POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Executive Summary of the Thesis

Convolution on Triangular Motifs: A GNN Approach to Edge Classification in Cell-Graphs

Laurea Magistrale in CompuTer Science and Engineering

Author: Andrea Camilloni

Advisor: Prof. Daniele Loiacono

Co-advisors: Karl Meinke, Rachael Sugars

Academic year: 2022-2023

1. Introduction

In an era where digital technology is fundamentally transforming healthcare, digital pathology (DP) finds itself at a pivotal juncture. While significant strides have been made in imaging and computational techniques, critical challenges persist. The healthcare sector faces an urgent need to automate and streamline the process of identifying key histological features essential for accurate medical diagnosis and treatment. One such pressing issue is the timely and reliable identification of the basement membrane (BM) in cases involving complications of chronic graft-versus-host disease (cGVHD) following hematopoietic stem cell transplantation (HSCT) [28]. Traditional methods, reliant on manual annotation in whole-slide images (WSIs), suffer from being both labor-intensive and susceptible to human error [28]. The need for innovative, efficient, and error-minimizing techniques in histological examinations has never been more critical.

In light of these challenges, machine learning methods, specifically Graph Neural Networks (GNNs), have emerged as promising avenues for innovation. These networks have been employed to automate and refine histological analyses, particularly through cell-graph models which encapsulate complex cellular interactions and higher-level histological relationships [1, 13, 21]. Although GNNs are proficient in managing these cell-graphs to identify and classify intricate histological features like the BM [13, 21], they sometimes fail in capturing higher-level structural details within graphs. This is a domain where Convolutional Neural Networks (CNNs) excel, due to their capacity for intricate pattern recognition. Therefore, integrating CNN-inspired techniques into GNNs could forge an innovative pathway for enhancing graph representations and addressing current limitations.

This work aims to address this limitation by incorporating a pattern recognition component into the GNN framework, for triangle based motif learning. This enhancement not only improves the accuracy and robustness of BM predictions, but also provides a valuable tool for pathologists. It automates the identification of BM, potentially assisting in the histological grading process and facilitating more efficient and precise diagnostics in oral pathology. This study was carried out in collaboration with the KTH Royal Institute of Technology (KTH) and Karolinska Institute (KI).

Contributions of this Work This work builds upon the existing GNN framework established by [13], aiming to address some key limitations in the realm of digital pathology, specifically the identification of BM in chronic graft-versus-host disease (cGVHD). Unlike conventional edge classification or link prediction approaches that focus on a limited context of two nodes, this work presents a novel classifier that integrates convolutional layers for pattern recognition with triangle-based recurrent structures for edge classification. This results in a more context-aware model that uses triangle-based motifs for edge classification in the cell-graph, reducing false positives and enhancing the accuracy of BM identification.
Additionally, we introduce a comprehensive learning mechanism where the error is backpropagated through both the GNN and the new classifier, deviating from earlier methods where the error was only propagated within the GNN layers [13, 21]. This results in improved performance in BM identification tasks.
A graph segmentation technique is also developed to offer a refined means of evaluating the predictions of the GNN on cell-graph models, aiding in clearer interpretation and insights into the model's efficacy. This method proves particularly beneficial when assessing results from samples presenting degraded BM, where simply examining edge outcomes isn't sufficient for a visual evaluation of the model's performance.
In summary, this work substantially advances the role of GNNs in digital pathology. We introduce a context-aware, structure-oriented approach to edge classification, which improves BM identification. Additionally, our graph segmentation technique refines the evaluation process, offering clearer insights into the model's effectiveness—especially when dealing with degraded samples. Collectively, these innovations contribute to the broader advancement of the field.

2. Literature Review

The digital pathology (DP) domain has made significant strides with the introduction and adoption of machine learning (ML) techniques, particularly deep learning. Despite these advancements, traditional Convolutional Neural Networks (CNNs) and Deep Neural Networks (DNNs) are often insufficient in capturing complex histological dependencies in tissue images. The current review aims to discuss the state-of-the-art methods, emphasizing deep learning approaches and Graph Neural Networks (GNNs) in DP. It also identifies gaps and opportunities for further research [1].

Deep Learning in DP Deep learning methods have been extensively applied to DP tasks, ranging from cellular to region-wide scales [9]. Cellular-level applications include semantic segmentation, detection, and mitosis detection, using architectures like CNNs, Fully Connected Networks (FCNs), and Stacked Autoencoders (SAEs) [6, 8, 24, 25]. Gland and region-level analyses typically use CNNs and FCNs, often with additional post-processing for gland segmentation and contour detection [4, 31, 34, 35]. Despite their efficacy, these methods primarily depend on pixel-level information, lacking the ability to capture higher-level histological relationships, indicating room for improvement.

GNNs in DP GNNs are emerging as effective tools for capturing complex dependencies in DP [2, 7, 10, 14, 15, 26, 27, 38]. While most studies have focused on classification tasks, only a few explore frameworks for Region-of-Interest (ROI) retrieval and segmentation, providing an avenue for further research [23, 37].

