Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Spatial GNN Analyzer: An AI-Powered Graph Neural Network Framework for Spatial Transcriptomics Domain Analysis and Visualization

Subhasankar Khilar
Biotecnika https://orcid.org/0009-0003-9280-7249
Dr. Natarajan Elamathi
elamathi@biotecnika.org
Biotecnika https://orcid.org/0000-0002-1944-9137

Research Article

Keywords: Spatial Transcriptomics, Graph Neural Networks, Graph Convolutional Networks, Spatial Domain Identification, Deep Learning, Spatial Omics, Tissue Architecture Analysis, Bioinformatics, Spatial Clustering, Computational Biology, Interactive Visualization, Representation Learning.

Posted Date: August 6th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10600213/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Spatial GNN Analyzer: An AI-Powered Graph Neural Network Framework for Spatial Transcriptomics Domain Analysis and Visualization

Subhasankar Khilar[1], Dr. Elamathi Natarajan[2]

Researcher[1], Guide[2]

Biotecnika Info Labs Pvt Ltd, Bangalore, India

subhasankarkhilar@gmail.com, elamathi@biotecnika.org

Abstract:

Spatial transcriptomics has developed as a game-changer technology for understanding tissue organization by converting them into spatial context of gene expression. Due to the high dimensionality & complex spatial dependencies within transcriptomics, that is challenging to accurately identify spatial domains and uncover biologically meaningful tissue structures. This work presents Spatial GNN Analyzer, which is an AI-powered analytical frame-work it integrated with Graph Neural Networks (GNNs) with spatial transcriptomics to enable robust spatial domain identification, interactive tissue analysis, representation learning & an end-to-end pipeline.

Here, my proposed framework models spatial transcriptomics data as a graph representation in which all the tissue spots are represented as nodes whereas the spatial neighbourhood relationships are encoded as graph edges. The system learns both local cellular interactions and global tissue architecture which are biologically informative latent embeddings capable of capturing by combining transcriptomic profiles with spatial topology through Graph Convolutional Networks (GCNs). The pipeline works a comprehensive preprocessing steps like quality control, normalization, highly variable gene selectin, dimensionality reduction, & spatial graph construction followed by graph-based clustering & domain segmentation.

It utilizes Scanpy & Squidpy for preprocessing & spatial graph generation, followed by graph-aware clustering & visualization. By using Leiden clustering & UMAP embedding the spatial coherent tissue was identified and visualization was done respectably. To figure out clustering performance & biological coherence, the framework integrated metrics such as Silhouette score & Spatial smoothness, to ensure transcriptomics separability & spatial continuity of predicted domains. The Silhouette score is 0.37 which indicates effective preservation of local tissue structure & the Spatial smoothness score is 0.93, which shows a strong spatial coherence and meaningful cluster separation of Visium breast cancer spatial transcriptomics. An interactive visualization platform was developed to do exploration of spatial clusters, tissue morphology, & gene expression patterns through dynamic multi-panel visual analytical and another best thing is automated report generation facility.

This framework was validated using publicly available 10x Genomics Visium datasets, including breast cancer spatial transcriptomics samples, that shows the improved preservation of tissue architecture & enhanced spatially coherent clustering compared with conventional non-spatial analytical approaches. This work will provide a scalable & interpretable computational solution for next-generation spatial biology research as it integrated with deep graph learning, spatial omics

analysis, & an interactive visualization into a uniform platform. The proposed work has a broad application in tumour microenvironment analysis, biomarker discovery, tissue heterogeneity characterization, & precision medicine.

Keywords:

Spatial Transcriptomics; Graph Neural Networks; Graph Convolutional Networks; Spatial Domain Identification; Deep Learning; Spatial Omics; Tissue Architecture Analysis; Bioinformatics; Spatial Clustering; Computational Biology; Interactive Visualization; Representation Learning.

I. Introduction

Within the biological tissue, the spatial organization of cells plays a critical role in regulating cellular communication, immune response, tissue development, & disease progression. To deciphering complex biological system & pathological mechanism we have to understand how cells interacting in their native microenvironment. In recent years, the high throughput sequencing technology were significantly developed transcriptomics profiling, however accurately preserving the spatial arrangement of gene expression in tissue remains a main challenge in modern computational biology [1] [2].

Conventional transcriptomics methods like, single-cell RNA (sc-RNA-seq), & bulk RNA sequencing, have used to understanding gene expression dynamics & cellular heterogeneity. They have analytical power, but still these methods inherently disrupt the spatial architecture of tissue during sample preparation, that results loss of positional information & neighbourhood interactions among cells. The absence of spatial context limits comprehensive interpretation of tissue organization, intercellular communication & surrounding microenvironmental conditions, these conditions are highly influencing to the cellular function [2].

Spatial transcriptomics has emerged as a transformative technology that can simultaneously capture genome-wide gene expression profiles while preserving the spatial localisation of tissue regions. Spatial transcriptomics integrates molecular information with spatial coordinates to enable detailed investigation of tissue architecture, cellular ecosystems, developmental biology, and tumour microenvironments at unprecedented resolution. The technology has shown remarkable potential across a range of diverse biomedical applications, such as cancer research, neuroscience, immunology, and precision medicine. Spatially resolved transcriptomic analysis has become increasingly important for studying tumour heterogeneity, immune infiltration, and spatially organized cellular interactions within complex tissue systems [2] [3].

Recent advancements in artificial intelligence and deep learning have introduced powerful methodologies for analysing complex biological datasets. Graph Neural Networks (GNNs) have garnered significant attention due to their capacity to effectively learn from graph-structured data. Several graph-based computational methods have been recently proposed for spatial transcriptomics analysis, including SpaGCN, STAGATE, GraphST, and DeepST [4] [5]. Graph learning approaches have shown the potential for capturing spatial dependencies and enhancing clustering performance. Many existing methods primarily focus on algorithmic modelling, but often provide limited support

for integrated visualization, interactive analytics, deployment-ready workflows, and end-to-end analytical frameworks. Some approaches also have limitations in biological interpretability, scalability, or user accessibility, which restricts their practical use in broader spatial omics research. To assuming these limitations, the work presents Spatial GNN Analyzer, an AI-powered graph learning framework that designed for spatial transcriptomics analysis & spatial domain characterization. This framework is integrated with transcriptomic preprocessing, spatial graph construction, graph neural network-based representation learning, clustering, quantitative evaluation, interactive visualization, and automated report generation into a unified analytical ecosystem. The spatial transcriptomics data is modelled as graph structure in which tissue spots are represented as graph nodes & spatial neighbourhood relations as graph edges. So that it can learn the biologically informative local tissue continuity & global spatial organization.

The proposed work pipeline focuses preprocessing strategy that includes quality control, normalization, highly variable gene selection, dimensionality reduction, and spatial graph generation using Scanpy and Squidpy. Leiden clustering is used for Graph-aware clustering & spatial domain segmentation. And Uniform Manifold Approximation and Projection (UMAP) is used for low-dimensional embedding visualization. To evaluate the clustering quality & biological coherence the framework integrates metrics like Silhouette Score & Spatial Smoothness. Along with an interactive visualization platform was developed to exploration of spatial domains, tissue morphology, & gene expression through multi-panel analytical visualization & automated report generation.

The proposed work has done by using publicly available Visium spatial transcriptomics datasets of breast cancer tissue sample of 10x Genomics. It integrates graph-based deep learning, spatial omics analysis, clustering evaluation, and interactive visualization into a scalable computational platform, Spatial GNN Analyzer provides an interpretable and deployment-ready solution for next-generation spatial biology research.

