

A Study on Graph-Theoretic-Based Classification to Explore Geometric Consistency in Tissue Architecture of Colon Histopathology

Dr. M. Siva Parvathi¹, Prof. T. Sudha², Dr. Shaheen Begum³, Dr. N.V. Muttu Lakshmi⁴, Mrs. T. Sukeerthi⁵, G. Keerthi⁶, P.V. Anand Krishna⁷

¹Assistant Professor, Dept. of Applied Mathematics, Sri Padmavati Mahila Visvavidyalayam, Tirupati, Andhra Pradesh, India; ²Professor, Dept. of Computer Science, Sri Padmavati Mahila Visvavidyalayam, Tirupati, Andhra Pradesh, India; ³Associate Professor, Institute of Pharmaceutical Technology, Sri Padmavati Mahila Visvavidyalayam, Tirupati, Andhra Pradesh, India; ⁴Assistant Professor, Dept. of Computer Science, Sri Padmavati Mahila Visvavidyalayam, Tirupati, Andhra Pradesh, India; ⁵Academic Consultant, Dept. of Statistics, Sri Padmavati Mahila Visvavidyalayam, Tirupati, Andhra Pradesh, India; ⁶Research Assistant, Dept. of Applied Mathematics, Sri Padmavati Mahila Visvavidyalayam, Tirupati, Andhra Pradesh, India; ⁷Student, Government Polytechnic, Palaripattu, Andhra Pradesh, India

ABSTRACT

This pilot study investigates conventional characterization of spatial organization in colon histopathology using graph theory which represents complex tissue images as networks of interconnected elements. It explores the potential of graph theory-based transformations in histopathological images of colon tissue to support medical image interpretation. Using a dataset of 40 images (20 benign and 20 adenocarcinoma), the research adopts five different graph modeling approaches such as proximity graph representation, Voronoi diagram, Delaunay triangulation, Minimum Spanning Tree (MST), and K-Nearest Neighbor (KNN) graph. These methods convert visual information into a spatially measurable network to improve on manual microscopic inspection, where superimposed structures often obscure diagnostic clarity. In the results, the Graph Convolution Network (GCN) shows 90% test accuracy, compared with Delaunay Triangulation and Minimum Spanning Tree models, which reveal perfect identification, thereby demonstrating that graph-theoretic thought captures related features effectively in tissue organisation.

ARTICLE HISTORY

Received February 13, 2026

Accepted April 15, 2026

KEYWORDS

Voronoi Diagram,
Delaunay Triangulation,
Minimum Spanning Tree
(MST),
K-Nearest Neighbor (KNN)

1. Introduction

Cancer is an important health problem worldwide and one of the most common causes of mortality globally. Among cancers, colorectal cancer, which involves colon and rectal malignancies, is the third most common cancer diagnosed and the second leading cause of cancer deaths worldwide. According to the World Health Organization, there were approximately 1.9 million new cases and 930,000 deaths due to colorectal cancer in 2020 (Organization, 2020). Many diseases progress slowly from benign polyps, and early stages are often asymptomatic. Thus, for any such malignancy, the key to good outcomes and survival is early detection and accurate diagnosis. While computational methods are being explored, the histopathological examination of tissue morphology remains the standard method for cancer diagnosis and grading. By microscopic examination of tissue architecture, glandular structures and morphological changes, pathologists decide whether a sample is benign or malignant (Gurcan et al., 2009). However, manual histopathological analysis is time-consuming and laborious, and subjective variability is somewhat inevitable. Given the large volume of slides pathologists must review, diagnostic reliability can be compromised by visual fatigue and interobserver differences.

