G-027

Fractal Graph Convolutional Network for Histopathological image classification

シヤ　シヨウキヨウ‡　　鎌田　清一郎‡　　ユウ　リンクン‡　　カク　ヨクション‡
XIANGJIANG　XIE　　Seiichiro Kamata　　LINKUN　YU　　YIJUN　GUO

Abstract

Histopathological image classification plays a crucial role in computational pathology, aiding disease diagnosis based on microscopic tissue morphology. With the growing prevalence of digital pathology and high-resolution histological imaging, there is a rising demand for automated, robust, and interpretable classification models.

While traditional convolutional neural networks (CNNs) have achieved success in natural image analysis, they often struggle to capture the complex spatial relationships and hierarchical organization present in histopathological images. Recent advances using Graph Convolutional Networks (GCNs) address this limitation by modeling cellular interactions as graph structures. However, many existing GCN-based methods rely on complex pipelines and lack flexibility in modeling multi-path feature hierarchies and diverse spatial fields-of-view.

We propose a lightweight Fractal Graph Convolutional Network (Fractal-GCN) tailored for histopathological image classification. Our model constructs k-nearest neighbor graphs at multiple magnification levels and performs fractal convolutions across multiple paths with varying depths, capturing both local and global spatial features. To enhance efficiency and generalization, we incorporate path-level dropout, soft supervisory signals derived from path behavior, and a gating mechanism for adaptive path fusion.

Experiments on the public Databiox dataset demonstrate that our method achieves competitive classification performance compared to standard GCN baselines, while significantly improving model efficiency and interpretability. The modular design of Fractal-GCN offers promising extensibility toward adaptive graph structure learning in future medical imaging tasks.

1. Introduction

Histopathological image classification is a key task in computer-aided diagnosis but remains challenging due to irregular cell structures and hierarchical tissue organization across multiple magnifications [1][5]. Traditional convolutional neural networks (CNNs) struggle to model long-range dependencies in such data due to their grid-based nature. Graph Convolutional Networks (GCNs) address this by modeling nuclei as nodes and spatial relationships as edges, but most existing models rely on fixed-depth, single-scale representations [2].

Figure 1 shows diverse tissue structures, and Figure 2 illustrates how low- and high-magnification H&E images reveal complementary information. These motivate the need for a scale-aware and efficient classification model.

We propose a lightweight Fractal Graph Convolutional Network (FGCN) that decouples graph construction from training, allowing flexible preprocessing [2] [6]. FGCN applies parallel graph convolutions of different depths to capture multi-level spatial features, integrates multi-magnification data via a gating-based fusion strategy, and employs path-level dropout for regularization.

Experiments on the Databiox dataset show that FGCN achieves competitive accuracy while reducing computational cost and improving interpretability [8]. Its modular design supports real-world deployment and future extensions in adaptive graph learning.

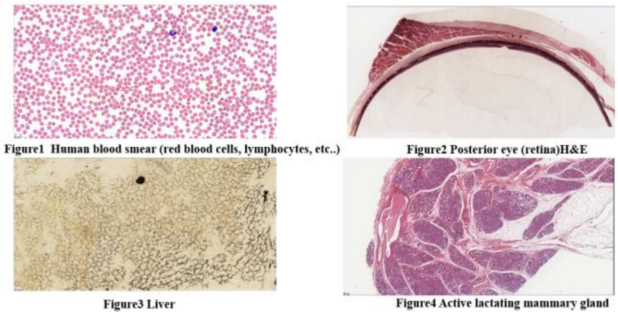

Figure 1. Histological examples of diverse tissue types. Images from the University of Michigan Virtual Histology Slide Collection[13].

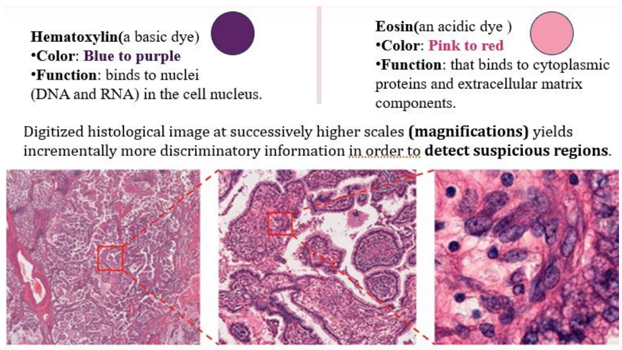

Figure 2. Multi-scale visualization of H&E-stained histopathological tissue. Hematoxylin (blue/purple) binds to nuclear structures, while eosin (pink/red) stains cytoplasmic components. Higher magnification levels reveal more detailed spatial patterns critical for accurate diagnosis. Adapted from [14].

