snapshot** of the current situation of **which genes are being actively expressed** and to what extent. In recent years, this information has been **correlated with spatial information**.



ST clustering

Clustering is a computational analysis that groups **spatial units** (spots or cells) in a tissue based on the similarity of their **gene expression profile and spatial information**.

Each group represents a **spatial domain**.



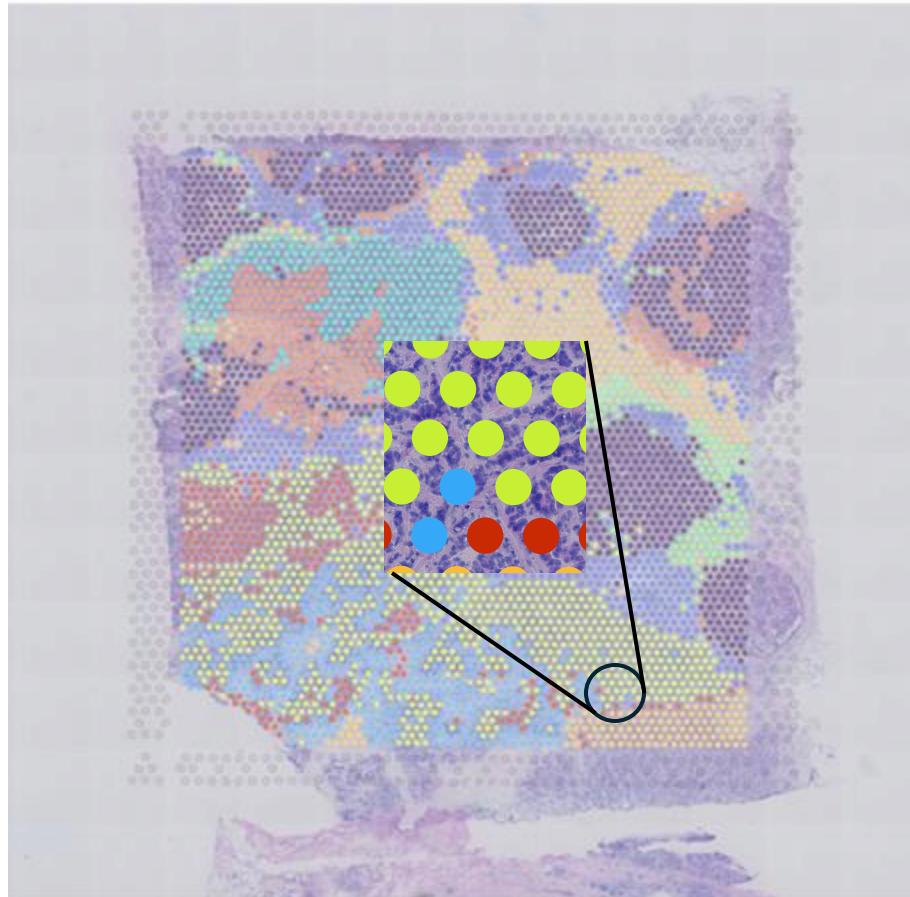
Spots

A **spot** is the fundamental unit of measurement on the surface of a tissue sample.

A **spot is not a single cell**. In many popular technologies it is larger than a single cell, so a **spot contains the genetic material for one or more cells**.

Each spot has thousands of gene expressions.

Each spot **has spatial coordinates** (x, y) that define its exact **position on the tissue**.

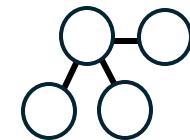
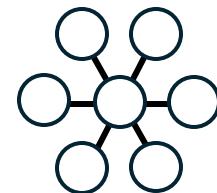


Spots representation

Spot information and spatial relationships are **represented using a graph**.

The **nodes represent the spots**, and the **edges represent the connections** between two spots.

The distance between the spots is computed using the **Euclidean distance**.