In colon histopathology, it is particularly difficult to distinguish benign tissue from adenocarcinoma by naked-eye inspection alone. Early malignancies may not show obvious visual differences from normal tissues. Sometimes, gland patterns, connective tissue, and background stroma can be very similar. The very challenge of spotting these slight anomalies underscores the importance of quantitative image analysis; it can expose hidden spatial and structural features that are not easily seen by the naked eye (De Matos et al., 2021; Moscalu et al., 2023). The drawback is that manual histopathological evaluation is very tedious, takes a long time to complete, and is inherently subjective. The sheer number

of slides that pathologists are required to review results in visual fatigue and inter-observer differences, which can hurt the reliability of the diagnosis. The very limitations already mentioned have pushed research into computational and mathematical approaches that can provide diagnostics with objective, reproducible, and quantitative support for decision-making.

Graph theory provides a mathematical framework for depicting the spatial relations and structural organisation of images. In the graph model, each region or feature represents a node, and the edges indicate the geometric or spatial relationships between them. Such representations are being extensively utilized in histopathology, where, through analyzing the different spatial features, such as connectivity, adjacency, and geometric anomalies, tissue organization might be investigated (Demir & Yener, 2005; Sharma et al., 2015). Voronoi diagrams, Delaunay triangulation, Minimum Spanning Tree (MST), and K-Nearest Neighbor (KNN) are some of the techniques that have been widely applied in the measurement of tissue structure and the detection of cancer-related conditions (Doyle et al., 2007; JS & VS, 2015). Recent studies now provide evidence of the effectiveness of graph-theoretic modelling in analysing histopathological images. In this regard, Sharma et al. (2015) have conducted a review of graph-based methods for digital pathology stating that geometric modeling can help to make the subtle topological differences between normal and tumor tissues clearer. Similarly, Gurcan et al. (2009) reported that the incorporation of mathematical frameworks into histopathology could lead to better diagnostic consistency and reproducibility. Furthermore, the recent advancements in computational pathology involving GNNs, (Lippoldt, 2020) and hybrid graph-based machine learning models (Xiang & Wu, 2021) have brought a remarkable improvement in the accuracy of automated tissue classification and grading. The five modeling approaches used by the authors included proximity graph representation, Voronoi diagram, Delaunay triangulation, MST, and KNN graph modeling. From the selected 40 images, five graph models were generated for each image, giving a total of 200 graph-based images for training, testing and analysis.

2. Methodology

2.1 Dataset Description

The dataset used for this work was sourced from the publicly available repository “Lung and Colon Cancer Histopathological Images” on Kaggle (Borkowski et al., 2021). This dataset includes digitised H&E-stained histopathology images of lung and colon tissue, curated and labelled by experienced pathologists into two classes. In this pilot evaluation, 40 colon tissue images (20 benign and 20 adenocarcinoma) were selected randomly and transformed into 200 graph samples.

2.2 Pre-Processing

Before implementing any graph-based transformations, all images underwent a structured preprocessing pipeline to maintain uniformity, clarity, and comparability across samples. Each image was rescaled to 512×512 pixels to reduce computational burden while preserving the images' structure. Reinhard colour normalisation was used to equalise the colour distribution in the data. This becomes necessary to address the staining issues that are common in histopathology slides. Afterwards, a Gaussian smoothing filter was applied to enhance tissue contrast and suppress background noise. Subsequently, key spatial points indicating changes in colour and shape were detected through feature extraction. These points would later be considered the nodes when creating the graph models. The preprocessing steps ensured that the detected points were not noise or staining artefacts, but rather the true structural patterns of the tissue.

2.3 Graph Construction

Graph-based modelling provides a sound mathematical framework for describing the spatial organisation observed in histopathology images. Unlike raw pixel-level analysis, which captures colour and texture, graphs encode tissue architecture as nodes and edges, thereby enabling quantitative analysis of structural patterns. In colon histopathology, malignant tissue is typically characterised by disrupted glandular boundaries, spacing, and loss of organised cellular arrangement. These irregularities can best be captured using a graph representation derived from spatial key points detected in each image. For this work, every histopathological image is transformed into a set of feature points obtained through key-point detection, which usually correspond to salient structural features such as high-contrast junctions, texture-rich regions, or morphological transitions, and form the vertex set V of the graph. The following step illustrates how these vertices are connected. In fact, different connection strategies highlight different structural properties of the tissue, and the adopted graph model will dictate which spatial relationships are highlighted.