2. Related Work

2.1 Histopathological Image Classification

Early studies employed convolutional neural networks (CNNs) to extract texture and morphological features from H&E-stained slides. For example, Ding et al. [2] reported the performance of ResNet21 and ResNet22 on the Databiox dataset, with accuracies of 77.46±3.63% and 77.73±4.26%, respectively. These implementations were based on prior studies utilizing ResNet architectures and the Databiox dataset [9][10]. More advanced attention-based multiple instance learning (MIL) methods later incorporated patch-level context to improve robustness [11]. However, these models still rely on regular image grids and often struggle to capture the irregular and hierarchical nature of cellular arrangements in histopathology.

2.2 Graph Neural Networks

Graph-based approaches have gained increasing attention for histopathological image analysis due to their ability to represent spatially irregular structures. In these models, nuclei or cells are treated as graph nodes, with edges defined based on spatial proximity or morphological similarity. GraphSAGE demonstrated the effectiveness of inductive graph learning in medical contexts such as lung tissue classification [12].

More recently, Ding et al. [2] proposed a fractal graph convolutional network (CNN-FGCN) that integrates multi-path GCN layers with a MLP-Mixer-based feature fusion unit (MMFFU), achieving 83.05 ± 1.95% accuracy on a histopathology dataset [2]. While effective, this method tightly couples graph construction with end-to-end CNN pipelines and relies on a computationally intensive fusion module, limiting interpretability and deployment flexibility.

2.3 Fractal Architectures

Fractal architectures employ self-similar multi-path designs to capture features at various depths and scales. FractalNet showed that such structures can enhance representational capacity in CNNs without significantly increasing model size [8]. This principle was later extended to GCNs by Ding et al., who introduced fractal paths into a graph-based framework and combined them with an MLP-Mixer-based fusion module for histopathological classification [2][6]. However, their approach relies on fixed-stage fusion and lacks flexibility in capturing multi-resolution spatial hierarchies.

In contrast, our method emphasizes scale-awareness—the ability to integrate information across multiple fields-of-view—by explicitly constructing graphs at different magnification levels and adaptively fusing multi-depth paths through a gating mechanism. Moreover, we replace the heavy MLP-Mixer fusion with a lightweight supervisory-guided strategy and path dropout, yielding a more computationally efficient yet expressive design suited for large-scale digital pathology workflows.

3. Proposed Method

3.1 Graph Construction

To construct input graphs for our model, we adopt a modular pipeline that separates image processing from graph learning. Specifically, we use a pre-trained CNN-based nuclei extractor to identify cell locations from histopathological images. This detector outputs a set of nuclei coordinates, denoted as$\{(x_i, y_i)\}_{i=1}^{N}$, where each (x_i, y_i) corresponds to the center of a detected nucleus.

Each detected nucleus is treated as a graph node, and spatial relationships are modeled by connecting each node to its k-nearest neighbors based on Euclidean distance. This results in an undirected graph G=(V,E), where V represents the set of nuclei and E encodes edges defined by the k-NN algorithm. The constructed graph captures local spatial dependencies among nearby cells and serves as the input to our FGCN model.

Importantly, this graph construction process is applied independently to images at different magnification levels (e.g., 10x and 40x), allowing the model to capture structural patterns across spatial scales. By decoupling graph construction from training, our approach ensures flexibility and interpretability, while avoiding the added complexity of embedding graph generation within an end-to-end neural network.

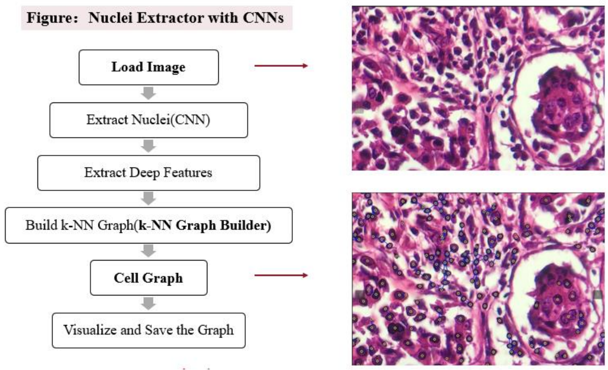

Figure3. Nuclei Extractor with CNNs

3.2 Fractal GCN Design

To efficiently model spatial features at multiple scales, we design a lightweight fractal graph convolutional network that integrates multi-depth path design, dynamic path fusion, and a path-level dropout mechanism. This structure enhances representation while maintaining low computational cost and strong generalization ability[2].