Contrastive learning

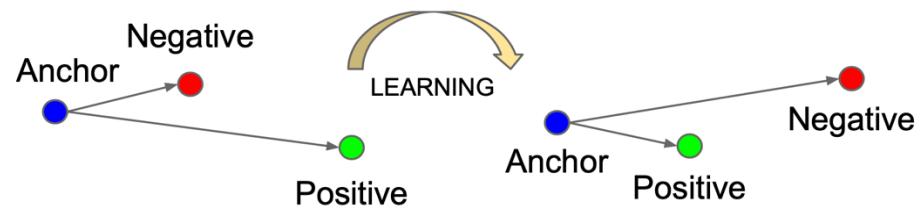
Contrastive learning is a **self-supervised learning technique**.

Its primary goal is to learn data representations (or embeddings) that are both **informative and discriminative**.

The fundamental principle is to teach a model to **distinguish what is ‘similar’ from what is ‘dissimilar’**.

To achieve this, **positive and negative pairs** are created.

The scope of the model is to **maximise the similarity of positive pairs while minimising the negative ones**.



State-of-the-art methods

State-of-the-art methods combine different technologies to maximise the embedding and the final clustering.

Among these can be found:

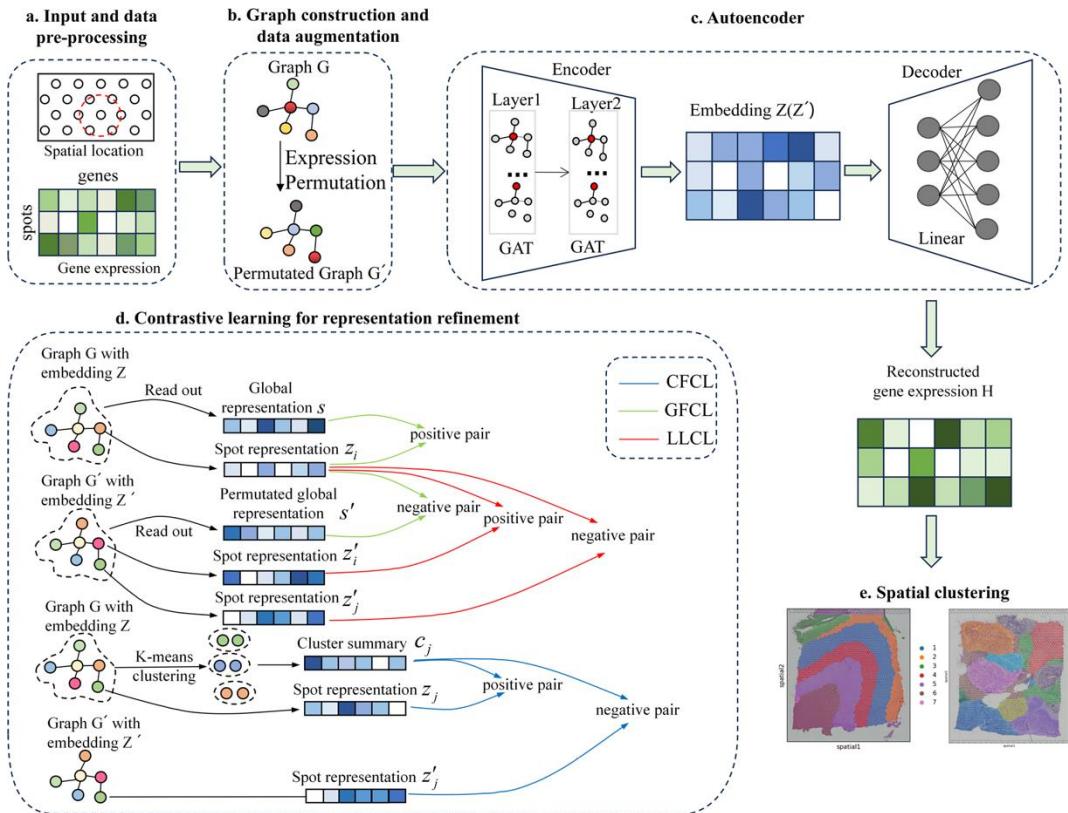
- GAAEST
- stCluster
- GraphST

GAAEST

GAAEST proposes as a **generalised deep learning method** that integrates both **spatial location details** and **gene expression data** from transcriptomics.