The major contributions of this work are summarized as follows:

1. Development of an AI-powered Graph Neural Network framework for spatial transcriptomics analysis and spatial domain identification.
2. Integration of transcriptomic features and spatial topology through graph-based representation learning using Graph Convolutional Networks.
3. Implementation of a comprehensive end-to-end analytical pipeline including preprocessing, graph construction, clustering, evaluation, and visualization.
4. Incorporation of quantitative evaluation metrics, including Silhouette Score and Spatial Smoothness, for assessing clustering quality and biological coherence.
5. Development of an interactive visualization and automated reporting platform for intuitive exploration of spatial transcriptomics datasets.
6. Validation of the proposed framework using publicly available 10x Genomics Visium datasets for biologically meaningful spatial domain characterization.

II. Related Work

The recent advanced fields in Spatial transcriptomics have accelerated in the tissue architecture analysis, spatial domain identification, & spatially informed representation learning. There are several machine learning and graph-based frameworks have been proposed to integrate transcriptomic information with spatial neighbourhood relationships for biologically meaningful tissue characterization due to the increasing availability of high-resolution spatial omics datasets [6] [7].

For Spatial transcriptomics, the traditional computational methods relied on conventional clustering & dimensionality reduction methods like PCA (Principal Component Analysis), t-SNE (t-distributed Stochastic Neighbour), & Leiden clustering [8]. Generally, these focuses on expression similarity without decently accumulating spatial continuity topology, while it demonstrated effectiveness in transcriptomic feature extraction & visualization. Therefore, spatially adjacent with same biological functions were often fragmented into different clusters, by limiting biological interpretability [2].

To pointing these disadvantages, the graph-based deep learning approaches have developed as a strongest method for spatial transcriptomics analysis. The addition of gene expression profiles, spatial coordinates, & high histological information using Graph Convolutional networks (GCNs) for Spatial domain identification these were introduced by SpaGCN among the traditional graph-driven frameworks [2]. SpaGCN shows that the graph convolution-based feature work could significantly increase spatial coherence & predicted domain in biological relevance. Yet, here my framework mainly works on spatially variable gene detection & identification of domain, with a short illustrate on interactive visualization & analytical workflow.

In addition with, some statistical & probabilistic frameworks also developed for spatial domain analysis. To improve spatial clustering resolution in Visium datasets BayesSpace integrated Bayesian statistical modelling by leveraging neighbourhood information during domain segmentation [9]. DeepST introduced a deep learning-based spatial transcriptomics framework which integrates spatial location & transcriptomics features for tissue architecture characterization [10]. As these methods improved spatial clustering performance. Among them some are complex in computationally difficulty in deploy, or limit in terms of interactive analytical support.

Recent research has shown that, the integrated computational ecosystem, that combines artificial intelligence, graph learning, visualisation & user-friendly diagnostic workflows are becoming increasingly crucial for next generation spatial omics research [7] [11] [12] [13]. Some existing frameworks mainly focus on secluded analytical tasks, rather than offering scalable end-to-end functionality, that can manage preprocessing graph construction, clustering, evaluation, visualisation & automated reporting within a single unified system.

Here proposed framework, Spatial GNN Analyzer, it integrates graph neural network-based representation networks, transcriptomic preprocessing, spatial graph creation, clustering assessment, interactive visualisation & automated report generation to fill research gaps. My system prioritises biological interpretability, spatial coherence, deployment-ready analytics, & user centric exploration over algorithmic performance in spatial transcriptomics datasets. The framework combines graph based deep learning with dynamic visualisation & evaluation methods to offer a scalable & interpretable solution for spatial domain analysis & tissue architecture characterization.

III. METHODOLOGY

3.1 Dataset

This proposed framework is done by using publicly available spatial transcriptomics datasets which is generated through 10x Genomics Visium platform. Mainly the Breast cancer spatial transcriptomics samples were used for model development, clustering analysis, & tissue architecture characterization. The spatial data set contains a high-dimensional gene expression with corresponding coordinates representing physical positions of tissue spots.

Every spatial spot contains transcriptomic information which is derived from multiple neighbouring cells, enables the analysis of localized tissue microenvironments & spatial gene expression pattern. Along with the dataset includes tissue image information, expression profiles, & spot coordinate metadata which is necessary for spatial graph construction & spatial visualization. The Visium dataset gives a robust benchmark for evaluating the biological relevance & spatial coherence of proposed graph based analytical framework. The dataset is taken for this proposed work contains, shape of the data is: 4898 * 36601, total number of genes: 36601, & the cells/spots: 4898.

3.2 Data Preprocessing

For preprocess the spatial transcriptomics data Scanpy & Squidpy computational module is used to improve the data quality and ensure reliable downstream analysis. At first, some spots & genes with low quality & little expression levels were filtered to remove noise and reduce technical variability. Spots containing insufficient transcript counts or low gene expression were eliminated by using quality control.

By doing quality control, the gene expression value was normalized to decrease sequencing depth variations across tissue spots. To stabilize the expression variance & improve the distribution of transcriptomics features logarithmic transformation was applied. Then the High Variable Genes (HVGs) were identified to retain the most biologically informative features while reducing.

For further optimization the feature expression, Principal Component Analysis PCA was done for dimensional reduction.

3.3 Graph Construction

3.3.1 Graph Construction from Spatial Transcriptomics Data

Spatial transcriptomics data were represented as graph structures to incorporate spatial neighbourhood information into the analytical pipeline. Each tissue spot in the constructed spatial graph was represented by a graph node, with edges connecting neighbouring spots based on their spatial proximity relationships.

Neighbourhood connectivity in space was generated using Squidpy's coordinate-based adjacency mapping. The K-nearest neighbour (KNN)-based graph construction method was used to identify tissue spots that are spatially adjacent and preserve the local tissue topology. The resulting adjacency matrix encoded spatial relationships among neighbouring regions and enabled graph-based propagation of transcriptomic information across tissue structures.

The graph can be represented as:

$$G = (V, E)$$

Where, V is the set of spatial spots & E is the spatial neighbourhood relationship.

By using processed gene expression profiles & by enabling the integration of both spatial & molecular information within the graph representation, the node features were defined.

Where,

Spots (cells) → Nodes,

Spatial proximity → Edges,

Gene expression (PCA) → Node Features

The graph representation enabled the framework to combine transcriptomic similarity with spatial continuity, thus enhancing the identification of biologically significant tissue domains. The generated spatial graph acted as the primary input structure for Graph Neural Network-based representation learning.

3.3.2 Adjacency Matrix Representation

After graph construction the Spatial neighbour relationships encoded in an adjacency matrix it results a connection among tissue spots. The adjacency matrix gives a mathematical function of graph topology & serving as a primary structure which is using for GNN computations.

Adjacency matrix is defined as:

$$A \in R^{N \times N}$$

Where, N is the total number of spatial spots. Each matrix element indicates whether two spots are spatially connected:

$$A_{ij} = \begin{cases} 1, & if\ i\ and\ j\ are\ neighbouring\ spots \\ 0, & otherwise \end{cases}$$

This representation facilitates efficient encoding of local tissue organisation while preserving neighbourhood information necessary for graph-based message propagation. Because the adjacency matrix stores using compressed sparse row (CSR) format, along with significantly decreases requirements & improves computational efficiency during downstream graph learning operations.

3.3.3 Conversion to Edge Index Representation

Although adjacency metrics gives a acceptable mathematical representation but still modern graph deep learning framework needs alternative data structure for perfect computation. So adjacency matrix converted into an edge index representation suitable to PyTorch Geometric.

The edge index represents:

$$edge_index \in Z^{2 \times |E|}$$

Each column defines a connection between two graph nodes. This format explicitly stores the source and target node indices for every edge in the graph, enabling highly efficient message-passing operations during graph neural network training.

This conversion method from adjacency matrix to edge index representation it significantly reduces computational overload & performs a scalable processing of large spatial transcriptomics dataset.

3.3.4 Graph Data Object Generation

The dataset was prepared for Graph Neural Network training by combining node features and graph connectivity information into a unified graph data object. Gene expression profiles were transformed using PCA to derive node features, and graph connectivity was represented using an edge index structure.