3.2.1 Multi-depth Path Design.

We construct multiple parallel paths, each with a different number of graph convolution layers. Shallow paths are used to extract local features, while deeper paths capture more global context. This allows the model to learn spatial relationships across multiple receptive fields within the tissue graph.

3.2.2 Dynamic Path Fusion.

To combine the outputs from different paths, We assign a learnable weight to each path and compute a weighted sum of their outputs. Let M be the number of paths and $H^{(i)}$ the output of the i_th path. The fused representation is calculated as:

$$H_{fused} = \sum_{i=1}^{M} \omega_i . H^{(i)}$$

The weights wiw_iwi are trained jointly with the network and normalized to ensure numerical stability. This mechanism allows the model to adaptively emphasize more informative paths during learning.

3.2.3 Path Dropout.

During training, we apply a path-level dropout strategy to reduce redundancy and improve robustness. A binary mask $m_i \in \{0,1\}$ is sampled for each path, and the fusion becomes:

$$H_{fused} = \sum_{i=1}^{M} m_i \cdot \omega_i \cdot H^{(i)}$$

Paths with $m_i = 0$ are disabled in that iteration. During inference, all paths are active. This regularization technique reduces overfitting and encourages diversity across paths.

Through this design, we aim to retain the expressive power of fractal architectures while improving efficiency and interpretability. The combination of multi-scale representation, adaptive fusion, and structured regularization makes this module both effective and lightweight.

3.3 Scale-Aware Path Gating with Soft Supervision

To enable adaptive fusion across different graph convolutional depths, we introduce a scale-aware gating mechanism guided by soft supervision.

During training, all fractal paths are retained by disabling path dropout. Each path is appended with a lightweight classifier to output a soft prediction signal. These soft labels reflect the path's individual performance and serve as supervision to train the gating module.

The gating module receives these path-wise performance signals and learns to assign sample-dependent weights for dynamic feature fusion. The final representation is computed as a weighted sum of all path outputs and then passed to the main classifier.

Although no explicit scale encoding is introduced, each path inherently captures features at different receptive field scales. Through this process, the model learns to selectively emphasize the most informative scale for each input—effectively achieving implicit scale-awareness through behavior-driven fusion.

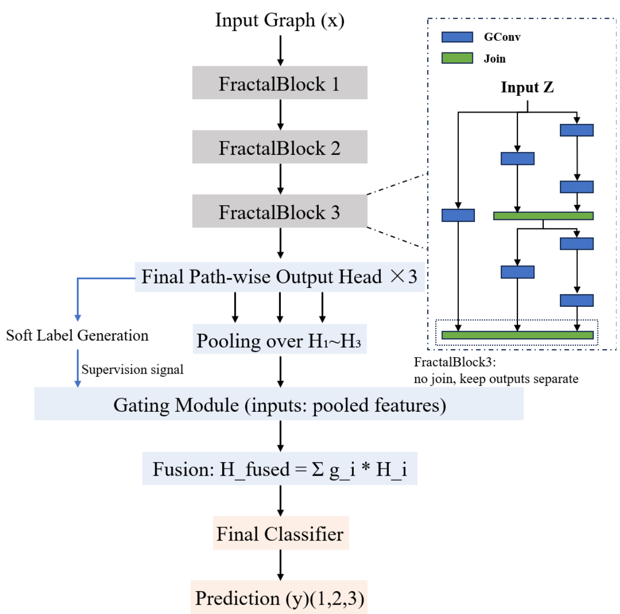

Figure4. Figure X: Overview of the proposed multi-path Fractal GCN architecture with path-level supervision and gating fusion.

4. Experiments and Results

4.1 Dataset

I evaluate my model on the Databiox dataset, which contains histological image patches categorized into three distinct classes. The dataset includes over 600 training samples and approximately 60 testing samples[8].