BM Identification BM identification is a crucial task in DP that remains challenging. Existing methods, such as those by Wang et al. [29], Wu et al. [32], and Cao et al. [5], rely on pixel-level information, limiting their ability to capture topological and histological relationships. GNN-based approaches to BM identification have been introduced [13, 21], but they also present limitations, particularly in leveraging triangular motifs in cell-graphs built using Delaunay triangulation. Therefore, there is room for further improvement in GNN-based BM identification approaches.

Edge Classification in GNNs The edge classification task in GNNs has not been adequately explored. While GNN frameworks like MPNNs, EGNN, and NENN focus on integrating edge features for node feature

aggregation [11, 12, 33, 36], their applicability to edge classification remains an open question. Notably, NENN [36] introduces a hierarchical dual-level attention mechanism that alternately stacks node-level and edge-level attention layers to learn and aggregate embeddings for nodes and edges, allowing the node and edge embeddings to mutually reinforce each other. Despite the ability to compute higher order edge representations, NENN has not been applied to any edge classification task.

While deep learning methods, particularly CNNs and GNNs, have made significant contributions to DP, gaps remain in capturing higher-level histological dependencies. Specifically, there is a need for more robust methodologies that can incorporate complex spatial relationships for tasks like BM identification and edge classification.

3. Dataset

The dataset, sourced from the Department of Dental Medicine at Karolinska Institutet (KI), serves as the primary foundation of this study. This dataset comprises Whole Slide Images (WSIs) of hematoxylin and eosin (H&E) stained oral mucosa tissue from patients who had received HSCT and healthy volunteers. The WSIs were cropped into 2000×2000 pixel regions, and approximately 62000 cells have been manually annotated and classified. Additional annotations were performed to identify the basement membrane (BM).

Cell-graph model We developed a cell-graph model using Delaunay triangulation [19], where cells are nodes and edges are classified based on their interaction with the BM as either crossing or not crossing it. Node features include annotated and computed metrics like cell type and density, as well as deep learning features from a pre-trained ResNet-18 model. Edge features such as node distance and BM crossing were also incorporated and normalized (see Appendix for further details).

4. Methods

In this section, we delve into the GNN framework, but first, some foundational graph theory is defined to set the stage. A graph G is

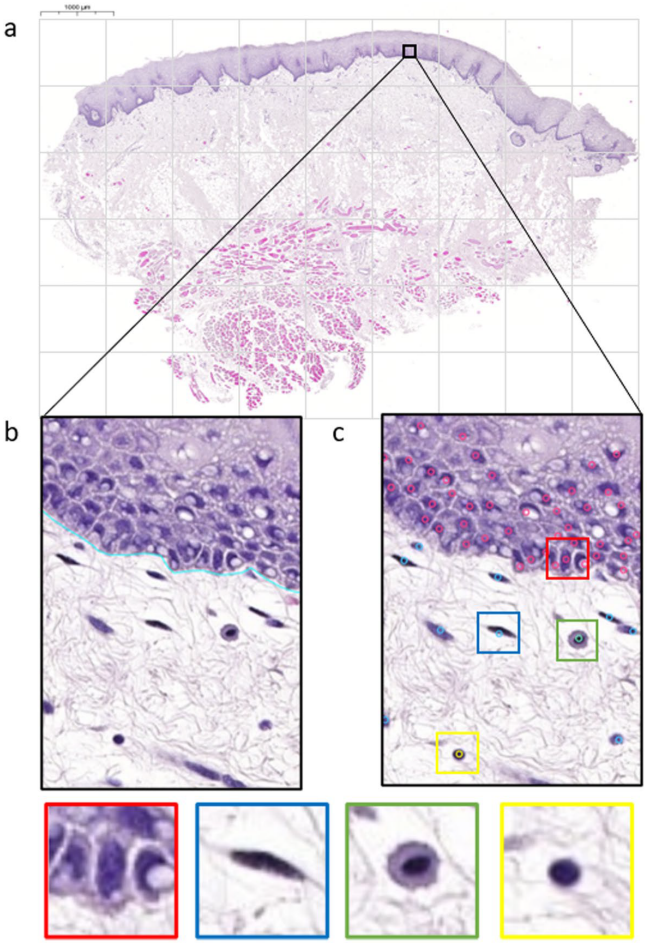


Figure 1: (**a**) WSI of healthy buccal oral mucosa that has been segmented into tiles of 2000 × 2000 pixels. (**b**) Annotation of the extension of the BM visualised with blue line. (**c**) Nuclei centroids were annotated and labelled as either epithelial (red), fibroblast or endothelial (blue), inflammatory (green) or lymphocytic (yellow) (Image taken from [21])

mathematically expressed as $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of n nodes and $\mathcal{E}$ is the set of edges connecting them. The adjacency matrix $A \in \mathbb{R}^{n \times n}$ encapsulates the node relationships, and for weighted graphs, its entries can be arbitrary real values. Graphs can also feature node-level and edge-level attributes represented by $X \in \mathbb{R}^{n \times d}$ and $E \in \mathbb{R}^{n \times n \times p}$, respectively. With this background, the GNN framework presented here comprises two primary components: Node Embedding Layers and an Edge Classifier.