The produced graph data object consisted of:

$$X \in R^{N \times d}$$

Represents node feature embeddings,

$$edge_index \in Z^{2 \times |E|}$$

It representing the graph connectivity relationships.

This unified representation allows for simultaneous access to transcriptomic features and spatial topology during model training, enabling the graph neural network to learn biologically meaningful embeddings that incorporate both molecular characteristics and spatial neighbourhood information.

3.4 Graph Neural Network Architecture

The proposed framework employed Graph Convolutional Networks (GCNs) to derive spatially informed transcriptomic representations from graph-structured tissue data, incorporating neighbourhood information across interconnected tissue spots to capture both local cellular interactions and global tissue organisation.

The graph neural network architecture was designed to process node-level transcriptomic embeddings in conjunction with spatial adjacency information. Feature information from neighbouring tissue spots was iteratively aggregated during graph convolution operations to generate biologically informative latent embeddings. The process improved the preservation of spatial continuity and enabled better characterization of the tissue domain.

Graph convolutional operation, mathematically represented as:

$$H^{(l+1)} = \sigma\left(\hat{D}^{-1/2}\hat{A}\hat{D}^{-1/2}H^{(l)}W^{(l)}\right)$$

Where, σ (activation function)

$\hat{A}$ represents adjacency matrix with self-connection,

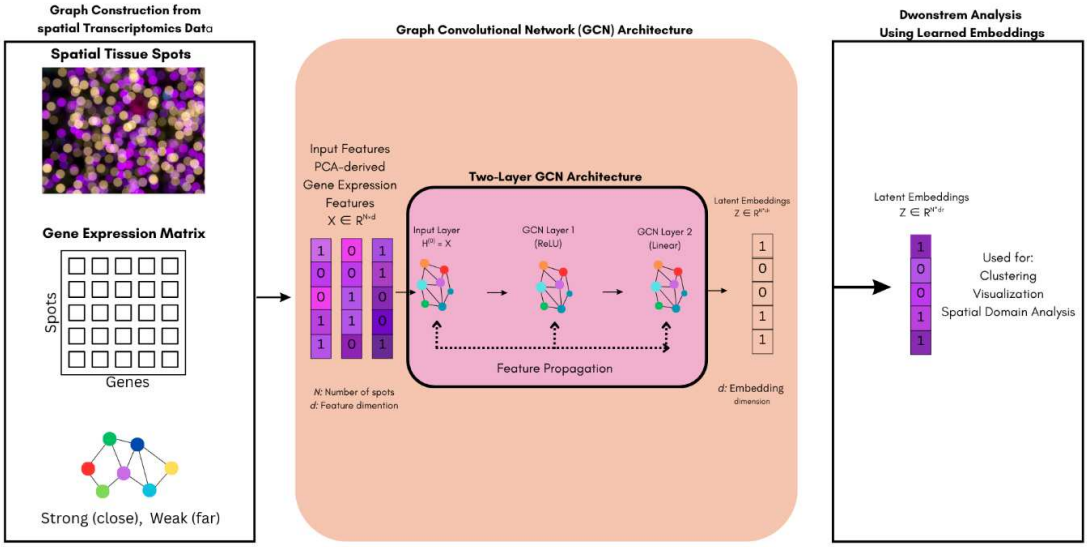
$\hat{D}$ represents degree matrix,

$H^{(l)}$ corresponds to node embeddings at layer l, and

$W^{(l)}$ represents trainable weight parameters.

The embeddings are generated by the GCN model, at where it is utilized for clustering, visualizing, & spatial domain analysis.

This architecture contains two convolutional layers. First layer uses Rectified Linear Unit (ReLU) that transforms PCA-derived gene expression features into a hidden representation space, whereas the second layer converts these representations into a lower dimensional latent embeddings space that captures biologically applicable spatial transcriptomics relationships.



[Fig 1: The diagram represents the Architecture of Graph Neural Network, which describes from graph construction to downstream analysis in details]

3.4.1 Feature Propagation and Embedding Generation

Graph Neural Networks rely on the fundamental mechanism of feature propagation to exchange information between adjacent nodes. In each graph convolution operation, a spatial location updates its representation by combining information from its neighbouring locations. This process preserves local neighbourhood structure by allowing biologically related tissue regions to influence one another, thereby enhancing representation quality.

The model learns latent representations that combine both transcriptomic features and spatial relationships as information passes through successive graph convolution layers. As a result, spots with comparable molecular properties and nearby spatial locations are positioned closer together within the learned representation space. The process effectively reduces noise while preserving biologically significant patterns within the tissue.

Here, each spatial spot was represented as a low-dimensional embedding vector as:

$$\boldsymbol{Z \in R^{N \times d'}}$$

Where, N is the total number of spatial spots & d′ represents the dimensionality of the latent embedding space. All the complex relationships were captured by these embeddings & serves as the foundation for downstream clustering, visualization & biological representations.

3.4.2 Spatial Domain Identification

The learned graph embeddings were then used for spatial domain identification. Biologically coherent tissue regions are represented by spatial domains, which are characterised by similar transcriptomic profiles and spatial organisation. In spatial transcriptomics, identifying these domains is a critical objective, as they frequently correspond to distinct cellular communities, tissue compartments, or functional microenvironments.

The Leiden community detection algorithm was used to perform clustering analysis on the generated latent embeddings. Leiden clustering was chosen for its robustness, scalability, and capacity to identify well-connected communities within graph-structured data. Clustering in the graph embedding space, as opposed to direct clustering on raw gene expression profiles, benefits from the integration of both molecular and spatial information, leading to more coherent and biologically meaningful tissue regions.

The cluster assignments were projected back onto the original tissue coordinates, resulting in spatial domain maps that visualise tissue organisation across the sample. The learned embeddings were further processed using Uniform Manifold Approximation and Projection (UMAP) to produce low-dimensional visual representations of cluster structure. The visualisations allowed for a qualitative evaluation of cluster separation, spatial continuity, and biological relevance, thereby offering insight into the underlying organisation of the tissue microenvironment.

3.5 Clustering and Spatial Domain Identification

Here, clustering analysis is done to perform identification of biologically coherent spatial domains within tissue sections. The Leiden clustering is utilized due to it's heftiness in detection of community structures in graph-based dataset. This type of clustering process group spatial spots with a similar spatial & transcriptomics character to distinct tissue domain.

For visualization of learned embeddings & clustering structures, the Uniform Manifold Approximation and Projection (UMAP) is used. The UMAP permits an effective visualization of transcriptomic similarity & cluster separation in 2D embedding space. Spatial cluster maps were generated to visualize the tissue architecture & domain organization from tissue coordinate.

To compare different clustering strategies & evaluate clustering performance, both K-Mean clustering & Leiden clustering is used. These two methods of visualizing tool is widely used in biological data for quality partitioning. As we know, the K-Mean clustering operates in Euclidean features by grouping all the observations according to centroid similarity. Whereas the Leiden clustering uses graph topology to identify highly connected clusters within learned graph structure.

3.5.1 K-Means Clustering

K-Means clustering was initially used as a baseline clustering method for the learned node embeddings. The algorithm partitions data points into k clusters by minimising the within-cluster sum

of squared distances between samples and their corresponding cluster centroids. Cluster centroids are updated iteratively until convergence, yielding groups of observations with similar feature characteristics.

This function is defined as:

$$J = \sum_{i=1}^{k} \sum_{x_j \in C_i} \| x_j - \mu_i \|^2$$

Where, C_i represents set of observations assigned to cluster *i,* & μ_i is the centroid of that particular cluster.

K-Means effectively identifies transcriptomically similar groups, but it does not take into account graph connectivity or spatial neighbourhood relationships explicitly. As a result, clusters produced by K-Means may display decreased spatial coherence, especially in intricate tissue environments where biological structure is heavily influenced by local spatial interactions. However, K-Means offers a useful benchmark for assessing the advantages of graph-aware clustering methods.

