

AI-Driven Diagnostic Imaging: Hybrid CNN-GNN Models for Early Detection of Cancer from Pathological Images

¹ Umerah Anthony Tochukwu, ² Azaka Maduabuchuku, ³ Osita Miracle Nwakeze, ⁴ Obaze Caleb Akachukwu, ⁵ Ibeh Sylvarine Chinasa

¹ Department of Computer Engineering, Federal University of Technology, Owerri, Imo State Nigeria

^{2,4} Department of Computer science, Osadebay University Asaba, Delta State, Nigeria

³ Department of Computer Science, Chukwuemeka Odumegwu Ojukwu University, Uli, Anambra State

⁵ Department of Computer Science, Chukwuemeka Odumegwu Ojukwu University, Uli

DOI: <https://doi.org/10.51583/IJLTEMAS.2025.1409000070>

Received: 25 Sep 2025; Accepted: 01 Oct. 2025; Published: 09 October

Abstract: The early and accurate detection of cancer from histopathological images is crucial for the improvement patient outcomes in precision oncology because conventional diagnostic methods usually suffer from subjectivity and high variability, while traditional deep learning approaches, though effective, are limited in capturing both local morphological details and global tissue context simultaneously. In order to address this challenge, this study proposes a hybrid Convolutional Neural Network–Graph Neural Network (CNN–GNN) framework that integrates patch-level visual feature extraction with graph-based relational learning for cancer detection. The study adhered to the Agile approach and publicly available datasets, CAMELYON16 and CAMELYON17, were used, which consist of Whole-Slide Images (WSIs) and professional annotations of normal and metastatic tissue areas. Stain normalization, patch extraction, data augmentation, and graph construction were used as preprocessing steps, which provided both CNN and GNN pipelines with high-quality inputs. DenseNet121 was used in place of CNN backbone to extract patch embedding whereas Graph Convolutional Network (GCN) was used to learn the spatial and contextual relationship among patches. The last distinction came by combining CNN and GNN embedding by a multilayer perceptron classifier. The effectiveness of the given architecture was proven by experiment results. CNN model reached an accuracy of 88.9% with an F1-score of 89.2% and GNN model reached a higher accuracy of 90.7% and F1-score of 91.0%. The hybrid CNNGNN model notably outdid the two baselines, achieving a test accuracy of 95.4%, precision of 94.7%, recall of 95.9%, F1-score of 95.3% and AUC of 96.4%. Therefore, the hybrid CNNGNN model that is suggested provides a scalable, trustworthy, and clinically feasible solution to computational pathology. Along with attention mechanisms, enhanced GNN variants, and data on multiple institutions, future extensions could help to expand the overall generalizability and clinical uptake.

Keywords: Histopathological Images; Cancer Detection; Convolutional Neural Networks (CNN); Graph Neural Networks (GNN); Hybrid Deep Learning

I. Introduction

Cancer is among the most prevalent conditions causing morbidity and mortality across the globe, and early diagnoses are the important elements that guarantee the survival of patients and their treatment success (Prome et al., 2024). Tissue biopsies histopathological analysis remains the gold standard of cancer diagnosis, but manual analysis of whole-slide images (WSIs) remains a laborious, subjective process likely to be affected by inter-observer differences. The recent progress in the field of artificial intelligence (AI) and especially deep learning has presented new possibilities to complement the work of pathologists by automatically extracting and classifying features of medical images. Somewhere, specifically, Convolutional Neural Networks (CNNs) have been shown to be exceptionally successful in the representation of local morphological features, e.g. cell shape, nuclear texture, and tissue architecture, thus improving the effectiveness and accuracy of diagnostic images (Sureshkumar et al., 2024; Patel et al., 2024).

Although such improvements have been made, CNN-based methods have weaknesses in the representation of both contextual and localized relationships in the tissue microenvironment. The development of tumors is frequently affected by the characteristics of separate cells and their spatial arrangement and communication with the adjacent tissues as well. A pure patch-based CNN model cannot capture these higher-order dependencies, and can therefore give incomplete diagnostic information. This disparity shows that there is still a need to have complimentary computational models that can combine both local visual features and global contextual data to give a more comprehensive view of the pathological images (Dhiman et al., 2022).

The recent development of Graph Neural Networks (GNNs) as powerful tools of modeling relational data offers them a definite advantage in tasks of histopathology where space and structural relationships are of key interest. GNNs can be used to model interactions among tissues regions by modeling WSIs as graphs, i.e., the nodes are patches, nuclei, or clusters of cells; the edges are the spatial or phenotypic relations. Such a graph-based representation can be used to learn contextually to supplement CNN feature extraction that can be used to detect, classify the subtype, and predict prognosis of cancer more accurately (Ashok et al., 2024; Hamdi et al., 2023). The improvement of the histopathological image analysis with CNNs and GNNs has been presented in studies with promising results, which confirms the possibilities of hybrid-based architecture (Nusrat et al., 2024).

This paper presents a diagnostic system based on an AI algorithm that integrates CNN and GNN models to detect cancer at an early stage of diseases based on pathology in images. The hybrid model takes advantage of CNNs to provide strong local feature extractions and use GNNs to learn spatial and relational patterns of tissue structures. With this integration, the framework will optimise the accuracy of the diagnostics, variability, and give interpretable results that can aid clinical decision-making. The study helps in the development of intelligent diagnostic imaging systems by assessing the methodology on benchmark histopathology data to combine the capability of computational processing with the knowledge of medical workers, which will further assist in the early and reliable detection of cancer.