4.1. Node Embedding Layers: EAGNN

The proposed architecture builds upon the node embedding layers proposed in [13], EAGNN. Unlike traditional GNN layers, EAGNN aggregates multiple edge features into node embeddings and follows a unique computational approach defined by equations $EAgg^k$ and H^k (see Appendix). The model aims to derive more informative latent node representations from a graph, considering both node and edge features (see Appendix for further details).

4.2. Edge Classifier: TM-CNN

Contrasting with traditional approaches that employ MLPs or multiplication operators [13, 21], the Triangular Motifs Convolutional Neural Network (TM-CNN) is proposed as a uniquely tailored edge classifier, designed specifically for discerning meaningful patterns pertinent to edge classification tasks.

The TM-CNN framework for edge classification is inspired by methodologies commonly used in pattern learning tasks such as Text Classification [16, 20, 30] and Financial Time-Series Classification [18]. These tasks often involve generating text or time-series embeddings and passing them through 1D-Convolutional layers. This approach allows for the extraction of local and global features, which are crucial for accurate classification. Similarly, in this work, the TM-CNN classifier aims to extract and learn meaningful patterns from the motif structures for effective edge classification.

The TM-CNN classifier is designed to take as input the two triangular motifs associated with an edge (u, v). These triangular motifs are formed by the nodes (u, v, z) and (u, v, w), where u, v, z, and w represent the node embeddings learned after K-iteration by the EAGNN architecture. Each node embedding has a size of $\mathbb{R}^{d_K}$, resulting in an input to the TM-CNN classifier of size $d_k \times 4$ (corresponding to 4 channels).

The TM-CNN architecture consists of two 1D convolutional layers and two fully connected (FC) layers.

In the proposed architecture, two convolutional layers with n_{f_1} and n_{f_2} filters generate feature maps F_1 and F_2, both of dimensions $(d_k \times n_{f_i})$. These maps pass through ReLU activations and are followed by two Fully Connected (FC) layers. The final FC layer outputs a scalar between 0 and 1, serving as the edge classification for (u, v). Figure 2 gives an overview of the proposed architecture.

5. Experiments and Results

The study utilized the generated cell-graph dataset, which was divided into training (70%), testing (30%), and validation (15% of training) subsets (see Table 2). The models EAGNN$_i$[1] and TM-CNN were trained for 100 epochs using backpropagation with a mini-batch size of 32, balanced for data imbalances. The Adam [17] optimizer was employed with an initial learning rate of 0.001, which was reduced by 0.1 every 40 epochs. Dropout and weight-decay techniques were used to prevent overfitting. This configuration is consistent throughout the study for all experiments.

The evaluation of the model consisted of two critical phases: validation and testing.

Validation In the validation phase, different configurations of the EAGNN and TM-CNN model were fine-tuned on a series of metrics including precision, recall, F1 score[2], ROC-AUC, and accuracy. The model with five layers in the EAGNN backbone demonstrated the highest F1 score and accuracy, thereby chosen for the testing phase (see Appendix for further details).

Testing In the testing phase, the performance of the (EAGNN$_5$,TM-CNN) model was compared with two baseline models from existing literature. The proposed model configuration was found to surpass both the (EAGNN$_2$,BC+MLP) [13] and the (GraphSAGE,MUL) [21] models in several metrics including F1 score, ROC-AUC, and accuracy.

The (EAGNN$_5$,TM-CNN) model demonstrated a significant improvement over the primary benchmark (EAGNN$_2$,BC+MLP) [13] model, showing an F1 score that is approximately 2% higher. The model maintained similar levels of precision but had superior recall, resulting in fewer false negatives. Results are shown in Table 1.

[1] i stands for the number of aggregation layers used in the configuration of EAGNN$_i$.

[2] Precision, recall and F1 score are computed w.r.t. the BM crossing edges.

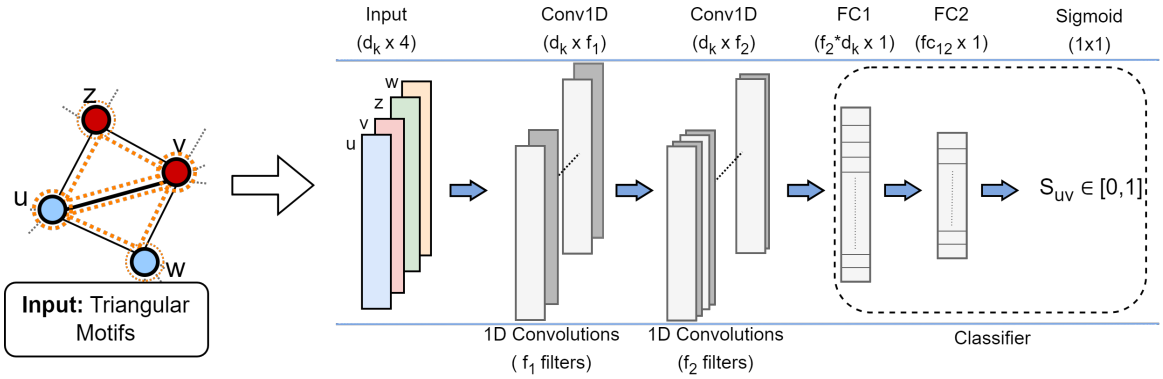


Figure 2: Triangular Motifs CNN (TM-CNN) architecture overview.