3.5.2 Leiden Clustering

Leiden clustering was used as the primary method to identify the spatial domains in the data, in order to better take advantage of the graph structure of spatial transcriptomics data. Leiden clustering is a graph-based algorithm that detects communities by identifying groups of nodes which are more densely connected to each other than to the rest of the graph. The Leiden algorithm directly incorporates graph topology and neighbourhood relationships, making it particularly well-suited for graph neural network embeddings.

The Leiden algorithm iteratively optimises network modularity by refining community assignments and improving partition quality. The algorithm produces well-connected and stable clusters through repeated local movement and community aggregation steps, thereby avoiding disconnected communities that can occur in other graph-partitioning methods.

The modularity optimization object can be represented as:

$$Q = \frac{1}{2m} \sum_{ij} \left(Aij - \frac{kikj}{2m} \right) \delta(ci, cj)$$

Here, A_{ij} is graph connectivity, k_i & k_j represents node degrees, m is the total number of graph edges, & $\delta\ (c_i, c_j)$ indicates the nodes belong to the same community.

3.5.3 Spatial Domain Mapping

The cluster assignments from clustering were then superimposed onto the original tissue coordinates to produce spatial domain maps. The process allowed for the direct visualisation of the spatial arrangement of identified clusters across the tissue section. The spatial domains that resulted revealed distinct tissue compartments and localized biological structures, offering insights into cellular heterogeneity and tissue architecture.

The Uniform Manifold Approximation and Projection (UMAP) was used to visualise the learned latent representations in a low-dimensional space, thereby further assessing cluster separation and embedding quality. This was achieved through the combined use of spatial domain mapping and

embedding visualisation, which facilitated both qualitative and quantitative evaluation of clustering performance, ultimately enabling comprehensive characterisation of spatially organised biological processes within the tissue.

3.6 Evaluation Metrics

To verify the framework's performance & accuracy, multiple evaluation metrics were used to measure the clustering quality & spatial coherence of the identified tissue domains. A single evaluation metrics is often insufficient to hold the biological relevance of clustering outcomes, so spatial transcriptomics analysis includes molecular & spatial information. That's why, this study utilizes both the Silhouette Score & Spatial Smoothness to evaluate the quality of learned embeddings & spatial domains.

3.6.1 Silhouette Score

This score is a wildly used clustering evaluation matric it measures how well a sample fits within their assigned cluster compared with neighbouring clusters. In the spatial transcriptomics, it estimates how effectively the learned embeddings separate distinct tissue regions on the basis their molecular characteristics.

The Silhouette Score for each tissue spot calculates the average distance to other spots within the same cluster and compares it to the average distance to the nearest neighbouring cluster.

It can be represented as:

$$s(i) = \frac{b(i) - a(i)}{max(a(i), b(i))}$$

Here in the representation, *a(i)* is the average distance between spot *i* & all other spots are from same cluster, while *b(i)* is the average distance between spot *i* & it's nearest cluster. The result score ranges from -1 to 1.

The Silhouette Score was calculated within the proposed framework using latent embeddings derived from a Graph Neural Network. The obtained score is 0.37 & it showed that the learned representations successfully captured transcriptomic variation and allowed for a meaningful separation of tissue domains. As a result, this metric functioned as a significant indicator of clustering quality and the effectiveness of the embedding.

3.6.2 Spatial Smoothness

This score evaluates the transcriptomic cluster separation. Biological tissues are inherently structured environments where neighbouring cells and tissue regions frequently display similar functional and molecular characteristics. A biologically meaningful clustering solution should therefore separate molecularly distinct regions while also preserving spatial continuity.

As a complementary evaluation metric, the Spatial smoothness is used to solve this problem. It measures the consistency of cluster assignments among neighbouring tissue spots. The metric assesses whether adjacent areas within the spatial graph are part of the same cluster, thus indicating the biological coherence of the resulting tissue domains.

The Spatial Smoothness can be expressed as:

$$\textbf{\textit{Spatial Smoothness}} = \frac{\sum(i,j)\in EI(ci=cj)}{|E|}$$

Here, E is the set of graph edges, c_i & c_j is the cluster assignments of neighbouring nodes, I is an indicator function that returns 1 when both nodes belong from same cluster or otherwise it returns 0.

This study Spatial GNN framework achieved a Spatial Smoothness score of 0.93, which indicates a strong preservation of neighbourhood relationships & highly coherent spatial domain. This spatial smoothness score demonstrates that the integration of graph-based learning & spatial regularization effectively captured tissue architecture.

3.7 Visualization Framework

Interpreting spatial transcriptomics data typically involves the concurrent evaluation of several analytical outputs such as spatial domain maps, gene expression patterns, clustering results, and embedding visualizations. An integrated visualization framework was developed to facilitate efficient exploration and interpretation of these complex datasets as part of the proposed Spatial GNN Analyzer platform. The framework was designed to bridge the gap between computational analysis and biological interpretation by offering an interactive environment for exploring spatially resolved transcriptomic data.

The developed framework, allows researchers to manually investigate tissue architecture along with differentiate the tissue clusters, identify the cluster distributions & spatially differing molecular patterns which is generated by analytical workflow.

3.7.1 Interactive Dashboard

A web-based interactive dashboard was created to offer centralised access to all stages of the analytical workflow. The dashboard functions as the main interface where users can upload datasets, start analysis pipelines, track computational progress, and examine the results that have been generated. The platform was designed to accommodate users with varying levels of computational expertise, thereby promoting broader accessibility to advanced spatial transcriptomics analysis.

The developed dashboard integrates multiple modules, which includes preprocessing summaries, clustering outputs, graph-based embedding visualizations, spatial domain maps, & evaluation metrics. Interactive controls are integrated on navigate bar, like guide in which user can learn how to use the web tool in details, along with upload, dashboard, visual & download, it helps to directly jump to that particular location for the analytical view. User can examine the individual clusters, & explore gene expression pattern within the specific tissue regions.

The dashboard facilitates a smooth transition from raw spatial transcriptomics data to biologically significant findings by integrating computational analysis with biological exploration. An interactive environment boosts user interaction and facilitates hypothesis development in spatial omics research.

3.7.2 Multi-Panel Visualization

A multi-panel visualization strategy is done only for provide a comprehensive view of tissue organization. This analysis generates a multiple complementary outputs & interpretations. The multi-

panel framework provides a circumstantial visualization of different analytical themes in a single workspace.

The visualization part incorporates spatial cluster maps, UMAP embedding plots, gene expression heatmaps, marker gene visualizations, & tissue morphology overlays. All the maps are present side by side, so that researchers can directly compare transcriptomic patterns with spatial domain organization & evaluate the biological relevance of identified clusters.

Interactive visualisation capabilities also enable users to explore individual spatial domains, examine cluster-specific gene expression patterns, and investigate tissue heterogeneity at various scales. The inclusion of modules for both two-dimensional and three-dimensional visualisation further enhances exploratory analysis by allowing for a detailed examination of spatial structures from different viewpoints. Comprehensive visualisation is particularly valuable for investigating complex biological systems, including tumour microenvironments and heterogeneous tissue landscapes.

3.7.3 Automated Report Generation

An automated report generation system was incorporated into the framework to improve reproducibility and facilitate the dissemination of analytical results. Upon completion of the analytical workflow, the platform automatically generates a structured report comprising key outputs, which includes dataset summaries, preprocessing statistics, clustering results, evaluation metrics, all other type of necessary results and visualisation outputs.

Generated reports offer a standardised depiction of analytical results, thereby decreasing the necessity for manual documentation and lessening the likelihood of reporting discrepancies. The reporting system simplifies result interpretation and facilitates efficient communication of findings among researchers and collaborators by consolidating multiple analytical outputs into a single document.