This framework consists of five main components:

- Data preprocessing.
- Neighbour graph construction and data argumentation.
- Autoencoder.
- Contrastive learning for embedding refinement.
- Spatial clustering.

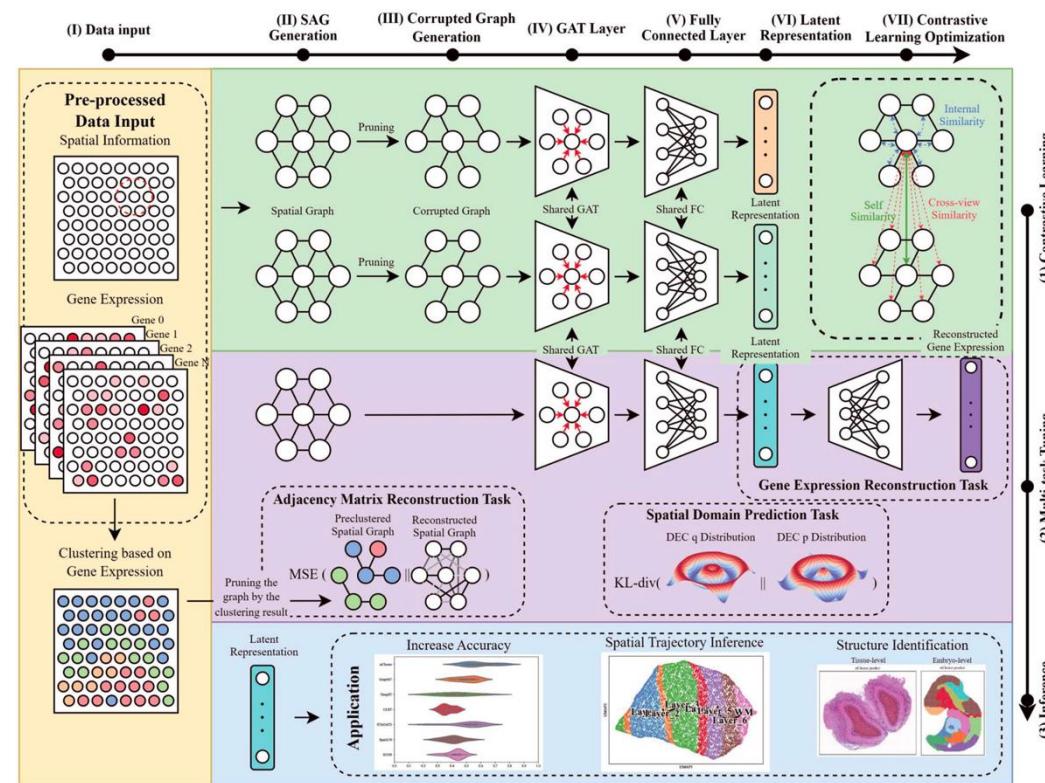


stCluster

stCluster proposes a novel method that **integrates graph contrastive learning and multi-task learning** to refine informative representation for **spatial transcriptomic data**.

This framework consists of three main modules:

- Contrastive learning.
- Multi-task tuning.
- Data inference/clustering.