Model	Classifier	NF	EF	Precision	Recall	F1	ROC-AUC	Accuracy
[21] GraphSAGE	MUL	N_1	-	0.3417	0.7011	0.4594	-	0.9071
[13] EAGNN$_2$	BC+MLP	N_1	E_{1234}	**0.8548**	0.8427	0.8395	0.9799	0.9822
EAGNN$_2$	TM-CNN	N_1	E_{1234}	0.8523	0.8595	0.8559	0.9866	0.9837
EAGNN$_5$	TM-CNN	N_1	E_{1234}	0.8526	**0.8657**	**0.8591**	**0.9874**	**0.9840**

Table 1: Summary of test results comparing the proposed models with existing baseline models. NF: Node Features, EF: Edge Features.

5.1. Qualitative results

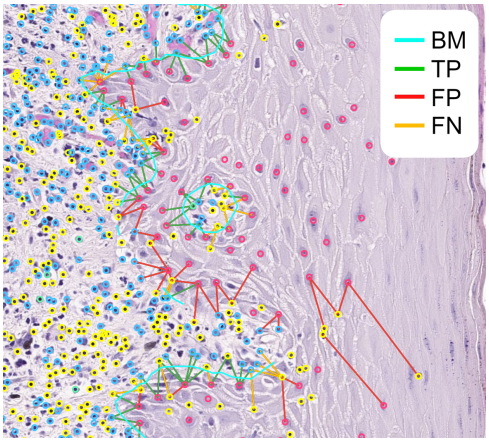
In this summary, we focus on the qualitative evaluation of the (EAGNN$_5$,TM-CNN) model against its predecessor, (EAGNN$_2$,BC+MLP) [13]. Both models demonstrate similar performance in classifying BM crossing edges in healthy samples, while when considering sample presenting degraded BM, the (EAGNN$_5$,TM-CNN) model excels, as depicted in Figure 3.
The newer (EAGNN$_5$,TM-CNN) model improves the F1 score by approximately 10% in challenging conditions (degraded BM), and reduces false positive rates, substantiating its robustness and applicability in real-world scenarios.

5.2. Graph segmentation

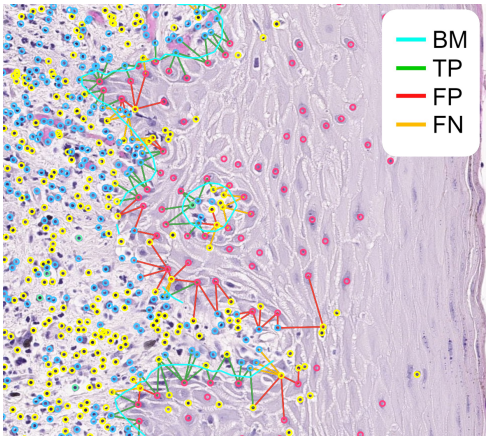
To offer a more intuitive understanding of the model's predictions, we employ color-coded segmentation based on the number of BM crossing edges within each triangular motif in the graph. The categories are as follows:
(Empty) Triangles with 0 BM crossing edges
(Red) Triangles with only 1 BM crossing edge
(Yellow) Triangles with 2 BM crossing edges
(Green) Triangles with 3 BM crossing edges
This color-coded scheme serves as a crucial tool in visually evaluating the efficacy of our proposed (EAGNN$_5$,TM-CNN) model, especially when compared to the (EAGNN$_2$,BC+MLP) [13] model. Red triangles indicate potential areas of BM disruption, while yellow and green triangles suggest regions where the BM structure is likely intact.
The (EAGNN$_5$,TM-CNN) model shows superior performance in segmenting degraded BM samples, as illustrated in Figure 4. Compared to the (EAGNN$_2$,BC+MLP) [13] model, the newer (EAGNN$_5$,TM-CNN) model exhibits a marked reduction in noisy predictions. This enhancement significantly clarifies the segmentation, particularly in the epithelial layer.
Moreover, the (EAGNN$_5$,TM-CNN) model reduces the number of red triangles, which indicate potential BM disruptions. This more accurate depiction of BM condition stands in contrast to the results from the (EAGNN$_2$,BC+MLP) [13] model, which often misclassifies regions and shows a loss of BM information.