By preserving critical information regarding analytical procedures & model outputs, the automated report generation supports reproducible research. The main & advance thing is the generated report can be downloaded, archived, shared or integrated into downstream scientific workflows, thereby enhancing transparency and facilitating future validation of experimental results. The proposed framework offers a comprehensive environment for spatial transcriptomics analysis and interpretation through the integration of visualization, quantitative evaluation, and automated documentation.

IV. System Architecture

The proposed framework is designed as an end-to-end computational ecosystem that integrates spatial transcriptomics preprocessing, graph-based representation learning, clustering, visualization, evaluation, & automated reporting in a unified analytical environment. The architecture was developed to cover a gap between advanced graph learning methodologies & practical biological interpretation by serving a seamless workflow from raw spatial transcriptomics data to a biological insight.

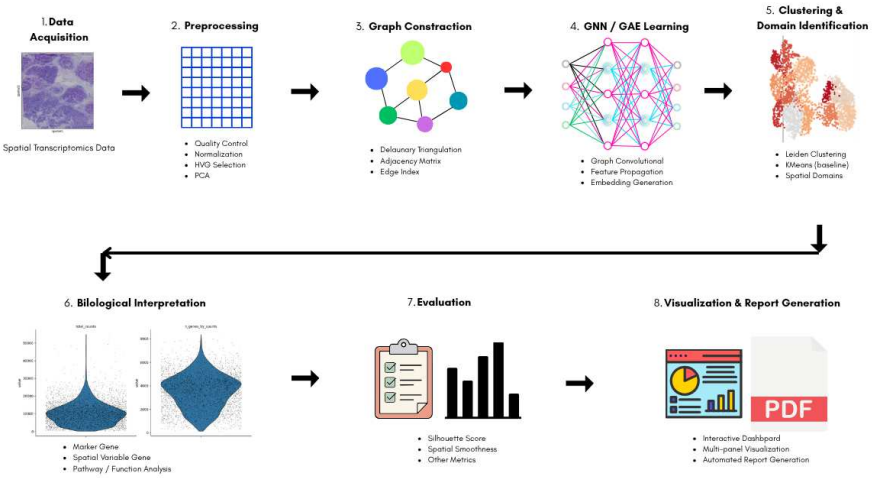
This workflow follows a modular architecture in which each analytical components works as independent interconnected module. This type of design improves scalability, maintainability, & reproducibility & enables efficient handling of large-scale spatial transcriptomics datasets. Here the overall system contains of four main layers: data processing layer, graph learning layer, analytical layer, and visualization layer. With the together of these four components, it forms a comprehensive platform for spatial domain identification & tissue architecture characterization.

4.1 Workflow Architecture

The Spatial GNN follows a sequential analytical pipeline that translates raw spatial transcriptomics data into an interpretable biological knowledge. The workflow starts with dataset acquisition & preprocessing, spatial graph construction & graph neural network-based representation learning. After that, the generated embeddings are subsequently utilized for clustering, spatial domain identification & biological interpretation.

The analytical workflow incorporates a number of preprocessing stages, that includes quality control, normalization, highly variable gene selection, dimensionality reduction, graph generation, graph neural network training, clustering, marker gene analysis, evaluation, and visualization. The framework ensures consistent analytical processing across different datasets by integrating these stages into a unified workflow and minimising manual intervention.

This design enables reproducible spatial transcriptomics analysis by facilitating efficient exploration of tissue architecture, spatial heterogeneity, and cellular organisation. The modular design of the workflow enables individual components to be updated or replaced without disrupting the entire analytical process.



[Fig 2: The image demonstrates how the workflow architecture works, Transcriptomics data enters as Data acquisition then it follows preprocessing, graph construction, GNN learning, clustering, interpretation, evaluation and at last visualization & report generation.]

4.2 Frontend–Backend Architecture

The system maintains a client-server architecture, by consisting a web-based frontend & a computational backend. This separation makes the app flexibility & computationally intensive. The frontend was developed by using react, which provides an intuitive user interface in which a researcher can upload datasets, figure out the parameters used in dataset, explore visualize outputs, & download generated reports. The interface was designed in a simple way with complex spatial transcriptomics workflows while maintaining easy accessibility for users with limited computational expertise.

The backend layer functions as the core computational component of the framework and is accountable for processing data, building graphs, training graph neural networks, performing clustering, evaluating results, and producing reports. The seamless transfer of data and analytical results between frontend and backend modules enables users to interact with complex graph-based analyses through a unified and user-friendly environment.

Separating frontend and backend components enhance scalability by enabling the independent management of computational resources from visualisation and user interaction services. The architecture enables the integration of further analytical modules and more advanced machine learning algorithms in the future.

4.3 Analytical Pipeline Architecture

This analytical pipeline represents as the core computational component of Spatial GNN Analyzer. The framework is mainly design to represent the spatial transcriptomic data in to a graph-based embedding representation. At first the raw Visium dataset undergo preprocessing step which includes quality control, normalization, highly variable gene selection, and dimensionality reduction. Then Graph Neural Networks performs neighbour-aware feature learning to generate embeddings that capture both transcriptomic & spatial information.

The embeddings are utilized for downstream clustering & spatial domain identification. Further support for a biological interpretation is provided through marker gene analysis, detection of genes with spatial variability, and quantitative evaluation using clustering and spatial coherence metrics. The outputs are then sent to the visualization and reporting modules for interactive analysis and result sharing.

4.4 Data Flow Architecture

Data flow in the proposed framework follows a structured way from raw input to final outputs. The spatial transcriptomics dataset initially contains gene expression matrices & spatial coordinates, which are uploaded in the system. These data are processed through quality control & normalization modules before transforming into graph-structured representations which is suitable for graph neural network learning.

After graph construction, node features & graph connectivity information are transferred to graph learning modules. Here, latent embeddings are generated through graph convolutional operations. The learned embeddings are subsequently transferred to clustering & evaluation for domain identification & performance assessment.

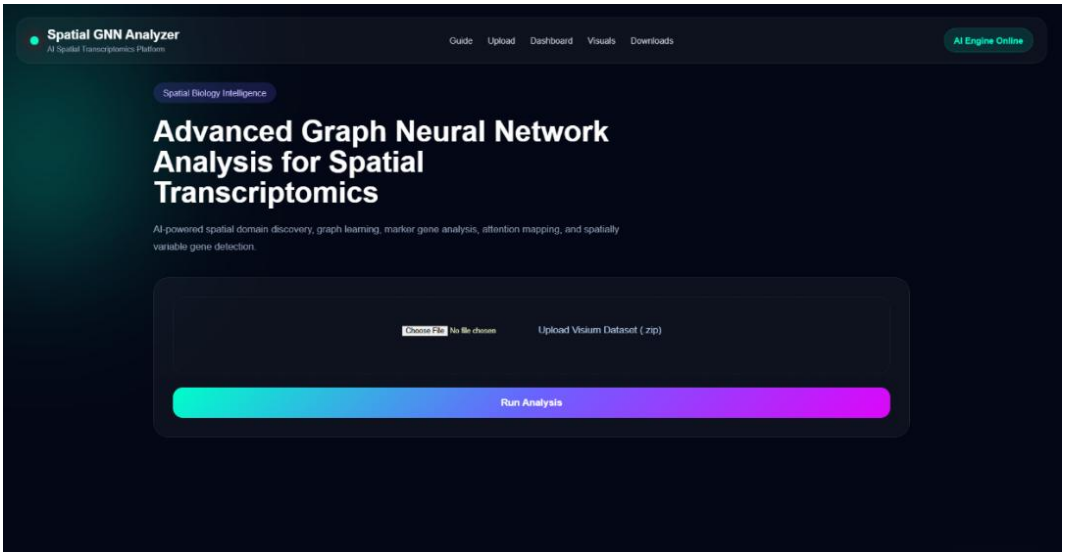
The all the resulted things like, cluster assignments, embedding representations, evaluation metrics, and biological annotations are forwarded to visualization framework. After that the final analytical outputs are integrated into automated reports & exportable result files. This systematics data flow occurs consistency across analytical stages while maintaining reproducibility throughput workflow.