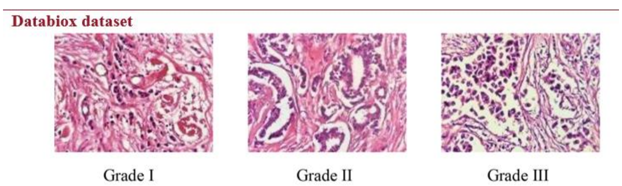

Figure5. Sample Histopathological Image from the Databiox IDC Dataset[2][8]

4.2 Implementation Details

To ensure robust performance evaluation, I adopt a 10-fold cross-validation strategy. The model is trained using the standard cross-entropy loss, which effectively measures the discrepancy between predicted and true class distributions in multi-class settings.

All experiments were conducted in a controlled environment featuring Pytorch 2.5.1 and DGL 2.1.0, executed on an NVIDIA RTX 3090 GPU. Key hyperparameters include a learning rate of 0.005, dropout rate of 0.1, batch size of 32, and a training duration of 40 epochs. These settings were selected based on empirical tuning, including adjustments to ensure stability and convergence.

The evaluation metrics include accuracy and F1-score, providing a comprehensive assessment of the model's performance.

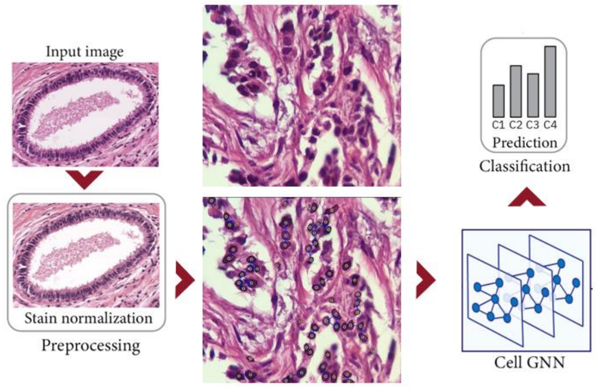

Figure 6. Input-output flow and general framework of the proposed approach.[8]

Hardware Configuration

CPU	Intel Core i7-10700F @ 2.90GHz
GPU	NVIDIA GeForce RTX 3090 (24GB VRAM)
GPU Driver	535.171.04
Memory	16 GB

Software Environment(Key Dependencies)

Operating System	Ubuntu 20.04.6 LTS
Kernel Version	Linux 5.15.0-105-generic
Python	3.9.0
CUDA Version	12.2
Pytorch	2.5.1
torchvision	0.20.1
DGL	2.1.0
NumPy	1.25.0

Figure7. Overview of the Experimental Environment and Training Parameters

4.3 Results

Our FGCN achieves an average classification accuracy of 79.0% and a precision of 79.3%, outperforming baseline models including ResNet21 [10], ResNet22, and a standard GCN [2][3].

As shown in Table 2, our method improves upon the best baseline (GCN) by a margin of 1.3% in accuracy.

Furthermore, Ablation studies demonstrate that both the fractal architecture and the scale-aware path gating mechanism contribute significantly, as removing either component results in a noticeable performance drop.

Model	Accuracy (%)
ResNet21	72.52
ResNet22	74.77
GCN	77.68
Ours	**79.02**

Figure 8. Classification accuracy of different models

5. Discussion

The proposed FGCN framework offers significant advantages in capturing complex spatial dependencies at multiple scales, an essential feature in histopathological analysis. Its interpretability aligns well with the clinical workflow of pathologists. Nonetheless, the increased model complexity due to fractal branching and multi-scale processing leads to greater computational demands and memory usage. Future work will explore attention-based or region-aware fusion strategies to mitigate this overhead, while better capturing the nested spatial structures often overlooked by conventional feature fusion approaches.

6. Conclusion

In this work, we developed a graph-based classification framework that integrates a fractal GCN design with multi-scale feature anchoring and fusion for histopathological image analysis. By incorporating multi-depth path structures, dynamic path fusion, and path-level dropout, the model effectively captures spatial features across different receptive fields while maintaining computational efficiency.

To further enhance contextual representation, I introduced a scale-aware path gating mechanism that adaptively combines features from different fractal branches. In addition, a cross-magnification fusion strategy aligns low- and high-resolution graphs through spatial anchoring, enabling the integration of global and local tissue information in a biologically meaningful way.

The results demonstrate that combining hierarchical graph modeling with structured, scale-aware fusion improves both classification accuracy and model interpretability.