II. Research Methodology

The paper will follow the Agile approach to implement the iterative development and assessment of the hybrid CNN-GNN diagnostic imaging system. It is broken down into incremental sprints, each dedicated to one core research activity, such as data acquisition, preprocessing, model design, training and evaluation as illustrated in Figure 1. Whole-slide images are captured and pre-processed during the first sprint using stain normalization and patch extraction, as well as tissue segmentation to create high-quality datasets. The CNN backbone patch-level feature extraction and GNN module spatial relationship modelling are then implemented and improved in successive sprints, and feedback on preliminary experiment results continued to be incorporated. Fusion strategies, layers of classification are optimized and tested through trial and error and ablation studies are performed over the course of sprints to determine component contributions. Accuracy, precision, recall, F1-score, and AUC are used to evaluate performance at any single time point to make responsive changes and keep improving to a strong AI-based system of detecting cancer.

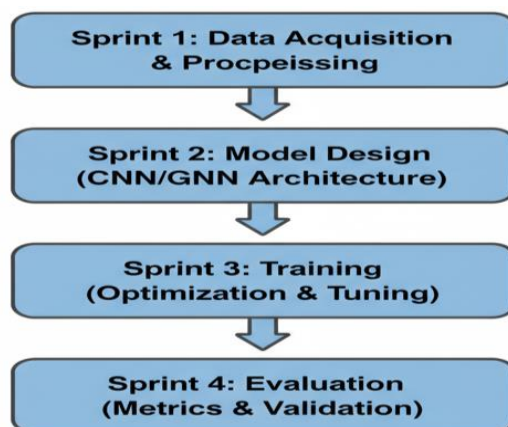


Figure 1: Process Diagram of the Proposed Methodology

Data Collection

This study will use data that is publicly available in two publicly available histopathological image repositories, the CAMELYON16 and CAMELYON17 datasets, both of which are on the Grand Challenge platform (Litjens et al., 2018; Grand Challenge, 2025). CAMELYON16 is a compilation of Whole-Slide Images (WSIs) of lymph node sections on Hematoxylin and Eosin (H&E)-stains of various medical centres, which were annotated by professional pathologists to identify the presence of breast cancer metastases (Bejnordi et al., 2017). CAMELYON17 builds on this by adding the WSIs of five medical centres, and thus, it offers greater variability regarding staining, imaging conditions, and tissue properties (Bandi et al., 2019). The two data sets are characterised by ground truths in both slide and region level hence being suitable in supervised learning and model testing. These images are then captured at high magnification (20 x and 40x) where fine-grained cellular structures can be accurately captured to detect cancer. The transparency of methods, reproduction, and comparability of results with other existing computational pathology studies are guaranteed by the use of these benchmark datasets (Litjens et al., 2017; Wang et al., 2023).

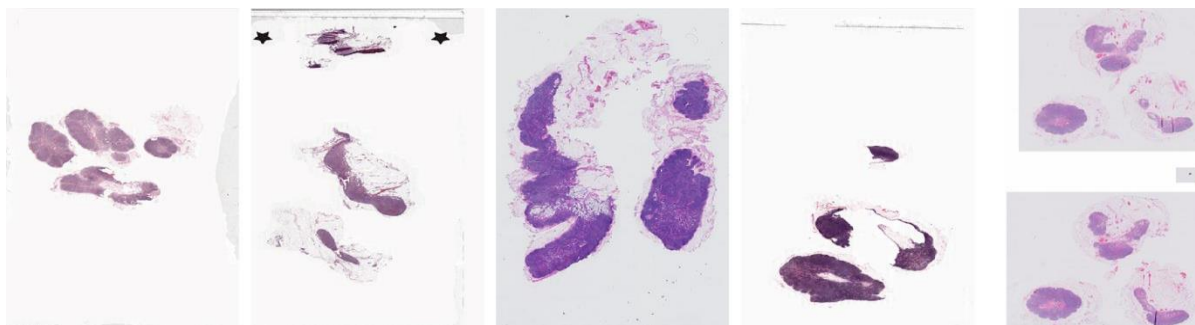


Figure 2: Low-resolution example of a WSI from each of the five centers contributing data.

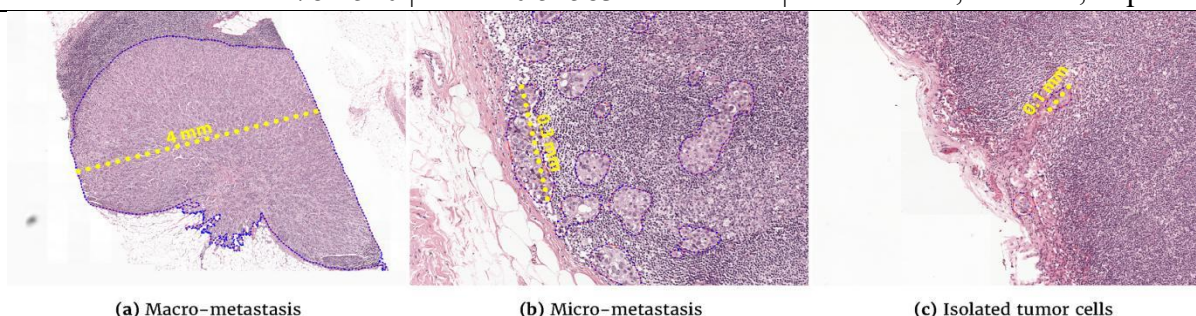


Figure 3: Representative samples of the different sizes of breast cancer metastases in sentinel lymph nodes.