Five graph construction methods were used: Proximity Graph, Voronoi Diagram, Delaunay Triangulation, Minimum Spanning Tree (MST), and K-Nearest Neighbour (KNN). Each of these methods puts a different geometric constraint on

edge formation. Let each histopathological image be defined over a finite two-dimensional spatial domain R^2 , $\Omega \subset R^2$. Following pre-processing, a key point detector extracts a finite set of salient spatial coordinates $P = \{p_1, p_2, \dots, p_n\}$, $p_i \in \Omega$ which serve as the vertex set of the graph representation. These points typically correspond to texture-rich or morphologically distinct regions in the tissue and capture essential geometric information about local structural patterns $G=(V,E)$, where $V = \{v_1, v_2, \dots, v_n\}$ and $v_i \leftrightarrow p_i$, and E is determined by the chosen geometric construction. Each of the five graph models imposes a distinct rule for edge formation, enabling the extraction of different structural properties of the tissue

2.3.1 Proximity Graph

Proximity graph connects two vertices if the Euclidean distance between their corresponding spatial points is less than a fixed threshold τ , $E_{prox} = \{(v_i, v_j) | \|p_i - p_j\|_2 < \tau\}$, where $\tau > 0$ controls the local neighbourhood radius. This model highlights local density and clustering, with edges reflecting immediate spatial adjacency in the tissue.

Algorithm 1:

Input: Image I

Parameters: threshold distance $\leftarrow 50$ # pixels

minArea $\leftarrow 10$ # for blob detector

$I_p \leftarrow \text{Preprocess}(I)$ # resize, stain-normalize, denoise

Keypoints $\leftarrow \text{DetectBlobs}(I_p, \text{minArea})$

$P \leftarrow [p_1, \dots, p_n] \leftarrow \text{coordinates}(\text{Keypoints})$

if $n < 2$: return empty graph

Initialize $G_{prox} = (V = \emptyset, E = \emptyset)$

for $i = 1 \dots n$:

add node v_i to V with $\text{pos} = P[i]$

for $i = 1 \dots n - 1$:

for $j = i + 1 \dots n$:

$d \leftarrow \text{EuclideanDistance}(P[i], P[j])$

if $d < \text{threshold_distance}$:

add edge (v_i, v_j) to E with $\text{weight} = d$

$A_{prox} \leftarrow \text{binarize_adjacency}(G_{prox})$

Save overlay image, save A_{prox} as CSV

Output: G_{prox}, A_{prox}

2.3.2 Voronoi Diagram

The Voronoi diagram partitions the domain Ω into convex regions such that each region contains all points closer to a specific region than to any other. For each point $p_i \in P$, its Voronoi region is defined as $V_i = \{x \in \Omega | \|x - p_i\|_2 \leq \|x - p_j\|_2, \forall j \neq i\}$. Two nodes v_i and v_j are adjacent if their corresponding Voronoi regions share a boundary of nonzero length $(v_i, v_j) \in E_{vor} \Leftrightarrow \partial V_i \cap \partial V_j \neq \emptyset$. This adjacency captures regions of mutual spatial influence, making Voronoi graphs useful for analysing spatial partitioning and morphological regularity.