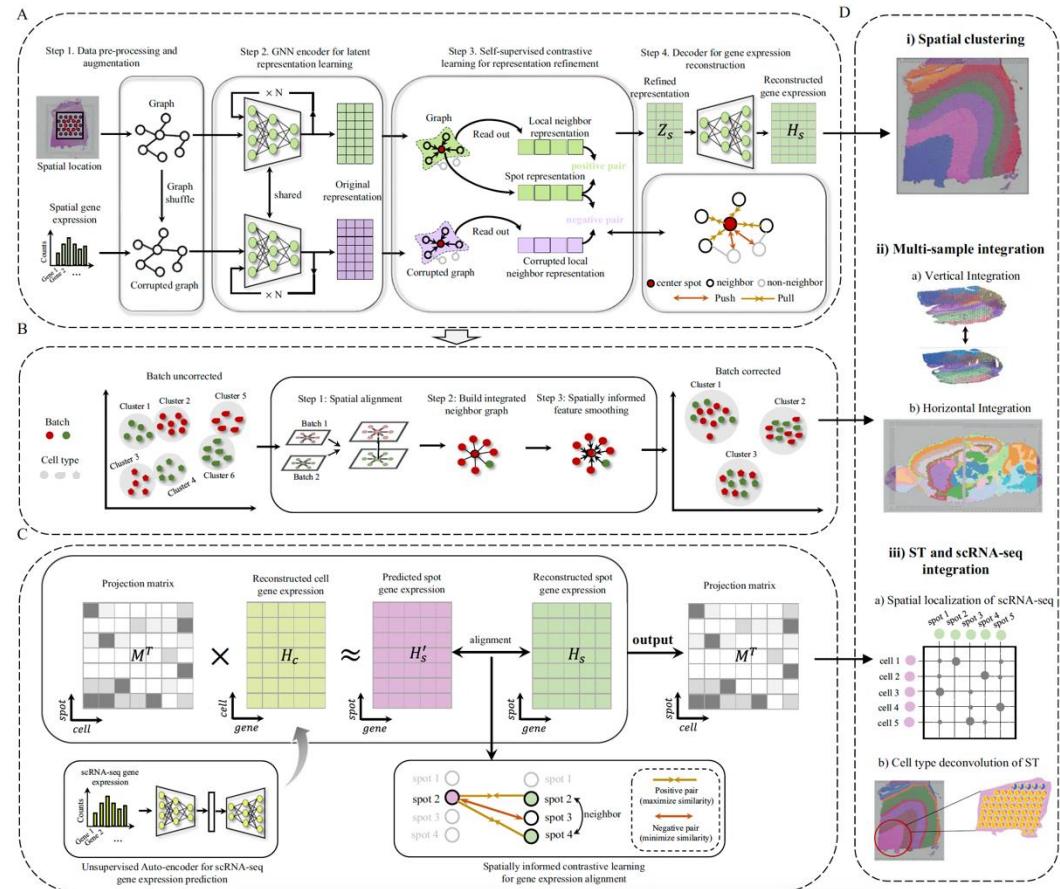


GraphST

GraphST proposes as **graph self-supervised contrastive learning** method that fully exploits spatial transcriptomics data, by **combining graph neural networks** with self-supervised contrastive learning.

This framework is composed of **three modules**.

Only the first one is considered.

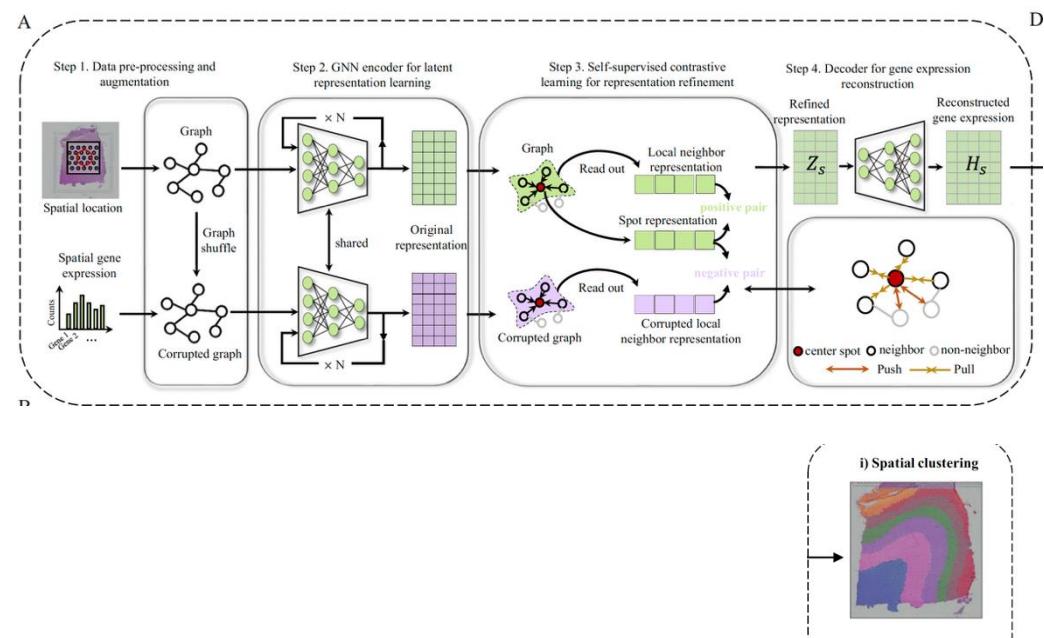


GraphST

GraphST's first module is devoted to **spatially informed clustering**.