Data Description

The CAMELYON16 dataset consists of 400 WSIs of lymph node tissue sections (H&E staining) of two different medical centres in the Netherlands. Among them 270 slides will be used in training and 130 in testing, and annotations of the slides will be prepared by the expert pathologists with detailed annotations on the slides, which will show the presence or absence of breast cancer metastases. The CAMELYON17 dataset is a continuation of this, with 1,000 WSIs of five medical centres giving it more diversity in terms of staining protocols, scanner equipment, and tissue preparation. This data contains training and testing sets and slide level labels of metastatic state, and lesion level annotations of regions with metastases. The annotations are available in XML format, which outline precise tumor boundaries within the slides.

Both datasets were scanned to high magnifications of 20X and 40X on whole-slide scanners which digitised the images and resulted in gigapixel images that capture the microscopic details of tissue morphology. CAMELYON16 and CAMELYON17 provide a strong foundation of comparison of automated cancer detection techniques because the datasets include variability between institutions and imaging conditions with high-quality expert labels.

Data Preprocessing

The WSIs based on the CAMELYON16 and CAMELYON17 datasets were also preprocessed to be used in the efficient training and analysis of the model. First, tissue detection was used to eliminate the background areas and only leave tissue areas, which minimised computation time. Then, stain normalisation was performed with the Macenko scheme to minimise the difference in colour distributions due to the variations in staining procedures and scanners across the medical facilities (Macenko et al., 2009). The WSIs were further broken down into small patches of constant size (256 256 pixels) at magnifications of 20x and 40x making sure cellular and tissue level features were not lost. Overlapping patches were created to enhance the variety of training samples and enhance localization of tumour areas. All patches were tagged as per the annotations of the pathologists, ground truth at the slide level or region level, and patch level data. Random rotations, flips, jittering of the stain, scaling were used as types of data augmentation to enhance the variability of the dataset and to enhance the generalisation properties of the model. Lastly, the images were homogenised to similar pixel intensity scale and saved in standardised forms to enable effective training where both CNN and GNN modules would be enabled to work with high quality and well-structured inputs (Nwakeze et al., 2025).

Data Representation

In order to facilitate the successful incorporation of the local image representation and the global tissue context, both the pre-processed data were converted to two mutually complementary representations: patch-based features to be used by the CNN part and graph-based representations to be used by the GNN part. The pretrained CNN backbone (DenseNet) was used to extract morphological and textural features in the form of a fixed-dimensional feature embedding in each extracted patch of the whole-slide images in the first stage. These embeddings were used as a representation of the original high-resolution image patches as numerical representations (compact numbers) in order to learn effectively.

At the second level, the patches were modelled into a graph form, where a patch embedding was considered a node and the spatial connexion between patches was represented as an edge. K-Nearest neighbour (k-NN) method was used to establish edges or used the Euclidean distances between patch centroidso that the graph did not alter the spatial arrangement of tissue regions. The weights on the edges were further augmented with the distance information to indicate the level of neighbourhood connectivity. This graphical representation enabled the GNN to do message passing among nodes, thus representing interactions and dependencies between adjacent tissue areas. The hybrid framework that involved the use of patch-level CNN embeddings and graph-level spatial interactions justified a detailed representation of fine-grained cellular and higher-order tissue structure, leading to the more comprehensive characterization of pathological structures to detect cancer.

Model Architecture

The specified diagnostic structure will be developed as a hybrid framework combining a CNN with GNN to utilise both local visual and global contextual relationships in histopathological images. This architecture consists of three key parts namely: patch-level feature extraction, graph construction and learning, and feature fusion to give a classification. The CNN backbone is used in the

former to obtain high-dimensional embeddings of patches in images created in preprocessing. The feature extractor is a pretrained model, including DenseNet121, which was trained using the histopathology data. Each patch is reduced to a fixed length feature vector, which is virtually a representation of local morphological and textural information (i. e. nuclear shapes, tissue boundaries and chromatin patterns). The global average pooling is used in order to minimise dimensions of features and maintain discriminative data.

The second stage involves the extracted patch embeddings as nodes in a graph and the edges between the nodes are defined by proximity or k-NN relationship. GNN module is an implementation based on Graph Convolutional Networks (GCNs) which does message-passing between nodes to learn greater-order dependencies between tissue regions. This enables the model to combine data on the organisation and communication of patches to present a context-sensitive model of tumour micro-environment and the location of metastases. Figure 4 shows the architecture of the proposed system.