Algorithm 2:

Input: Image I , $P = \{p_1 \dots p_n\}$ Parameters: image_bounds B

if $n < 2$: return empty graph

$Vor \leftarrow \text{Voronoi}(P)$ # use SciPy Voronoi

Initialize $G_{vor} = (V, E)$ with $V \leftarrow \{v_1 \dots v_n\}$

for each ridge_points pair (i, j) in $Vor.\text{ridge_points}$:

if ridge is valid (not infinite or clipped within B):

add undirected edge (v_i, v_j) to E

Optionally clip Voronoi polygons to B and render

$A_{vor} \leftarrow \text{binarize_adjacency}(G_{vor})$

Save overlay and CSV

Output: G_{vor}, A_{vor}

2.3.3 Delaunay Triangulation

The Delaunay triangulation is the geometric dual of the Voronoi diagram. A Delaunay graph connects points that share

an edge in the triangulation, which is defined such that the circumcircle of each triangle contains no other point of P . Formally, $(v_i, v_j) \in E_{duy} \Leftrightarrow p_i, p_j$ is an edge of a Delaunay simplex. Equivalently, three points p_i, p_j, p_k form a Delaunay triangle if $circ(p_i, p_j, p_k) \cap P = \{p_i, p_j, p_k\}$ where $circ(\cdot)$ denotes the circumcircle. Delaunay graphs provide stable approximations of local geometric structure, capturing spacing, orientation, and triangle-based connectivity of tissue regions.

Algorithm 3:

```
Input: Image I, P = {p1...pn}
if n < 3: return G_del with V only
Tri ← Delaunay(P) # SciPy Delaunay
Initialize G_del = (V = {v1...vn}, E = ∅)
for each simplex [i,j,k] in Tri.simplices:
    add edges (i,j), (j,k), (k,i) to E
Compute optional features: edge lengths, triangle angles, area
A_del ← binarize_adjacency(G_del)
Save visualization and CSV
Output: G_del, A_del
```

2.3.4 Minimum Spanning Tree (MST)

Given the complete weighted graph $G_w = (V, E_w)$, $E_w = \{(v_i, v_j) \mid i \neq j\}$, where edge weights represent Euclidean distances $w_{ij} = \|p_i - p_j\|_2$, the minimum spanning tree is defined as the acyclic subgraph $G_{mst} = (V, E_w)$, that minimise total edge weight, highlight global structural connectivity, and are sensitive to tissue disorganisation, often reflecting malignant spatial dispersion.

Algorithm 4:

```
Input: Image I, P = {p1...pn}
if n < 2: return empty graph
Initialize complete weighted graph G_full with nodes V
for i = 1...n-1:
    for j = i+1...n:
        add edge (i,j) with weight = EuclideanDistance(P[i], P[j])
G_mst ← MinimumSpanningTree(G_full) # Kruskal or Prim (NetworkX)
A_mst ← binarize_adjacency(G_mst)
Save overlay and CSV
Output: G_mst, A_mst
```

2.3.5 k-Nearest Neighbour (KNN) Graph

For a fixed integer $k \geq 1$, the KNN graph connects each point to its k closest neighbours in Euclidean space. Formally, $(v_i, v_j) \in E_{knn} \Leftrightarrow p_j \in N_k(p_i)$, where $N_k(p_i)$ denotes the set of k nearest points to p_i . The resulting graph expresses local topological consistency, which is independent of point density, and therefore allows for robust characterisation of micro-architectural patterns.

Algorithm 5:

```
Input: Image I, P = {p1...pn}
Parameters: k ← 5
if n < 2: return empty graph
KD ← KDTree(P)
Initialize G_knn = (V = {v1...vn}, E = ∅)
for i = 1...n:
    distances, indices ← KD.query(P[i], k+1) # includes self
    for idx in indices [1:]: # skip self
        add undirected edge (i, idx) to E
(Optional) symmetrize E: E ← {(i,j) | (i,j) in E or (j,i) in E}
A_knn ← binarize_adjacency(G_knn)
Save overlay and CSV (include k in filename)
Output: G_knn, A_knn
```