V. Platform and Visualization Module

The main contribution of this proposed framework Spatial GNN Analyzer lies in development of an integrated platform that combines an advanced graph based spatial transcriptomic analysis with an interactive visualization. While, existing other platforms are providing a limited support for intuitive result exploration & user interaction. To address this limitation, I developed this visualization module to enable researchers to efficiently analyse, interpret & communicate spatial transcriptomics findings through a unified platform.

5.1 React-Based Visualization Workspace

The visualization environment is done as a web-based analytical workspace by using React-based frontend architecture. The workspace works as a centralizes interface in which users can upload spatial transcriptomics datasets, execute the analytical workflow in the background, monitoring processing status & interactive exploration of results.



[Fig 3: Here, the image shows the UI visualization of the framework, user can upload his .zip file here to do analysis the spatial transcriptomics data]

If the user gets confuse, about how to use the platform, how to arrange the files for feed to the workspace, then the user can learn from the guide button from the navigation bar, that contains how a user can download the dataset, how to make feedable .zip file and all the details along with an tutorial video also their so that user can learn easily.

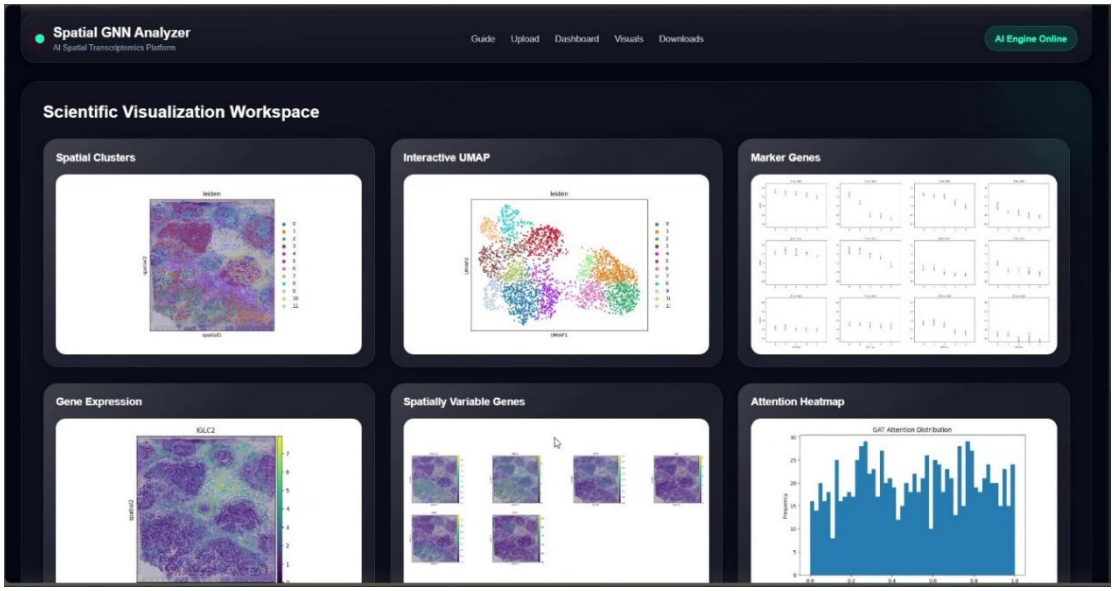
The React-based workspace supports a dynamic rending of analytical outputs & provides a responsive communication in a large-scale spatial transcriptomics dataset. The framework significantly improved usability & accessibility as compared with command-line based bioinformatics workflow.

Furthermore, the workspace allows a real-time interactive analytical result. It allows to investigate a specific tissue region, compare outputs, & evaluate the patterns of biology without requiring another software tool. This framework enhances the user experience & along with supports reproducibility spatial transcriptomics analysis.

5.2 Interactive Visualization and Analytical Exploration

To explore spatial transcriptomics data, the platform incorporates an interactive visualization module. This allows users to dynamically examine clustering outcomes, graph embeddings, evaluation metric & tissue morphology while maintain direct linkage between transcriptomics & spatial information.

The attractive plots were integrated to visualize UMAP embeddings, clustering distributions, spatial maps, graph-based representations which are generated by Graph Neural Networks. The users can delve into individual clusters, inspect neighbourhood relationships, and evaluate spatial organisation through responsive visual interfaces. Interacting with analytical outputs facilitates a more in-depth examination of tissue heterogeneity and enhances the interpretation of graph-derived visualisations.



[Fig 4: The image shows different type of interactive visualization, plot & graphs which are generated by this framework. By using this, a researcher can easily find all information from this workspace.]

In addition, the workspace supports both 2-D & 3-D visualization modes of graphs with multi scale exploration of tissue architecture. These capabilities offer a more detailed comprehension of spatial relationships and enable the identification of biologically significant structures within intricate tissue environments.

5.3 Downloadable Reports and Result Dissemination

The platform includes an automated framework for generating and exporting analytical reports to enhance reproducibility and facilitate the dissemination of findings. Upon completion of the analysis, the system automatically compiles key outputs into structured reports which include preprocessing summaries, clustering statistics, evaluation metrics, spatial domain visualisations, and gene expression analyses.

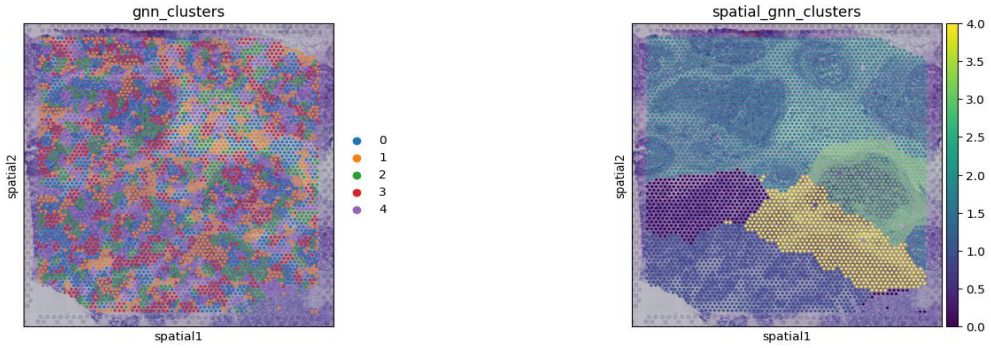
Researchers can download reports in portable formats for purposes such as archiving, collaboration, publication support, and downstream analyses. The platform enables the export of processed datasets, cluster assignments, marker gene results, and visualisation outputs in addition to report generation.

A complete research platform is formed by this reporting and export functionality, as it turns the framework into a comprehensive data analysis and scientific communication tool. The integration of visualization, automated documentation, and result dissemination improves transparency, reproducibility, and practical usability in spatial transcriptomics research.

VI. Results and Discussion

6.1 Spatial Domain Identification Results

The proposed Spatial GNN Analyzer framework successfully identified biologically significant spatial domains from the Visium Human Breast Cancer dataset, Visium Human Breast Cancer filtered feature bc matrix.h5 using graph-based representation learning and Leiden community detection. The tissue was partitioned into five distinct spatial domains by Leiden clustering, following graph construction and embedding generation, with each domain representing a heterogeneous tissue region of unique transcriptomic characteristics. The identified clusters showed a clear spatial arrangement and maintained neighbourhood continuity throughout the tissue structure.