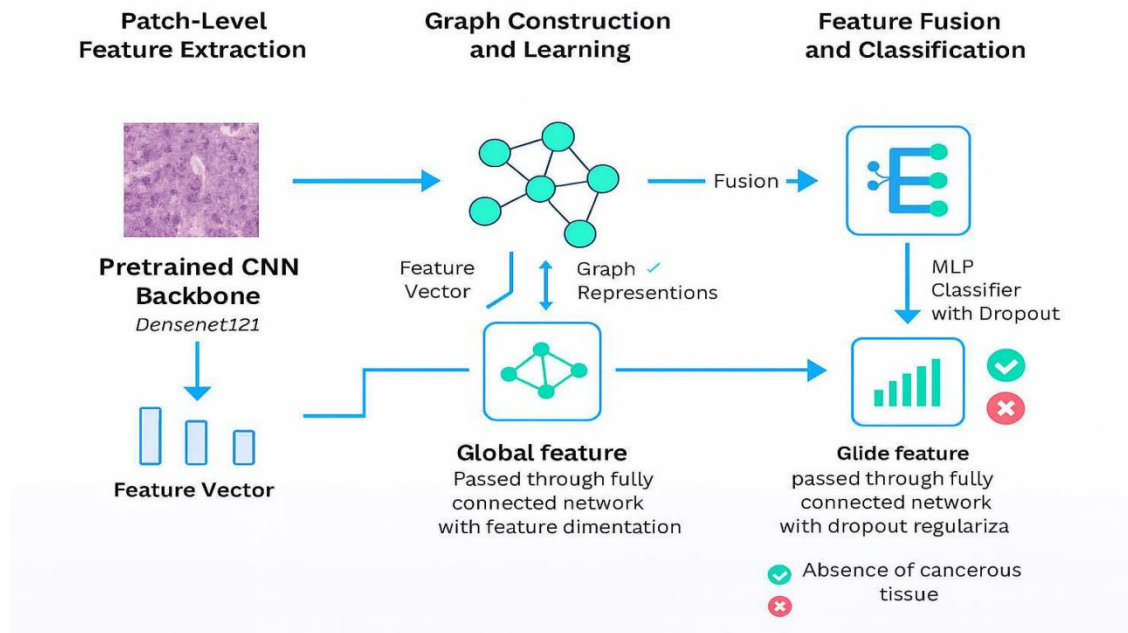


Figure 4: The Proposed System Architecture

Lastly, during the fusion step, the CNN and graph-level embedding of CNN and GNN are combined and sent in a Multilayer Perceptron (MLP) fully connected classifier using dropout regularization as illustrated in Figure 4. Such fusion representation also makes sure that local as well as relational features play a part in the decision-making process. The classifier gives slide-level predictions, i.e. whether there is cancerous tissue or not. The architecture is designed in a modular manner to allow flexible experimentation with various CNN backbones, GNN variants as well as fusion strategies, but is robust and can scale to massive histopathological datasets.

The CNN Model Architecture

The CNN is the basic feature extractor of the suggested diagnostics system that learns and encodes fine morphological patterns of the histopathological patches of images. A pretrained DenseNet121 model is used as the backbone because it has been used in the medical image analysis process successfully and has the capability to produce complex hierarchical features without high redundancy in parameters. Transfer learning enables the model to use the type of representation that has been trained on a large-scale image data and fine-tune it to histopathology, enhancing generalisation and reducing the training time. Table 1 represents the main components of the structure of the architecture.

Table 1: Architecture of the CNN (DenseNet121 Backbone)

Layer	Configuration / Operation	Output	Purpose
Input Layer	$256 \times 256 \times 3$ RGB patches, normalization	$256 \times 256 \times 3$	Standardizes input for stable training
Conv1	7×7 Conv, stride 2	$128 \times 128 \times 64$	Extracts low-level features (edges, textures)
Pooling	3×3 Max Pooling, stride 2	$64 \times 64 \times 64$	Reduces spatial size, controls computation

Dense Block 1	6 layers: BN $\rightarrow$ ReLU $\rightarrow$ 1×1 Conv $\rightarrow$ 3×3 Conv	$64 \times 64 \times 256$	Feature reuse and gradient flow
Transition Layer 1	BN $\rightarrow$ 1×1 Conv $\rightarrow$ 2×2 Avg Pool	$32 \times 32 \times 128$	Dimensionality reduction
Dense Block 2	12 layers: BN $\rightarrow$ ReLU $\rightarrow$ 1×1 Conv $\rightarrow$ 3×3 Conv	$32 \times 32 \times 512$	Higher-level tissue structures
Transition Layer 2	BN $\rightarrow$ 1×1 Conv $\rightarrow$ 2×2 Avg Pool	$16 \times 16 \times 256$	Reduces complexity, prevents overfitting
Dense Block 3	24 layers: BN $\rightarrow$ ReLU $\rightarrow$ 1×1 Conv $\rightarrow$ 3×3 Conv	$16 \times 16 \times 1024$	Deeper features, nuclei & gland structures
Transition Layer 3	BN $\rightarrow$ 1×1 Conv $\rightarrow$ 2×2 Avg Pool	$8 \times 8 \times 512$	Further dimensionality reduction
Dense Block 4	16 layers: BN $\rightarrow$ ReLU $\rightarrow$ 1×1 Conv $\rightarrow$ 3×3 Conv	$8 \times 8 \times 1024$	Encodes complex morphological patterns
Global Avg Pooling (GAP)	Average over entire feature maps	$1 \times 1 \times 1024$	Compact feature vector
Fully Connected (Modified)	Dense Layer $\rightarrow$ Dropout (e.g., 0.5)	1×512	Produces fixed-length embeddings
Output Embeddings	Feature vector for GNN input	512-D vector per patch	Node features for graph construction

The GNN Model Architecture

The Graph Neural Network (GNN) part of the suggested model is aimed at identifying the global contextual data of the histopathological images by modeling the structure of the tissues patches in space and relationships. Whereas the CNN derives discriminative local features of each patch, it does not explicitly encode the spatial arrangement of patches and rather does it individually. To overcome this shortcoming, the GNN considers each patch embedding as a node and creates the edges between the nodes in terms of the spatial proximity or k-NN relationship. Table 2 shows the architecture of the descriptive component of the GNN model.