In future work, I plan to explore adaptive graph construction and extend this framework to other domains involving complex spatial hierarchies. I also aim to enhance the interpretability of the model—an essential aspect in medical applications—by integrating tools such as Grad-CAM and GNNExplainer, as well as designing multi-scale visualization techniques tailored to histopathological images, including heatmap overlays and case-based evidence. These enhancements will help bridge the gap between automated predictions and clinical trust.

Acknowledgments

I would like to express my sincere gratitude to the Kamata Laboratory at Waseda University, especially to Professor Kamata, for his continuous guidance and encouragement throughout this research. I also thank my lab members and friends for their helpful discussions and support during the development and evaluation of this work. I am grateful to Waseda University and the Graduate School of Information, Production and Systems for providing a stimulating academic environment and the necessary resources. Lastly, I acknowledge my own efforts in designing, implementing, and validating the proposed model, and I remain committed to further advancing this research in the future.

References

[1] Ahmedt-Aristizabal, D., Armin, M. A., Denman, S., Fookes, C., & Petersson, L. (2022). A survey on graph-based deep learning for computational histopathology. In *Computerized Medical Imaging and Graphics* (Vol. 95). Elsevier Ltd. https://doi.org/10.1016/j.compmedimag.2021.102027

[2] Ding, S., Gao, Z., Wang, J., Lu, M., & Shi, J. (2023). Fractal graph convolutional network with MLP-mixer based multi-path feature fusion for classification of histopathological images. *Expert Systems with Applications*, *212*. https://doi.org/10.1016/j.eswa.2022.118793

[3] Gao, Z., Lu, Z., Wang, J., Ying, S., & Shi, J. (2022). A Convolutional Neural Network and Graph Convolutional Network Based Framework for Classification of Breast Histopathological Images. *IEEE Journal of Biomedical and Health Informatics*, *26*(7), 3163–3173. https://doi.org/10.1109/JBHI.2022.3153671

[4] Gour, M., Jain, S., & Sunil Kumar, T. (2020). Residual learning based CNN for breast cancer histopathological image classification. *International Journal of Imaging Systems and Technology*, *30*(3), 621–635. https://doi.org/10.1002/ima.22403

[5] 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. https://doi.org/10.1109/RBME.2009.2034865

[6] Larsson, G., Maire, M., & Shakhnarovich, G. (2016). *FractalNet: Ultra-Deep Neural Networks without Residuals*. http://arxiv.org/abs/1605.07648

[7] Zhao, Y., Zhang, J., Hu, D., Qu, H., Tian, Y., & Cui, X. (2022). Application of Deep Learning in Histopathology Images of Breast Cancer: A Review. In *Micromachines* (Vol. 13, Issue 12). MDPI. https://doi.org/10.3390/mi13122197

[8] Araujo, T., Aresta, G., Castro, E., Rouco, J., Aguiar, P., Eloy, C., Polonia, A., & Campilho, A. (2017). Classification of breast cancer histology images using convolutional neural networks. *PLoS ONE*, *12*(6). https://doi.org/10.1371/journal.pone.0177544

[9] Computer Vision and Pattern Recognition Workshops, 2009. CVPR Workshops 2009. IEEE Computer Society Conference on : [Fontainebleau Resort Miami Beach, Florida : June 20-25, 2009]. (2009). IEEE.

[10] He, K., Zhang, X., Ren, S., & Sun, J. (n.d.). *Deep Residual Learning for Image Recognition*. http://image-net.org/challenges/LSVRC/2015/

[11] Wang, X., Chen, H., Gan, C., Lin, H., Dou, Q., Tsougenis, E., Huang, Q., Cai, M., & Heng, P. A. (2020). Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. *IEEE Transactions on Cybernetics*, *50*(9), 3950–3962. https://doi.org/10.1109/TCYB.2019.2935141

[12] Hamilton, W. L., Ying, R., & Leskovec, J. (2017). *Inductive Representation Learning on Large Graphs*. http://arxiv.org/abs/1706.02216

[13] University of Michigan Medical School. Virtual Histology Slide Collection. https://histologyslides.med.umich.edu/Histology/

[14] He, W., Liu, T., Han, Y., Ming, W., Du, J., Liu, Y., Yang, Y., Wang, L., Jiang, Z., Wang, Y., Yuan, J., & Cao, C. (2022). A review: The detection of cancer cells in histopathology based on machine vision. *Computers in Biology and Medicine*, *146*, 105636. https://doi.org/10.1016/J.COMPBIOMED.2022.105636