[Fig 5: The image shows, two different graphs, one is GNN Cluster & another is spatial GNN Cluster]

The clustering map shows that GNN effectively captured local tissue relationships. The distinct clusters were organised in specific regions within the tissue section, not randomly dispersed across it, which indicates that the learned embeddings had managed to preserve biologically important spatial arrangement. Understanding tissue organisation and identifying functionally distinct microenvironments requires spatially coherent segmentation.The analysis of cluster composition showed variability in domain size, which reflects the inherent heterogeneity of the tissue.

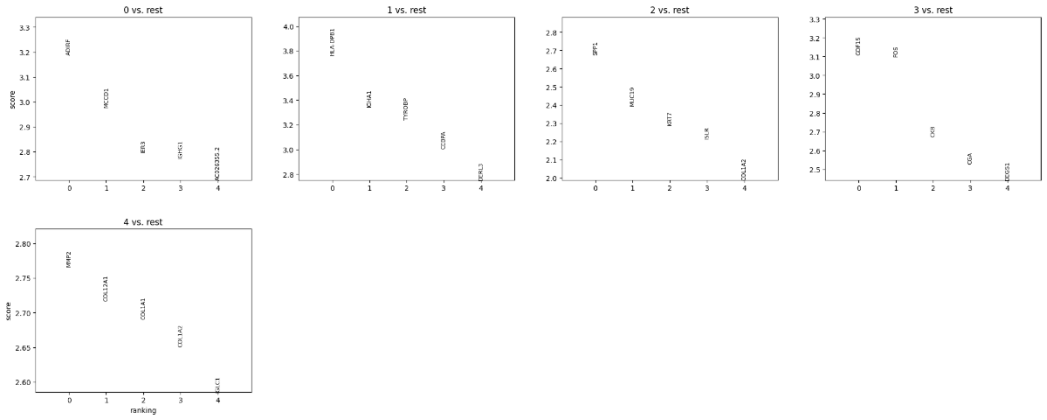
6.2 UMAP Embedding Analysis

The quality of the learned graph embeddings was investigated using Uniform Manifold Approximation and Projection (UMAP) for low-dimensional visualisation. The UMAP representation showed distinct clusters, confirming that the Graph Neural Network created useful latent representations that could differentiate between distinct tissue regions.

The UMAP visualisation also demonstrated the proposed framework's capacity to preserve both local and global data structures. The formation of compact cluster groups indicates effective feature extraction and dimensionality reduction, and the relative positioning of clusters implies biologically significant relationships among tissue regions. The results show that graph-based representation learning is effective for spatial transcriptomics analysis.

6. .3 Marker Gene Characterization

The Leiden clustering algorithm divided the tissue into five distinct clusters, each displaying unique transcriptional profiles. Cluster 0 was distinguished by increased expression of ADIRF, KCCD1, IER3, and IGHG1. The presence of IGHG1, a gene related to immunoglobulins, implies potential immune-related activity within this region. The genes HLA-DPB1, IGHA1, TYROBP, and CEBPA, which are typically involved in immune regulation, antigen presentation, and inflammatory responses, were strongly expressed in Cluster 1, suggesting the presence of immune-enriched tissue compartments.



[Fig 6: The image demonstrates, identification of marker genes of the Visium file of human breast cancer dataset]

Cluster 2 was characterised by increased expression of SPP1, MUC19, KRT7, ISLR, and COL1A2. These genes are associated with epithelial organisation, extracellular matrix interactions, and tissue

remodelling, suggesting a biologically distinct structural tissue region. Cluster 3 demonstrated an enrichment of GDF15, FOS, CKB, and CGA, suggesting active transcriptional programmes linked to cellular signalling and stress-response pathways. Cluster 4 showed a strong expression of genes associated with the extracellular matrix, such as MMP2, COL12A1, COL1A1, and COL1A2, indicating the presence of stromal or connective tissue-related areas within the breast cancer microenvironment.

The framework's ability to identify distinct marker gene profiles across all five clusters shows that it effectively captured biologically significant transcriptomic variation. The results show that the learned embeddings successfully maintained tissue heterogeneity and allowed for the identification of transcriptionally distinct spatial domains.

6.4 Spatially Variable Gene Analysis

Differential expression analysis of cluster-specific marker genes showed significant spatial heterogeneity within breast cancer tissue. The tissue microenvironment contains region-dependent transcriptional programs, as indicated by several genes exhibiting localized enrichment patterns across specific spatial domains.

Genes like SPP1, MMP2, COL1A1, and COL12A1 displayed robust cluster-specific expression patterns linked to specific tissue areas. The enrichment of extracellular matrix-related genes in Cluster 4 implies localized stromal remodelling and structural tissue organisation. The expression of SPP1 within Cluster 2 suggests spatially restricted biological activity that could be linked to tissue remodelling and tumour-associated microenvironments.

Genes related to immunity, such as HLA-DPB1, IGHA1, TYROBP, and IGHG1, showed enrichment in Clusters 0 and 1, indicating the presence of distinct immune-associated regions in space. The localised expression of these genes implies diverse immune cell infiltration patterns across the tissue section, which is a characteristic feature of the spatial organisation of breast cancer.

The observed spatial variation in gene expression shows that the proposed Spatial GNN framework effectively maintained biologically relevant spatial information throughout graph-based learning. The resulting spatial domains were not only distinct at the transcriptomic level but also reflected meaningful biological compartmentalisation within the tissue microenvironment. The discovery of these patterns provides additional evidence for the effectiveness of graph-based representation learning in identifying spatially organised molecular patterns in breast cancer tissues.

VII. Comparative Analysis

7.1 Comparison with Existing Spatial Transcriptomics Frameworks

In last recent years, several graph-based & deep learning approaches for transcriptomics analysis have developed, including SpaGCN, STAGATE, GraphST, and DeepST. All these methods were significantly improved the identification of spatial domains by integration of transcriptomic information with spatial context. Most of the existing methods mainly focus on algorithmic

performance & some provides a limited support for visualization, automated reporting, and user-oriented analytical workflows.

A qualitative comparison was carried out against representative state-of-the-art methods to position the proposed Spatial GNN Analyzer within the current spatial transcriptomics tools landscape. The comparison centred on graph learning capability, spatial domain identification, biological interpretability, visualisation support, and platform integration.

Feature	SpaGCN	STAGATE	GraphST	DeepST	Spatial GNN Analyzer
Spatial Information Integration	Yes	Yes	Yes	Yes	Yes
Graph Neural Network-Based Learning	Yes	Yes	Yes	Partial	Yes
Spatial Domain Identification	Yes	Yes	Yes	Yes	Yes
Marker Gene Analysis	Limited	Limited	Yes	Yes	Yes
Spatially Variable Gene Analysis	Limited	Limited	Partial	Partial	Yes
Interactive Visualization	No	No	No	No	Yes
Web-Based Analytical Platform	No	No	No	No	Yes
Automated Report Generation	No	No	No	No	Yes
Downloadable Results	Limited	Limited	Limited	Limited	Yes
User-Friendly Interface	No	No	No	No	Yes
End-to-End Analysis Pipeline	Partial	Partial	Partial	Partial	Yes

[**Table 1**: Comparison of Spatial Transcriptomics Frameworks]

7.2 Current Limitations

Along with its advantages, several limitations are there. This framework is currently evaluating the Visium spatial transcriptomics datasets and requires validation across additional spatial omics platforms and tissue types. Additionally, while graph neural networks can accurately represent local spatial connections, combining multimodal data including histological image characteristics may enhance biological understanding.

The computational complexity associated with graph construction and graph neural network training may also rise significantly for extremely large datasets. Implementing future scalable graph learning architectures and optimization strategies could address these challenges.

VIII. Novelty and Key Contributions

8.1 Novelty of the Proposed Framework

The main novelty of this research, this framework incorporates a spatial regularization mechanism that encourages neighbouring tissue spots to learn same latent representation, and also preserve a local tissue architecture.