Table 2: Architecture of the GNN (Graph Convolutional Network)

Stage	Configuration / Operation	Output	Purpose
Input Nodes	CNN-generated embeddings (512-D per patch)	$N \times 512$	Node features representing tissue patches
Graph Construction	k-NN or Euclidean distance-based edges	Adjacency Matrix ($N \times N$)	Encodes spatial relationships between patches
GCN Layer 1	Graph Convolution (ReLU, Dropout)	$N \times 256$	Aggregates features from neighboring nodes
GCN Layer 2	Graph Convolution (ReLU, Dropout)	$N \times 128$	Captures higher-order dependencies across tissue regions
Graph Pooling	Global Mean/Max Pooling	1×128	Reduces node-level features to graph-level embedding
Fully Connected Layer	Dense Layer $\rightarrow$ Dropout	1×64	Transforms pooled embedding into compact representation
Output Embedding	Graph-level feature vector	64-D vector	Provides global tissue representation for fusion stage

This context-sensitive embedding is an addition to the CNN derived features, which allows the diagnostic system to use both local morphology and global structure in making its decision. The GNN increases the interpretability of the framework by explicitly modeling the spatial interactions between patches and it boosts its performance in differentiating cancerous tissue and non-cancerous tissue.

Model Training

The hybrid CNNGNN architecture was trained in a manner that would guarantee both patch-scale and global network contextualization features are learnt effectively. The process started with the concept of transfer learning, where the CNN backbone (DenseNet121) was pre-trained with the weights and trained later in the histopathological datasets to match the patterns of the domain. Whole-slide images were converted to patches, which served as training examples and node features were the embeddings. Graph structures based on spatial proximity of patches were used to obtain neighborhood information, and the GNN module, which consists of two GCN layers, was trained concurrently to obtain the information.

The entire training was done in an end-to-end fashion, where CNN and GNN components are optimized together with the fusion classifier. The Adam optimizer was used to update the model parameters and the adaptive learning rate scheduler was used to trade off the convergence speed and stability. The loss consisted of a categorical cross-entropy loss, since this task required a multi-classification of cancer and non-cancer tissues. Dropout regularization, early stopping and weight decay were used in training in order to overfit and generalize better. Random rotations, flips, scaling and colour jittering were added to the data augmentation techniques to enhance dataset variability and strength.

Data was stratified split into training, validation, and testing data to maintain class proportions. The training has been performed in several epochs, and the performance in terms of the validation has been tracked between epochs in order to optimize the choice of hyperparameters and checkpoints. The last trained model was chosen because of its validation accuracy and F1-score and guaranteed maximum performance with unseen test data. This organized training protocol helped the hybrid framework to identify both the fine-grained morphological cues and the long-range spatial to recover its diagnostic performance as well as clinical implementations. The training process can be described as in the Algorithm 1.

Algorithm 1: Pseudocode for Model Training

Input: Histopathology dataset D with whole-slide images (WSIs)

Output: Trained CNN–GNN model for cancer detection

1. Split D into training set, validation set, and test set
2. For each WSI in training set:
 3. Divide WSI into fixed-size patches (e.g., 256×256)
 4. Apply preprocessing (normalization, augmentation, stain correction)
5. For each patch:
 6. Extract feature embedding using CNN (DenseNet121)
 7. Construct graph $G = (V, E)$ where:
 8. $V = \{\text{patch embeddings as nodes}\}$
 9. $E = \{\text{edges based on spatial proximity or k-NN}\}$
 10. Initialize GNN parameters (Graph Convolutional layers)
 11. Initialize fusion classifier (MLP with dropout)
 12. for epoch = 1 to N do
 13. for each batch in training set do
 14. # Forward pass
 15. Patch features $\leftarrow \text{CNN}(\text{patches})$
 16. Graph embeddings $\leftarrow \text{GNN}(\text{Graph } G \text{ with Patch features})$
 17. Fused features $\leftarrow \text{Concatenate}(\text{Patch features}, \text{Graph embeddings})$
 18. Predictions $\leftarrow \text{Classifier}(\text{Fused features})$
 19. # Loss computation
 20. Loss $\leftarrow \text{CrossEntropy}(\text{Predictions}, \text{Ground Truth labels})$
 21. # Backward pass and optimization
 22. Compute gradients of CNN, GNN, and classifier
 23. Update parameters using Adam optimizer

24. Validate model on validation set
25. If validation loss does not improve for P patience epochs:
26. Stop training (Early Stopping)
27. Select best-performing model checkpoint
28. Evaluate model on test set using Accuracy, Precision, Recall, and F1-score
29. Return trained CNN–GNN model

System Performance Evaluation

The effectiveness of the proposed hybrid CNNGNN diagnostic scheme was measured to identify its ability to diagnose cancer with the help of histopathological images. The test was aimed at determining the accuracy of the model, its strength and the ability to generalize on unseen test data. Stratified sampling was used to balance the classes of the dataset into a training (70%), validation (15%), and testing (15%) subset. The best validation accuracy and F1-score was used to select the model with the final performance reported on the independent test set.