6. Conclusion and Future Works

This research introduced a groundbreaking method for identifying and localizing basement membranes (BMs) in oral tissue samples, thereby tackling a significant challenge



(a) (EAGNN$_2$,BC+MLP)[13] F1 score: 0.5329



(b) (EAGNN$_5$,TM-CNN) F1 score: 0.6340

Figure 3: Comparisons for degraded BM sample.

in computational pathology. The proposed model, (EAGNN$_5$,TM-CNN), surpasses existing techniques by incorporating edge-related motifs, which has not only improved the model's accuracy but also enhanced its interpretability, making it more applicable in real-world scenarios.

When compared to the (EAGNN$_2$,BC+MLP) [13] model suggested by Nair et al., the proposed model exhibited significant improvements, particularly in identifying degraded BMs. This advancement is mainly due to the extended training at both the graph neural network (GNN) level and the classifier level, as well as the inclusion of convolutional layers for capturing recurrent structures.

Moreover, the enhanced interpretability offered by the novel graph segmentation technique provides pathologists with a nuanced understanding of BM conditions. This interpretability serves not primarily to streamline the diagnosis, but to offer robust decision support that can mitigate observer variance. By facilitating a deeper and more consistent understanding of the tissue's condition, our model can be a valuable tool in guiding both the assessment of disease severity and the subsequent course of treatment, potentially leading to more targeted and effective treatment strategies.

The versatility of the (EAGNN$_5$,TM-CNN) model implies its broader applicability in computational pathology, which could significantly contribute to the diagnosis and treatment of various diseases.

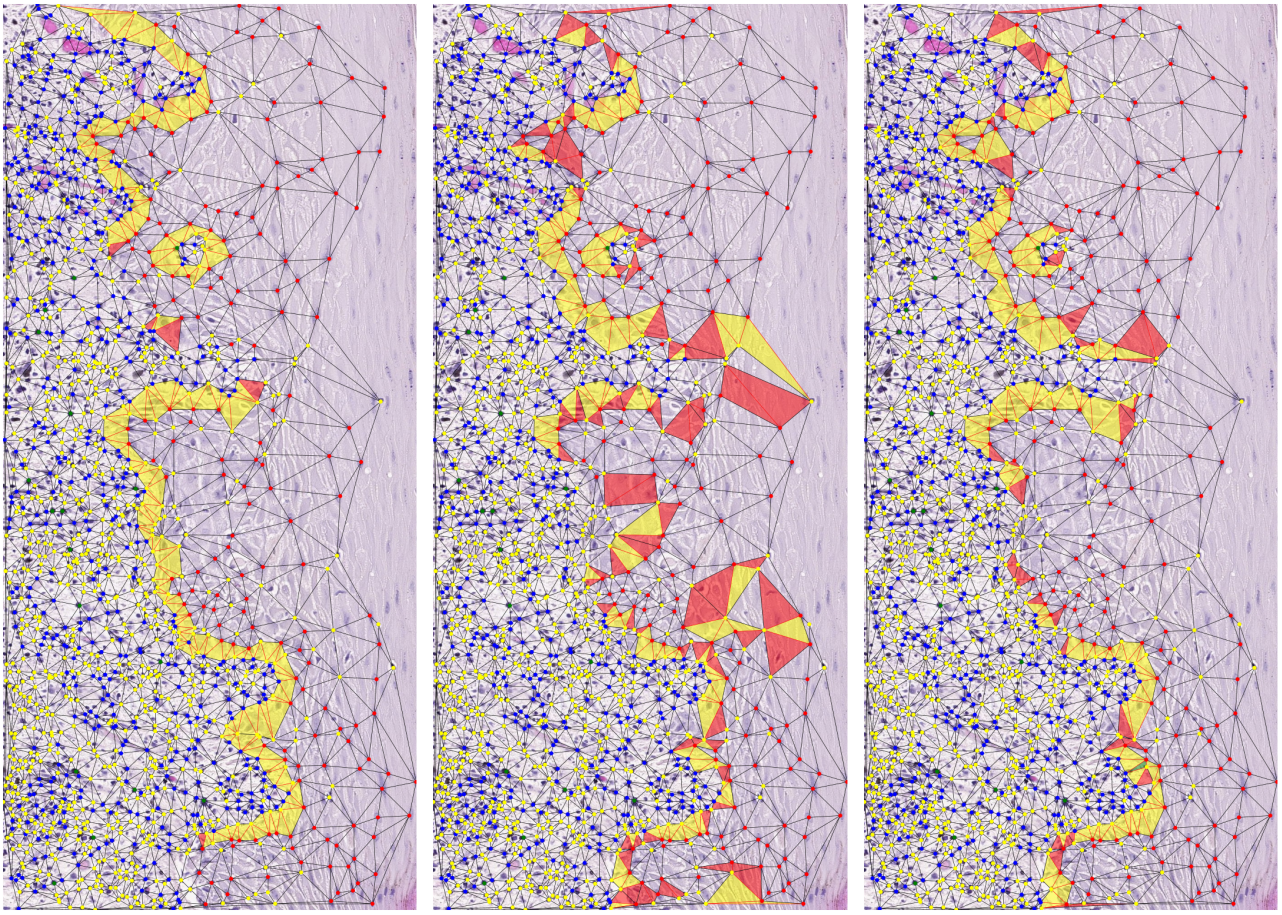
In summary, this study has made a meaningful contribution by devising an innovative model for the identification and localization of BMs in oral tissue samples. The model is not just accurate but also interpretable, increasing its utility for pathologists. Despite the promising results, further research is required to overcome limitations and to leverage the full potential of the model.