The five different graph models all show the distinct geometric sides of tissue organisation and thus the histopathology

of the colon can be analysed from different structural angles. Proximity graphs show dense, clustered areas by linking locations within the specified spatial radius. Voronoi graphs indicate how space is divided among the points, providing a visualisation of the areas of influence and their distribution. Delaunay triangulations capture detailed local geometric structure and a degree of spatial uniformity by forming stable triangular connections among neighbouring points. The Minimum Spanning Tree (MST) focuses on global connectivity, measuring how dispersed or compact the overall tissue structure is. K-Nearest Neighbour (KNN) graph provide consistent local neighbourhood information by connecting each point to its closest neighbours, independent of density variations. The authors selected these five methods because together they provide a balanced mix of local, intermediate, and global structural measurements, making them particularly well-suited to distinguishing subtle architectural differences between benign and adenocarcinoma tissues. These graph models transform complex tissue patterns into quantifiable mathematical structures, enabling a thorough and objective comparison of cancer-related disorganization in colon histopathology.

3. Results and Discussion

This pilot study evaluated whether graph-theoretic representations of histopathology images can capture structural differences that distinguish benign colon tissue from adenocarcinoma. From the 40 selected images, five graph models per image were generated, yielding 200 graph samples for training and testing. The results are presented below, along with visual examples of each transformation.

3.1 Key Point Detection

Figure1 shows an example of the key points extracted from a colon tissue section using the Simple Blob Detector. The detector identifies rounded, high-contrast regions that often correspond to nuclei or small glandular structures. These key points form the basis for constructing the graph representations.

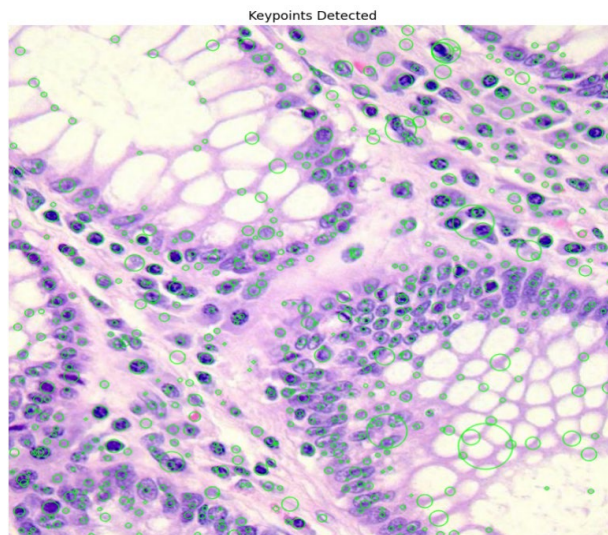


Figure 1: Key Points Detected Across the Colon Tissue Sample. The Green Circles Indicate the Detected Points

Variations in the distribution of key points among images reflect mainly tissue density and ductal pattern. In benign tissue, there is generally more uniform spacing between landmarks. In contrast, an adenocarcinoma sample exhibits irregular and deformed arrangements due to disrupted glandular boundaries and a loss of organised cellular architecture.

3.2 Graph-Based Representations

Five graph variants have been generated for each image using extracted key points, each emphasising a different aspect of spatial organisation.

3.2.1 Proximity Graph

In this graph representation, edges are drawn between the points that lie within a fixed Euclidean distance threshold. It captures local neighbourhoods and clustering tendencies. In benign samples, proximity graphs often appear more evenly connected, whereas areas of adenocarcinoma tend to exhibit dense, irregular connectivity reflecting structural distortion.

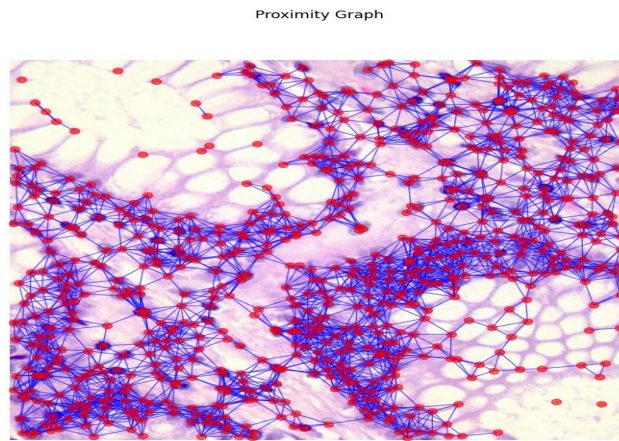


Figure 2: Proximity Graph is Constructed by Connecting the Key Points within a Predefined Spatial Threshold