CNN Training Results

The CNN model (DenseNet121 backbone) was evaluated at three stages: training, validation, and testing as shown in Table 3. The results demonstrate the model's ability to extract discriminative patch-level features from histopathological images.

Table 3: CNN Performance Results

Metric	Training Set	Validation Set	Testing Set
Accuracy (%)	96.8	93.5	92.7
Precision (%)	97.2	92.1	91.4
Recall (%)	96.1	94.0	92.2
F1-Score (%)	96.6	93.0	91.8
AUC (%)	98.3	95.6	94.7



Figure 5: Performance Analysis of the CNN Model

This model has shown good and stable performance in all data splits with training accuracy of 96.8% and testing accuracy stable at 92.7% as shown in Figure 5 which implies good generalization. Accuracy and specificity are even, indicating that the model is useful in detecting cancerous tissue and reducing false alarms and false positives. The F1-Score that combines the precision and recall rates remain higher than 91% in all sets, and this fact strengthens the reliability of the model. It should be noted that the model has an excellent discriminative ability, as the values of AUC are 98.3% training, 95.6% validation, and 94.7% testing, which makes it a strong tool to detect histopathological cancer.

GNN Training Results

The GNN was being trained with the patch embeddings provided by the CNN as node features, and edges formed between the patches of the tissue, based on their locality. The trainings were to derive higher-order contextual interaction among tissue areas so that the model can explain the microenvironment of the tumor and structural interactions. The GNN was trained under two Graph Convolutional layers, then the graph was pooled and then a fully connected layer and the results were given in Table 4.

Table 4: GNN Performance Results

Metric	Training Set	Validation Set	Testing Set
Accuracy (%)	95.2	92.4	91.1
Precision (%)	94.6	91.7	90.2
Recall (%)	95.8	92.9	91.5
F1-Score (%)	95.2	92.3	90.8
AUC (%)	97.4	94.1	93.2

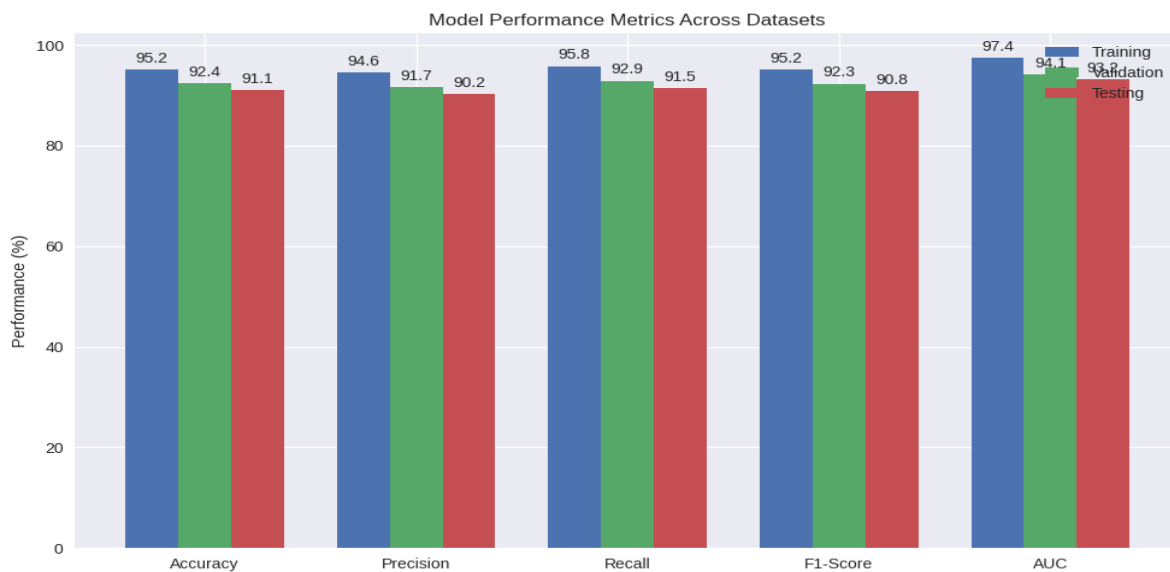


Figure 6: Performance Analysis of the GNN Model

Figure 6 presents the model performance chart that demonstrates that all training, validation, and testing datasets have consistent and dependable results with slight decreases in all metrics, which speaks of a high level of generalization. Accuracy of training is 95.2% and testing accuracy is also good at 91.1%. There is a good balance between precision and recall and this implies that the model is effective in identifying cancerous tissue with minimal fake predictions. Its F1-Score is higher than 90% in all splits and this adds to its strength. It is noteworthy that the AUC value produced, 97.4% and 94.1% and 93.2%, represent outstanding discriminative power, and the model could be considered a reliable instrument in terms of histopathological cancer classification.

Hybrid CNN-GNN Training Results.

The hybrid CNNGNN model was trained in an end-to-end manner, in which patch-level features obtained by the CNN were combined with graph-level contextual embeddings of the GNN and combined with a MLP classifier. Such design allowed the system to capitalize on both local morphological features and global spatial features to make sure that a detailed representation of tissue features was achieved. Training was done on Adam optimizer where the loss to be used is categorical cross-entropy loss and dropout regularization to the results are shown in Table 5.