The integration of both spatial information and gene expression similarity is an important contribution during graph construction. Results from experiments showed that combining these two approaches enhances biological understanding and makes it possible to identify accurate similarities between different types of tissue at the transcriptional level. The research suggests that spatial information on its own is not enough for a complete characterisation of tissue, and that combining multiple types of data provides a more accurate picture of tissue structure.

8.2 Platform Innovation

This work introduces an end-to-end spatial transcriptomics analytical platform that integrates data preprocessing, graph construction, graph neural network learning, clustering, marker gene analysis, spatially variable gene detection, visualisation, and automated reporting within a unified environment.

The platform includes interactive dashboards, multi-panel visualisations, gene expression exploration, downloadable analytical outputs, and automatic report creation. The proposed framework's capabilities transform it from a standalone computational method into a comprehensive research platform that supports both computational analysis and biological interpretation.

8.3 Key Contributions

The main findings of this research are outlined below.

- Learning spatial representations for spatial transcriptomics using a spatially aware Graph Autoencoder.
- Modelling spatial interaction strength through distance-based edge weighting.
- Characterisation of tissue is enhanced through the integration of spatial proximity and transcriptomic similarity.
- Identification of spatial domains in the graph-based spatial domain is implemented using Graph Neural Networks and Leiden clustering.
- The inclusion of marker gene detection and spatially variable gene analysis for biological interpretation purposes.
- Development of a comprehensive web-based analytical platform for spatial transcriptomics research.
- The integration of interactive visualisation, gene expression mapping, and multi-panel analytical dashboards.
- Automated report generation and downloadable analytical outputs are implemented.
- The analysis achieved strong spatial coherence with a Spatial Smoothness score of 0.93 and biologically relevant clustering with a Silhouette Score of 0.37.

IX. Future Work and Limitations of the Study

9.1 Limitations

Although this framework provides a lot of performance, still it has some limitations, like every good thing are not 100% perfect as our prospective. Like this, the proposed GNN Analyzer have several limitations.

First, the framework only works or evaluates by using 10x Genomics Visium spatial transcriptomics datasets. Although the framework does a strong spatial coherence & biologically meaningful clustering, additional validation across diverse spatial omics platforms.

Secondly, the framework does not currently work on histopathological image features or multimodal biological information, while it relies on spatial coordinates & transcriptomic similarity.

Third, Here the model is designed for unsupervised spatial domain discovery so it does not use prior biological knowledge or annotated labels during depends on downstream marker gene analysis & expert validation.

The computational complexity of constructing graphs and training graph neural networks may escalate for very large spatial transcriptomics datasets comprising hundreds of thousands of spatial locations. Large-scale applications may necessitate the use of optimisation and scalable graph learning strategies.

Despite achieving high spatial smoothness and meaningful clustering performance, the framework's comprehensive benchmarking against multiple state-of-the-art methods across standardised datasets was beyond the scope of the current study and remains an important area for future research.

9.2 Future Work

The proposed GNN Analyzer demonstrated effective spatial domain identification & biological interpretation & several things for future development. Future work will focus on integrating of multimodal spatial omics data, including histopathology images, spatial proteomics, and single-cell transcriptomics, to provide a more comprehensive understanding of tissue organization.

Along with the framework can be extended by advanced architecture such as Graph Attention Networks (GATs), self-supervised learning approaches to improve it. Along with more emerging technologies like Slide-seq, MERFISH, SeqFISH, and Stereo-seq can increase versatility of the platform.

As a platform point of view, it can be upgraded as cloud-based deployment, real-time analysis, AI-assisted biological interpretation & advanced visualization modules. Along with validation across diverse tissue type & disease conditions will be conducted to evaluate.

The proposed framework will undergo large-scale validation across various tissue types, disease conditions, and publicly available benchmark datasets to further evaluate its robustness and generalisability. Developments in the making will enable the Spatial GNN Analyzer to function as a comprehensive and scalable platform for next-generation spatial omics research, biomarker discovery, tissue heterogeneity analysis, and various precision medicine applications.

X. Conclusion

This study introduces the Spatial GNN Analyzer, a unified AI-powered framework for spatial transcriptomics analysis, combining graph-based representation learning, spatial domain identification, biological interpretation, and interactive visualisation. The framework effectively captures both local cellular interactions and global tissue architecture by combining transcriptomic information with spatial neighbourhood relationships modelled as graph structures.

The proposed approach integrates spatially aware graph learning, distance-based edge weighting, and graph-based clustering to discover biologically significant tissue domains. The experimental evaluation of 10x Genomics Visium breast cancer datasets showed strong preservation of spatial organisation, achieving a Spatial Smoothness score of 0.93 and a Silhouette Score of 0.37. The biological relevance of the identified spatial domains was confirmed by marker gene analysis, spatially variable gene detection, and gene expression visualisation.

The framework offers an end-to-end analytical environment which encompasses interactive visualisation, multi-panel investigation, automated report creation, and downloadable results. This integration transforms a proposed standalone computational model into a practical research platform for spatial omics investigations.

The Spatial GNN Analyzer showcases the potential of integrating Graph Neural Networks, spatial transcriptomics, and interactive analytics to improve the comprehension of tissue heterogeneity and spatially resolved biological processes. A proposed framework offers a scalable and interpretable solution for next-generation spatial biology research, which has potential applications in tumour microenvironment analysis, biomarker discovery, tissue architecture characterisation, and precision medicine.

XI. References

[1] P. L. S. F. V. S. e. a. Ståhl, “Visualization and analysis of gene expression in tissue sections by spatial transcriptomics,” *Science,* no. 353(6294), pp. 78-82, 2016.

[2] A. E. &. I. S. Moor, “Spatial transcriptomics: paving the way for tissue-level systems biology,” *Current Opinion in Biotechnology,* no. 46, p. 126–133, 2017.

[3] S. E. G. S. F. e. a. Vickovic, “High-definition spatial transcriptomics for in situ tissue profiling,” *Nature Methods,* no. 16, p. 987–990, 2019.

[4] J. L. X. C. K. e. a. Hu, “SpaGCN: Integrating gene expression, spatial location and histology to identify spatial domains by graph convolutional network,” *Nature Methods,* no. 18, pp. 1342-1351, 2021.

[5] K. Z. S. e. a. Dong, “Deciphering spatial domains from spatially resolved transcriptomics with an adaptive graph attention auto-encoder,” *Nature Communications,* no. 13, p. 1739, 2022.

[6] V. Marx, “Method of the Year: spatially resolved transcriptomics,” *Nature Methods,* no. 18, pp. 9-14, 2021.

[7] L. &. P. L. Moses, “Museum of spatial transcriptomics,” *Nature Methods,* no. 19, pp. 534-546, 2022.

[8] L. H. J. &. M. J. McInnes, “UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction,” *arXiv preprint,* p. 1802.03426., 2018.

[9] E. S. M. R. R. X. e. a. Zhao, “Spatial transcriptomics at subspot resolution with BayesSpace,” *Nature Biotechnology,* no. 39, p. 1375–1384, 2021.

[10] J. M. A. C. Y. e. a. Wang, “DeepST: identifying spatial domains in spatial transcriptomics by deep learning,” *Nucleic Acids Research,* no. 50(22), p. e131, 2022.

[11] B. L. Y. &. Z. L. Liu, “nalysis and visualization of spatial transcriptomic data,” *Briefings in Bioinformatics,* 2021.

[12] D. E. N. Subhasankar Khilar, “Explainable Graph Neural Networks (XGNNs) For Protein-Ligand Interaction Interpretation,” *IJRTI,* vol. 11, no. 1, 2026.

[13] D. E. N. Subhasankar Khilar1, “Graph Neural Networks (GNNs) for Protein-Ligand Interaction Prediction,” *bioRxiv,* 2026.

[14] Y. L. K. X. J. e. a. Long, “Spatially informed clustering, integration, and deconvolution of spatial transcriptomics with GraphST,” *Nature Communications,* no. 14, p. 1155, 2023.