3.2.2 Voronoi Diagram

Figure 3 shows the Voronoi diagram generated from the same key points. Each cell encloses the region closest to a specific key point. This representation is sensitive to local spatial irregularities. Benign tissue generally exhibits more uniform cell shapes, whereas adenocarcinoma samples yield a mixture of elongated and compressed polygons due to disrupted glandular boundaries.

3.2.3 Delaunay Triangulation

Figure 4 shows the Delaunay triangulation that connects the points into a mesh of non-overlapping triangles. The model accentuates geometric regularity. Benign regions exhibit balanced triangulation patterns, whereas malignant tissue shows sharp angles and uneven triangulation due to abnormal cellular arrangement.

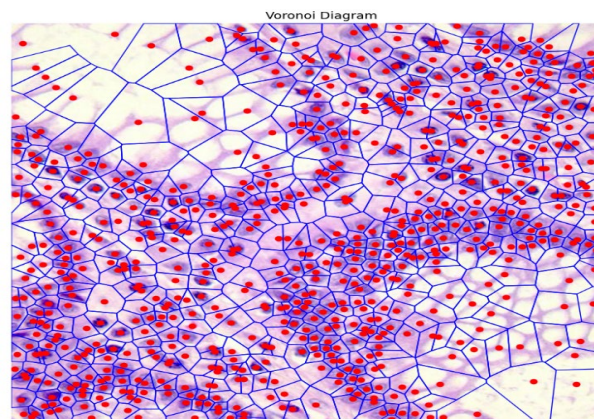


Figure 3: Voronoi Diagram Showing Spatial Partitioning around Extracted Key Points

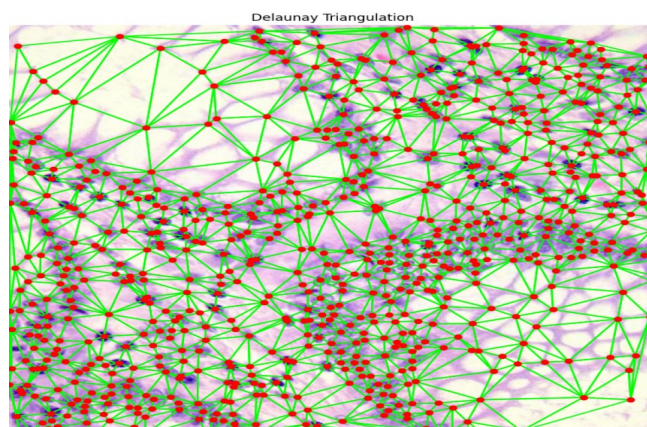


Figure 4: Delaunay Triangulation Showing Spatial Partitioning around Extracted Key Points

3.2.4 Minimum Spanning Tree (MST)

Figure 5 shows the MST, which connects all key points using the smallest possible total edge length. The MSTs are especially useful to identify the structural "skeleton" of the underlying tissue region. Adenocarcinoma samples produce longer, irregular branches, while benign tissue has smoother branching patterns.

3.2.5 K-Nearest Neighbours ($k = 5$)

Figure 6 shows the KNN graph for $k = 5$, in which every point is connected to its nearest neighbours. The KNN graph captures local micro-architectural features. In malignant areas, the KNN graph becomes densely packed due to increased nuclear clustering, a well-documented histological hallmark of dysplasia and carcinoma.