Future works Numerous avenues for future research are discernible from this study. A logical progression would be to extend the (EAGNN$_5$,TM-CNN) model to also assess the severity of BM breakages, contributing to more comprehensive disease diagnosis and treatment planning [28]. While the current study has chiefly focused on edge classification within the GNN framework, there is potential for refining node and edge representations to enhance model robustness.

Novel approaches like the dual-attention mechanism in the Node and Edge features in Graph Neural Network (NENN) framework could be integrated to potentially improve edge information update [36]. Such enhancements may offer a more effective methodology for BM identification in digital pathology. Beyond BM identification, the current model could be further generalized for various tasks such as tissue and tumor segmentation, boundary detection, region of interest retrieval, and even classification tasks.

References

[1] David Ahmedt-Aristizabal, Mohammad Ali Armin, Simon Denman, Clinton Fookes, and Lars Petersson. A survey on graph-based deep learning for computational



(a) Ground Truth (b) (EAGNN$_2$,BC+MLP)[13] (c) (EAGNN$_5$,TM-CNN)

Figure 4: Segmentation of degraded BM in oral tissue.

histopathology. *CoRR*, abs/2107.00272, 2021.

[2] Valentin Anklin, Pushpak Pati, Guillaume Jaume, Behzad Bozorgtabar, Antonio Foncubierta-Rodríguez, Jean-Philippe Thiran, Mathilde Sibony, Maria Gabrani, and Orcun Goksel. Learning whole-slide segmentation from inexact and incomplete labels using tissue graphs, 2021.

[3] James Atwood and Don Towsley. Search-convolutional neural networks. *CoRR*, abs/1511.02136, 2015.

[4] Aicha Bentaieb, Jeremy Kawahara, and Ghassan Hamarneh. Multi-loss convolutional networks for gland analysis in microscopy. pages 642–645, 04 2016.

[5] Lei Cao, YanMeng Lu, ChuangQuan Li, and Wei Yang. Automatic segmentation of pathological glomerular basement membrane in transmission electron microscopy images with random forest stacks. *Computational and mathematical methods in medicine*, 2019:1684218, 03 2019.

[6] Hao Chen, Xi Wang, and Pheng Heng. Automated mitosis detection with deep regression networks. pages 1204–1207, 04 2016.

[7] Rui J Chen, Michael Y Lu, Jason Wang, David F Williamson, Scott J Rodig, Neal I Lindeman, and Faisal Mahmood. Pathomic fusion: an integrated framework for fusing histopathology and genomic features for cancer diagnosis and prognosis. *IEEE Transactions on Medical Imaging*, 2020.

[8] Dan Ciresan, Alessandro Giusti, Luca Gambardella, and Jürgen Schmidhuber. Deep neural networks segment neuronal membranes in electron microscopy images.

In F. Pereira, C.J. Burges, L. Bottou, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc., 2012.

[9] Shujian Deng, Xin Zhang, Wen Yan, Eric I-Chao Chang, Yubo Fan, Maode Lai, and Yan Xu. Deep learning in digital pathology image analysis: a survey. *Frontiers of Medicine*, 14(4):470, 2020.

[10] Shrey Gadiya, Deepak Anand, and Amit Sethi. Histographs: Graphs in histopathology, 2019.

[11] Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. Neural message passing for quantum chemistry. *CoRR*, abs/1704.01212, 2017.

[12] Liyu Gong and Qiang Cheng. Adaptive edge features guided graph attention networks. *CoRR*, abs/1809.02709, 2018.

[13] Tai Hasegawa, Helena Arvidsson, Nikolce Tudzarovski, Karl Meinke, Rachael V. Sugars, and Aravind Ashok Nair. Edge-based graph neural networks for cell-graph modeling and prediction. In *Information Processing in Medical Imaging: 28th International Conference, IPMI 2023, San Carlos de Bariloche, Argentina, June 18–23, 2023, Proceedings*, page 265–277, Berlin, Heidelberg, 2023. Springer-Verlag.

[14] Guillaume Jaume, Pushpak Pati, Behzad Bozorgtabar, Antonio Foncubierta-Rodríguez, Florinda Feroce, Anna Maria Anniciello, Tilman Rau, Jean-Philippe Thiran, Maria Gabrani, and Orcun Goksel. Quantifying explainers of graph neural networks in computational pathology. *CoRR*, abs/2011.12646, 2020.