It involves four main parts:

- a) Data pre-processing and argumentation
- b) GNN encoder for latent representation learning.
- c) Self-supervised contrastive learning for representation refinement.
- d) Decoder for gene expression reconstruction.
- e) Clustering



High level structure

All the framework proposes almost the same structure:

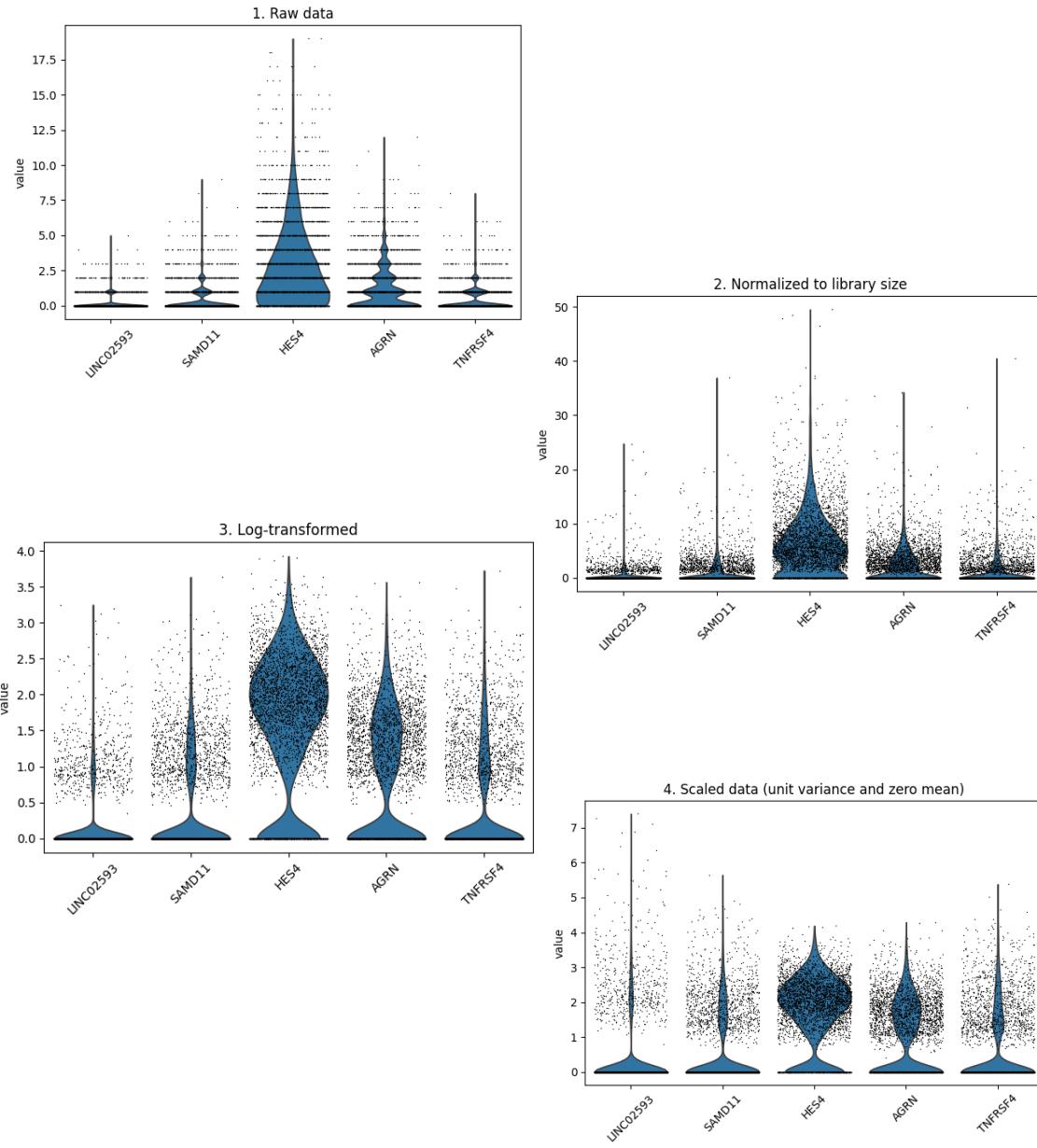
- a) Pre-processing.
- b) Graph construction.
- c) Data argumentation (generation of a corrupted version of the graph).
- d) The spots' information (graphs) and the gene expression matrix are encoded in a latent representation.
- e) Latent representation is used to reconstruct the original expression matrix.
- f) Contrastive learning is used to improve the encoder.
- g) Additional techniques used to improve the encoder.