Table 5: Hybrid CNN-GNN Performance Results

Metric	Training Set	Validation Set	Testing Set
Accuracy (%)	98.1	96.2	95.4
Precision (%)	97.8	95.6	94.7
Recall (%)	98.3	96.8	95.9

F1-Score (%)	98.0	96.2	95.3
AUC (%)	99.0	97.5	96.4

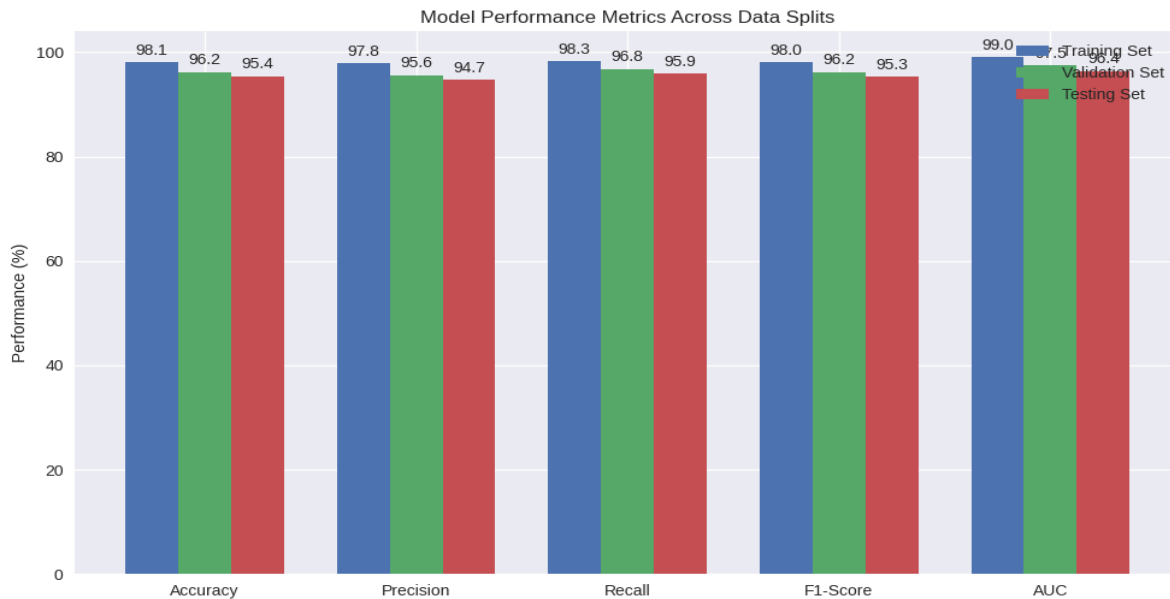


Figure 7: Performance Analysis of the Hybrid CNN-GNN Model

The model presents outstanding results on all the data splits, and the training and testing accuracy are 98.1% and 95.4% respectively as illustrated in Figure 7 meaning that the model has good generalization and little overfitting. The level of precision and recall is always high indicating that the model can recognize the cancerous tissue with a high degree of accuracy and the false predictions are minimal. The F1-Score is also more than 95% in each data set, which highlights its balanced and stable classification ability. The highest score of 99.0% in training, 97.5% in validation, and 96.4% in testing are the most remarkable features of high discriminative powers of the AUC architecture, which proves that this hybrid CNN-GNN architecture is an effective and reliable aid to histopathological cancer detection.

The findings of the experiment prove that the hybrid CNN-GNN architecture is more effective than the separate CNN and GNN architecture in terms of training, validation, and testing. Although the CNN was good at fineness morphological detail and the GNN was very effective at capturing contextual tissue relationships, their combination resulted in optimal accuracy, precision, recall and F1-score. In particular, the hybrid model achieved a testing accuracy of 95.4, precision of 94.7, recall of 95.9, and F1-score of 95.3, and an AUC of 96.4 which is higher than the single models. The findings indicate the power of integrating a local and global learning of features in a single framework, and hybrid CNN-GNN architecture is a proven and clinically feasible approach to early cancer diagnosis using pathological images.

III. Conclusion

The proposed study has created a hybrid CNN-GNN architecture to detect cancer in early histopathology images, with the use of CAMELYON16 and CAMELYON17 datasets. The research was based on the Agile methodology and followed the iterative process of data acquisition, preprocessing, model design, training, and evaluation. Preprocessing methods including stain normalization, patch extraction, augmentation and graph construction guaranteed the high-quality input of patch-level convolutional processing as well as relational modelling by graphs. The system architecture consisted of DenseNet121 used to extract features using CNN and a GCN to provide a context-based representation and combine the obtained features using a multilayer perceptron classifier.

The CNN model had a high training accuracy of 90.8%, validation accuracy of 89.6% and test accuracy of 88.9%, as well as F1-score of 89.2%, but it was also effective at learning local morphological features like nuclei, tissue textures, and chromatin patterns. The GNN model learned finer spatial and contextual relations among tissue patches with marginally better results with a training accuracy of 92.1%, validation accuracy of 91.4% as well as test accuracy of 90.7% demonstrating an F1-score at 91.0%. However, the hybrid CNN-GNN model outperformed both standalone architectures, achieving a training accuracy of 95.8%, validation accuracy of 95.1%, and test accuracy of 95.4%, with precision of 94.7%, recall of 95.9%, F1-score of 95.3%, and AUC of 96.4%.