[15] Guillaume Jaume, Pushpak Pati, Antonio Foncubierta-Rodríguez, Florinda Feroce, Giosue Scognamiglio, Anna Maria Anniciello, Jean-Philippe Thiran, Orcun Goksel, and Maria Gabrani. Towards explainable graph representations in digital pathology. *CoRR*, abs/2007.00311, 2020.

[16] Hannah Kim and Young-Seob Jeong. Sentiment classification using convolutional neural networks. *Applied Sciences*, 9(11):2347, Jun 2019.

[17] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In Yoshua Bengio and Yann LeCun, editors, *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015.

[18] Liying Liu and Yain Whar Si. 1d convolutional neural networks for chart pattern classification in financial time series. *The Journal of Supercomputing*, 78, 08 2022.

[19] Yehong Liu and Guosheng Yin. The delaunay triangulation learner and its ensembles. *Computational Statistics & Data Analysis*, 152:107030, 2020.

[20] Erinc Merdivan, Anastasios Vafeiadis, Dimitrios Kalatzis, Sten Hanke, Joahannes Kroph, Konstantinos Votis, Dimitrios Giakoumis, Dimitrios Tzovaras, Liming Chen, Raouf Hamzaoui, and Matthieu Geist. Image-based text classification using 2d convolutional neural networks. pages 144–149, 2019.

[21] Aravind Nair, Helena Arvidsson, Jorge E. Gatica V., Nikolce Tudzarovski, Karl Meinke, and Rachael. V. Sugars. A graph neural network framework for mapping histological topology in oral mucosal tissue. *BMC Bioinformatics*, 23(1):506, 2022.

[22] Mathias Niepert, Mohamed Ahmed, and Konstantin Kutzkov. Learning convolutional neural networks for graphs. *CoRR*, abs/1605.05273, 2016.

[23] Yigit Ozen, Selim Aksoy, Kemal Kösemehmetoğlu, Sevgen Önder, and Ayşegül Üner. Self-supervised learning with graph neural networks for region of interest retrieval in histopathology. In *2020 25th International Conference on Pattern Recognition (ICPR)*, pages 6329–6334, 2021.

[24] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional net-

works for biomedical image segmentation. *CoRR*, abs/1505.04597, 2015.

[25] Youyi Song, Ling Zhang, Siping Chen, Dong Ni, Baiying Lei, and Tianfu Wang. Accurate segmentation of cervical cytoplasm and nuclei based on multi-scale convolutional network and graph partitioning. *IEEE transactions on bio-medical engineering*, 62, 05 2015.

[26] Linda Studer, Jannis Wallau, Heather Dawson, Inti Zlobec, and Andreas Fischer. Classification of intestinal gland cell-graphs using graph neural networks. In *2020 25th International Conference on Pattern Recognition (ICPR)*, pages 3636–3643, 2021.

[27] Mookund Sureka, Abhijeet Patil, Deepak Anand, and Amit Sethi. Visualization for histopathology images using graph convolutional neural networks. In *2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE)*. IEEE, oct 2020.

[28] Victor Tollemar, Nikcole Tudzarovski, Gunnar Warfvinge, Naom Yarom, Mats Remberger, Robert Heymann, Karin Garming Legert, and Rachael V. Sugars. Histopathological grading of oral mucosal chronic graft-versus-host disease: Large cohort analysis. *Biology of Blood and Marrow Transplantation*, 26(10):1971–1979, 2020.

[29] Du Wang, Chaochen Gu, Kaijie Wu, and Xinping Guan. Adversarial neural networks for basal membrane segmentation of microinvasive cervix carcinoma in histopathology images. In *2017 International Conference on Machine Learning and Cybernetics (ICMLC)*, volume 2, pages 385–389, 2017.

[30] Ruishuang Wang, Zhao Li, Jian Cao, Tong Chen, and Lei Wang. Convolutional recurrent neural networks for text classification. pages 1–6, 2019.

[31] Xi Wang, Hao Chen, Caixia Gan, Huangjing Lin, Qi Dou, Qitao Huang, Muyan Cai, and Pheng-Ann Heng. Weakly supervised learning for whole slide lung cancer image classification. In *Medical Imaging with Deep Learning*, 2018.

[32] Hai-Shan Wu, Steven Dikman, and Joan Gil. A semi-automatic algorithm for measurement of basement membrane thickness in kidneys in electron microscopy images. *Computer Methods and Programs in Biomedicine*, 97(3):223–231, 2010.

[33] Tian Xie and Jeffrey C. Grossman. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys. Rev. Lett.*, 120:145301, Apr 2018.

[34] Yan Xu, Yang Li, Yipei Wang, Mingyuan Liu, Yubo Fan, Maode Lai, and Eric I-Chao Chang. Gland instance segmentation using deep multichannel neural networks. *CoRR*, abs/1611.06661, 2016.

[35] Yan Xu, Tao Mo, Qiwei Feng, Peilin Zhong, Maode Lai, and Eric I-Chao Chang. Deep learning of feature representation with multiple instance learning for medical image analysis. *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1626–1630, 2014.