Pre-processing

During pre-processing, the **gene expression matrix** is **normalised** and reduced to the 3,000 highly variable genes.

The normalization process includes normalization to library size (2), log-transformation (3), and scaling to have unit variance and zero mean (4).

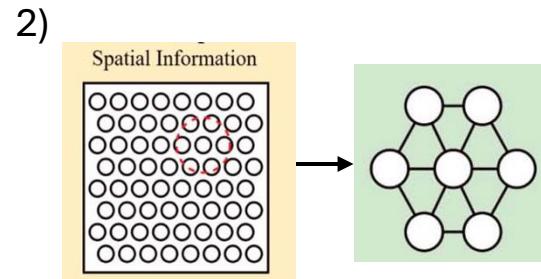
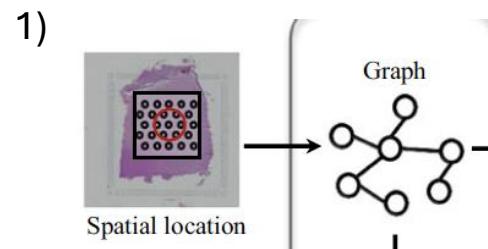


Graph construction

Graphs are built using different techniques.

GAAEST and GraphST consider as neighbours the **k-nearest spots** (1).

stCluster considers **two spots neighbours** if their **distance is less than a hyperparameter** (2).

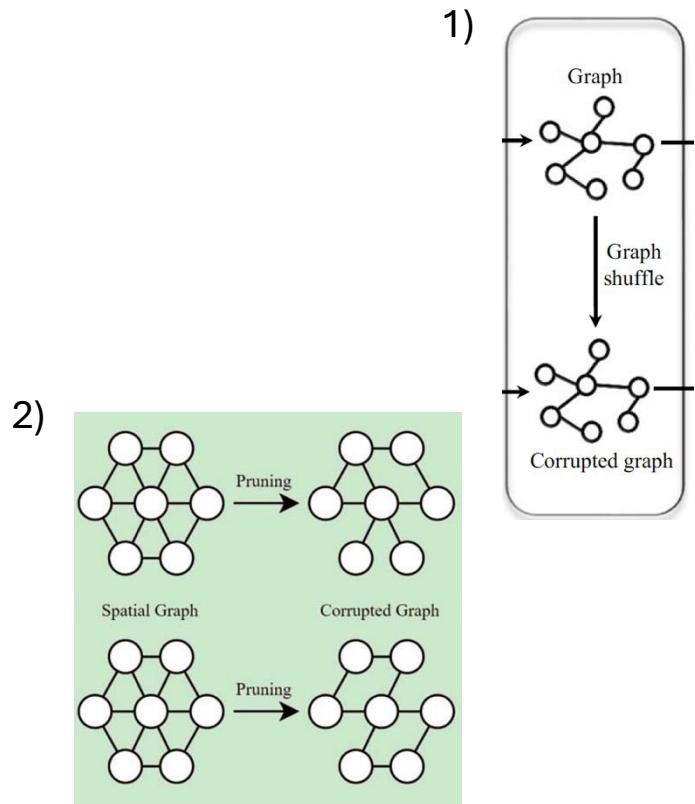


Data argumentation

Data are argued using different approaches.