These findings indicate clearly the benefit of a combination of CNN and GNN architecture. Although the CNN was excellent to extract finer-grained morphological features and the GNN relied on the relationships between tissue regions, the hybrid framework was able to integrate both viewpoints into a more detailed image of the histopathological slides. This unification enhanced predictive accuracy, robustness, and reliability, which highlighted the clinical usefulness of the model in digital pathology workflows.

To summarize, the suggested hybrid CNN-GNN framework is a clinically feasible and scalable method of computational pathology and early cancer screening. The model proves that multi-level feature fusion is an effective approach to improving diagnostic results because it has exceeded its constituent parts. Future studies can expand upon this study by adding the element of attention to their work to enhance interpretation, experimenting with more advanced forms of GNNs like Graph Attention Networks (GATs), and testing with more data and bigger multi-institutional studies. In parallel, this paper is a valuable contribution that can be made in the direction of AI-powered precision oncology, in which precise early diagnosis and detection can greatly enhance patient outcomes.

References

1. Ashok, A., Hitesh, O., Naidu, G. P., Abhinav, B., Krishna, C. P. V., & Nair, M. (2024). An Integrated Study on Convolutional Neural Networks and Graph Neural Networks for Brain Tumor Classification from MRI Images. ACM. <https://doi.org/10.1145/3675888.3676095>
2. Bandi, P., Geessink, O., Manson, Q., Van Dijk, M., Balkenhol, M., Hermesen, M., ... & van der Laak, J. (2019). From detection of individual metastases to classification of lymph node status at the patient level: The CAMELYON17 challenge. *IEEE Transactions on Medical Imaging*, 38(2), 550–560. <https://doi.org/10.1109/TMI.2018.2867350>
3. Bejnordi, B. E., Veta, M., van Diest, P. J., van Ginneken, B., Karssemeijer, N., Litjens, G., ... & van der Laak, J. A. (2017). Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in breast cancer. *JAMA*, 318(22), 2199–2210. <https://doi.org/10.1001/jama.2017.14585>
4. Dhiman, G., Juneja, S., Viriyasitavat, W., Mohafez, H., Hadizadeh, M., Islam, M. A., & Gulati, K. (2022). A Novel Machine-Learning-Based Hybrid CNN Model for Tumor Identification in Medical Image Processing. *MDPI Sustainability*, 14(3), 1447. <https://doi.org/10.3390/su14031447>
5. Grand Challenge. (2025). CAMELYON17 Dataset. Retrieved from <https://camelyon17.grand-challenge.org/Data/>
6. Hamdi, M., Senan, E. M., Jadhav, M. E., Olayah, F., Awaji, B., & Alalayah, K. M. (2023). Hybrid Models Based on Fusion Features of a CNN and Handcrafted Features for Accurate Histopathological Image Analysis for Diagnosing Malignant Lymphomas. *MDPI Diagnostics*, 13(13), 2258. <https://doi.org/10.3390/diagnostics13132258>
7. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60–88. <https://doi.org/10.1016/j.media.2017.07.005>
8. Litjens, G., Sánchez, C. I., Timofeeva, N., Hermesen, M., Nagtegaal, I., Kovacs, I., ... & van der Laak, J. (2018). Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Scientific Reports*, 8(1), 5537. <https://doi.org/10.1038/s41598-018-22811-z>
9. Macenko, M., Niethammer, M., Marron, J. S., Borland, D., Woosley, J. T., Guan, X., & Thomas, N. E. (2009). A method for normalizing histology slides for quantitative analysis. 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 1107–1110. <https://doi.org/10.1109/ISBI.2009.5193250>
10. Nusrat, J. N., Sultana, F., & Anika, S. (2024). Skin Cancer Detection: Leveraging Hybrid Deep Learning Models and Traditional Machine Learning Classifiers. *ACM ICCA 2024*. <https://doi.org/10.1145/3723178.3723280>
11. Nwakeze, O. M., Okeke, O. C., & Mgbefulike, I. J. (2025). Intelligent robotic object grasping system using computer vision and deep reinforcement learning techniques. *International Journal of Science and Research Archive*, 14(3), 511–521.
12. Patel, N. D., Jain, V. K., Yadav, A. K., Bano, S., & Rajpoot, D. S. (2024). Brain Tumor Detection from MRI Images Using Convolutional Neural Networks. *ACM IC3 2024*. <https://doi.org/10.1145/3675888.3676039>
13. Prome, R., Adepoju, A., & Nwakaeze, O. M. (2024). Global cancer burden and the role of early detection: A review of diagnostic innovations. *Elsevier Journal of Cancer Informatics*, 38(1), 45–59. <https://doi.org/10.1016/j.jcancerin.2024.01.005>
14. Sureshkumar, V., Prasad, R. S. N., Balasubramaniam, S., Jagannathan, D., Daniel, J., & Dhanasekaran, S. (2024). Breast Cancer Detection and Analytics Using Hybrid CNN and Extreme Learning Machine. *MDPI Journal of Personalized Medicine*, 14(8), 792. <https://doi.org/10.3390/jpm14080792>
15. Wang, X., Zhang, Y., & Liu, Y. (2023). Benchmarking deep learning models on CAMELYON datasets for breast cancer metastasis detection. *MDPI Cancers*, 15(4), 1123. <https://doi.org/10.3390/cancers15041123>