Minimum Spanning Tree (MST)

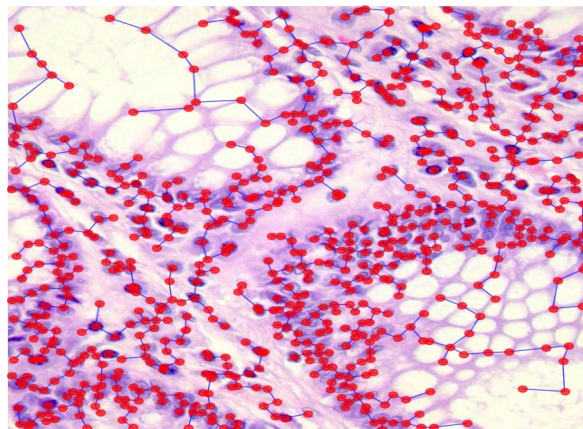


Figure 5: Minimum Spanning Tree and Capturing the Most Efficient Connectivity across the Tissue

k-Nearest Neighbors ($k=5$)

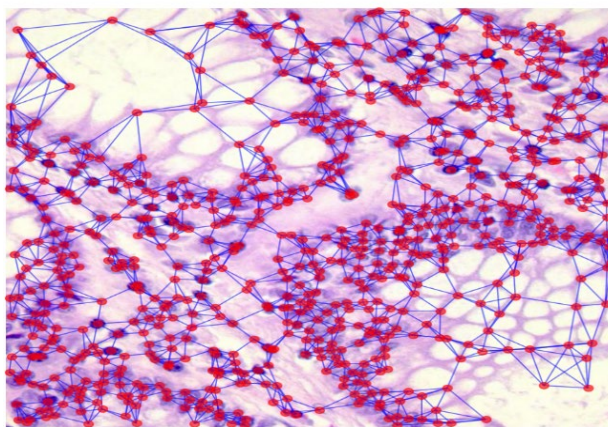


Figure 6: k-Nearest Neighbour Graph Capturing the Most Efficient Connectivity across the Tissue

3.3 Model Performance

These 200 graph samples were first converted to their adjacency matrices, standardised to 100×100 dimensions, and then used to train a Graph Convolutional Network (GCN), a specialised machine learning model for automated classification of colon tissue images. Using an 80:20 train-test split (160 training graphs, 40 test graphs), the network achieved a training accuracy of more than 98% and a final test accuracy of 90%. It follows that even with a rather small dataset, structural information encoded by the graph representations is enough to perform meaningful classification

3.4 Confusion Matrix Analysis

Figure 7 shows the confusion matrix for the 40 test graphs. Analysis of the confusion matrix reveals that most classification results were on the diagonal, confirming that the GCN accurately identified the majority of the graph samples.

Notably, the Delaunay triangulation and Minimum Spanning Tree representations achieved perfect identification, highlighting their superior geometric consistency in capturing the tissue's structural properties. In contrast, some minor classification errors occurred among the proximity, Voronoi, and K-Nearest Neighbour models; this is expected, as these specific construction methods can sometimes produce overlapping connectivity patterns that are more difficult for the model to distinguish.

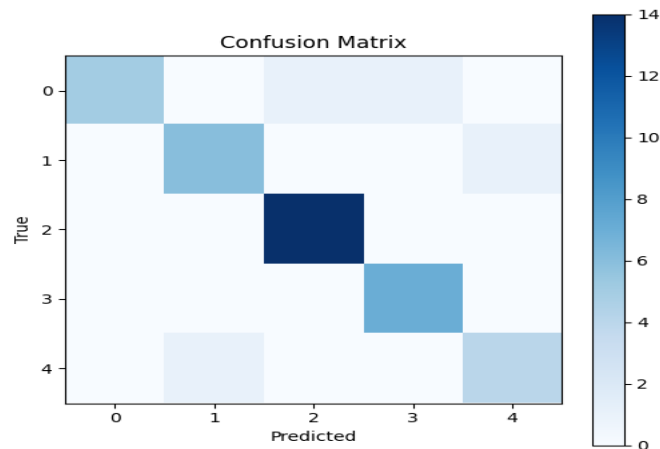


Figure 7: Confusion Matrix Showing Classification Performance Across the Five Graph Types