GAAEST and GraphST generates a **corrupted graph by randomly shuffling gene expressions among the spots** (1).

stCluster generates **two corrupted version** of the original graph by **randomly pruning edges** (2).



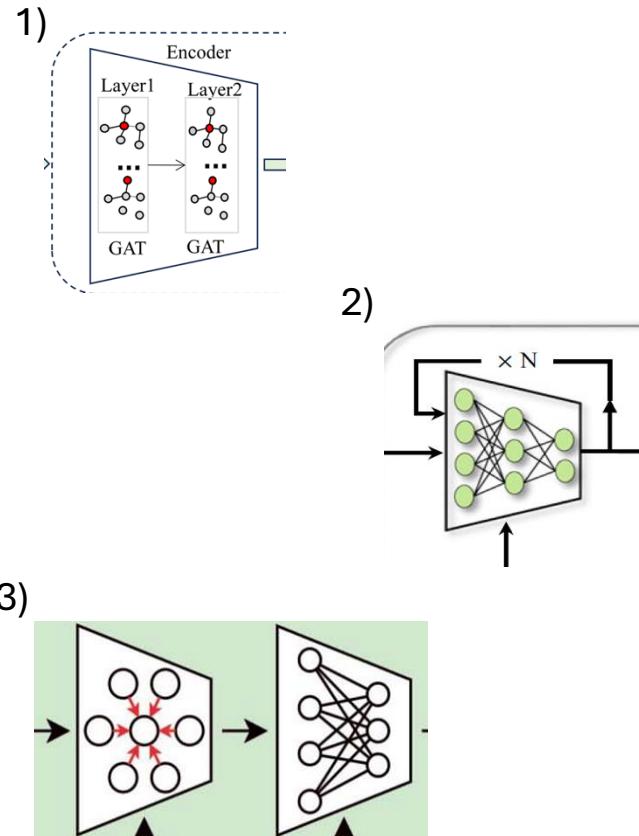
Encoding

The **first dissimilarity** between the methodologies is in the encoding step. It is devoted **to learn spot representations** that **capture** the informative parts of the **gene expression and spatial locations**.

GAAEST (1) provides an encoder with **two layers of GATs**.

GraphST (2) provides an encoder with **two layers of GNN**.

stCluster (3) provides **a layer of GAT** and a **fully connected layer**.