4. Conclusion

This pilot study investigated the application of graph-theoretic representations to characterise structural patterns in colon histopathology images. Key point extraction from each image and transformation into five different graph models, Proximity, Voronoi, Delaunay, Minimum Spanning Tree, and k-Nearest Neighbours, were performed. The researchers inquired whether these spatial abstractions contain sufficient information to differentiate non-cancerous tissue from adenocarcinoma. The visualizations indicate that the areas of malignancy are characterized by irregular and very interconnected graphs, while the areas of non-cancerous tissues are more likely to show the emergence of regular and predictable patterns. Taking into account the classifier's performance, it is inferred that the geometry of cellular arrangement, as depicted via graph relationships, can serve as an attractive feature space for computational pathology. Despite the dataset being purposely restricted, these outcomes not only confirm the effectiveness of such a method but also indicate that this can be extended to larger image collections and more complex graph-learning models. Overall, the findings point to the use of graph-based modelling as a tool for interpreting complex histological architecture and, perhaps even more, as a support for conventional image analysis techniques when building future diagnostic systems.

5. Acknowledgements

We sincerely thank Pradhan Mantri Uchchatar Shiksha Abhiyan (PM-USHA) Project, under the Multi-Disciplinary Education and Research Universities (MERU) grant sanctioned to Sri Padmavati Mahila Visvavidyalayam, Tirupati, Andhra Pradesh.

6. References

- Borkowski, A. A., Bui, M. M., Thomas, L. B., Wilson, C. P., DeLand, L. A., & Mastorides, S. M. (2021). Lung and colon cancer histopathological images dataset| Kaggle. In.
- De Matos, J., Ataky, S. T. M., de Souza Britto Jr, A., Soares de Oliveira, L. E., & Lameiras Koerich, A. (2021). Machine learning methods for histopathological image analysis: A review. *Electronics*, 10(5), 562.
- Demir, C., & Yener, B. (2005). Automated cancer diagnosis based on histopathological images: a systematic survey. *Rensselaer Polytechnic Institute, Tech. Rep.*
- Doyle, S., Hwang, M., Shah, K., Madabhushi, A., Feldman, M., & Tomaszewski, J. (2007). Automated grading of prostate cancer using architectural and textural image features. 2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro,
- Gurcan, M. N., Boucheron, L. E., Can, A., Madabhushi, A., Rajpoot, N. M., & Yener, B. (2009). Histopathological image

analysis: A review. *IEEE reviews in biomedical engineering*, 2, 147-171.

JS, P., & VS, S. (2015). Structural Representation of Colon Tissue using Delaunay Triangle for Cancer Detection and Grading. *International Journal of Computer Applications*, 128(13), 12-15.

Lippoldt, F. (2020). Efficient colon cancer grading with graph neural networks. *arXiv preprint arXiv:2010.01091*.

Moscalu, M., Moscalu, R., Dascălu, C. G., Țarcă, V., Cojocaru, E., Costin, I. M., Țarcă, E., & Șerban, I. L. (2023). Histopathological images analysis and predictive modeling implemented in digital pathology—current affairs and perspectives. *Diagnostics*, 13(14), 2379.

Organization, W. H. (2020). *WHO report on cancer: setting priorities, investing wisely and providing care for all*. World Health Organization.

Sharma, H., Zerbe, N., Lohmann, S., Kayser, K., Hellwich, O., & Hufnagl, P. (2015). A review of graph-based methods for image analysis in digital histopathology. *Diagnostic pathology*, 1(1).

Xiang, X., & Wu, X. (2021). Multiple instance classification for gastric cancer pathological images based on implicit spatial topological structure representation. *Applied Sciences*, 11(21), 10368.