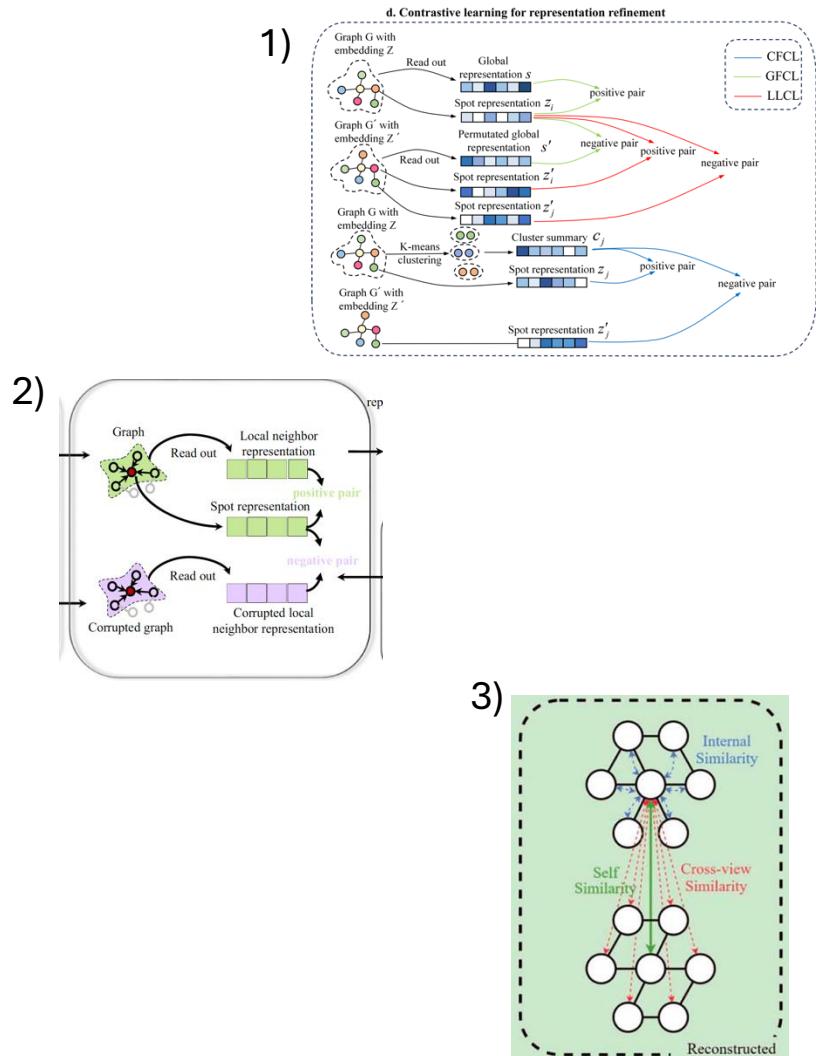
Contrastive learning

All methodologies present at least one contrastive learning technique used to enhance the model's ability to create a more accurate latent representation.

GAAEST (1) provides a **three-level strategy**: local, global and context.

GraphST (2) provides a strategy that focuses on **local spatial context**, considering a neighbourhood as a spot **microenvironment**.

stCluster (3) provides a strategy that focuses on **differentiating spots**.

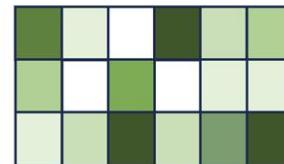


Clustering

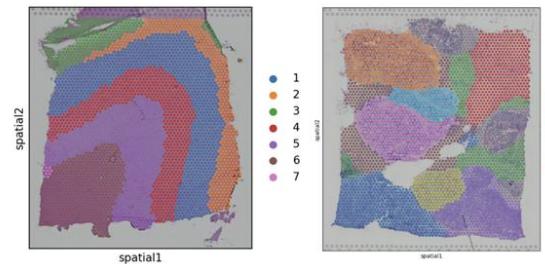
After the training, all the frameworks are used in conjunction with a community detection algorithm.

This algorithm works on the reconstructed gene expression matrix created by the decoders from the latent representation.

↓
Reconstructed gene expression H



e. Spatial clustering



Training

The models are **trained using a combination of loss functions.**

The losses are categorizable as **reconstruction loss** or **contrastive loss**.

These **keep in consideration contrastive learning approaches** used by the framework and **additional tasks** like reconstruction of gene expression matrix or multi-task tuning (stCluster).

References

- Williams, C.G., Lee, H.J., Asatsuma, T. et al. An introduction to spatial transcriptomics for biomedical research. *Genome Med* **14**, 68 (2022). <https://doi.org/10.1186/s13073-022-01075-1>
- Wang, T., Zhu, H., Zhou, Y., Ding, W., Ding, W., Han, L., Zhang, X.: Graph attention automatic encoder based on contrastive learning for domain recognition of spatial transcriptomics. *Communications Biology* **7**(1), 1351 (2024) <https://doi.org/10.1038/s42003-024-07037-0>
- Long, Y., Ang, K.S., Li, M., Chong, K.L.K., Sethi, R., Zhong, C., Xu, H., Ong, Z., Sachaphibulkij, K., Chen, A., et al.: Spatially informed clustering, integration, and deconvolution of spatial transcriptomics with GraphST. *Nature Communications* **14**(1), 1155 (2023) <https://doi.org/10.1038/s41467-023-36796-3>
- Wang, T., Shu, H., Hu, J., Wang, Y., Chen, J., Peng, J., Shang, X.: Accurately deciphering spatial domains for spatially resolved transcriptomics with stcluster. *Briefings in Bioinformatics* **25**(4), 329 (2024) <https://doi.org/10.1093/bib/bbae329>
- Veličković, Petar, et al. "Graph attention networks." *arXiv preprint arXiv:1710.10903* (2017).

Thanks for the attention