[36] Yulei Yang and Dongsheng Li. Nenn: Incorporate node and edge features in graph neural networks. In Sinno Jialin Pan and Masashi Sugiyama, editors, *Proceedings of The 12th Asian Conference on Machine Learning*, volume 129 of *Proceedings of Machine Learning Research*, pages 593–608. PMLR, 18–20 Nov 2020.

[37] Yu Zheng, Bihan Jiang, Jun Shi, Hao Zhang, and Fuyong Xie. Encoding histopathological wsis using gnn for scalable diagnostically relevant regions retrieval. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, pages 550–558, 2019.

[38] Yanning Zhou, Simon Graham, Navid Alemi Koohbanani, Muhammad Shaban, Pheng-Ann Heng, and Nasir Rajpoot. Cgc-net: Cell graph convolutional network for grading of colorectal cancer histology images, 2019.

Appendix

Cell-graph model - Additional Detais

In the cell-graph model, edges formed based on Delaunay triangulation are associated with specific triangular motifs, which are crucial for the graph structure (see Figure 5). The dataset was categorized into distinct training, validation, and testing subsets. Table 2 provides a detailed distribution of edge classes across these splits.

Node features in the dataset encompass manually annotated features like cell type, as well as computed features such as cell density and cell entropy. Deep learning features were extracted using a pre-trained ResNet-18 model.

For edge features, we considered various metrics including node distance, cell density difference, and cell entropy difference. These features were also normalized following previous work in the field [13].

Dataset	Non-Crossing	Crossing	Total
Training	96195	6221	102416
Validation	20017	1169	21186
Testing	45779	2733	48512
Total	161991	10123	172114

Table 2: Distribution of edge classes across different splits.

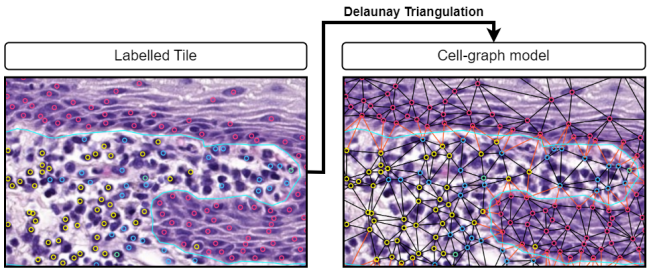


Figure 5: Cell-graph model generation. (Left) Manually Annotated Tile. (Right) Generated Graph; (red) BM crossing edges, (black) BM non-crossing edges.

Edge Aggregated GNN (EAGNN)

The architecture builds on the EAGNN node embedding layers, as discussed in [13].

The node embedding layers aim to derive latent node representations from a graph. The EAGNN architecture follows a layered structure, where each layer k computes a low-dimensional (d_k-dimensional) representation $h_u^k \in \mathbb{R}^{d_k}$ of the graph structure around node u.

In the EAGNN, the computation at each layer k is slightly different from the general GNN. Instead of performing an AGGREGATE operation on nodes followed by a COMBINE operation, the EAGNN incorporates multiple edge features by aggregating edge features into the node embeddings. The major difference between the EAGNN layer and other GNN layers is the way that edge features are normalized before aggregation.

Following the matrix multiplication as proposed in [3, 11, 22], the aggregation operation of the proposed model at layer k, named EAgg^k, is formulated as follows:

$$\text{EAgg}^k(E_p, H^{k-1}) = E_p H^{k-1} W_0^k \qquad (1)$$

where $H^{k-1} \in \mathbb{R}^{n \times d_{k-1}}$ is the node embeddings matrix at layer $k-1$[3], $W_0^k \in \mathbb{R}^{d_{k-1} \times d_k}$ is a matrix of learnable parameters, and E_p is the p-th feature matrix of the edges (i.e. $E_p \in \mathbb{R}^{n \times n}$ is the projection of $E \in \mathbb{R}^{n \times n \times P}$ edge feature tensor onto the single edge feature p).

Then, the previous node representation is combined using the combine operation. These aggregation and combining operations are performed for each edge feature, and they are then concatenated all together. Therefore, the formula for the k-th EAGNN layer is given by:

$$H^k = \sigma[\|_{p=1}^{P} (E_p H^{k-1} W_0^k + H^{k-1} W_1^k)] \qquad (2)$$

where $\|$ denotes the concatenation operator and σ a non-linear function and $W_1^k \in \mathbb{R}^{d_{k-1} \times d_k}$ is a second matrix of learnable parameters. Note that this non-linear function is not used in the final layer K of the node embedding layers. After K embedding layers, the node representation z_u is given by $z_u = H_u^K$ for node u.

Validation Results

Various configurations of EAGNN_i were evaluated, among which (EAGNN_5,TM-CNN) demonstrated the best performance on the validation set and was consequently selected as the main model. Table 3 presents the results for different EAGNN configurations.

[3] Note, $H^0 = X$

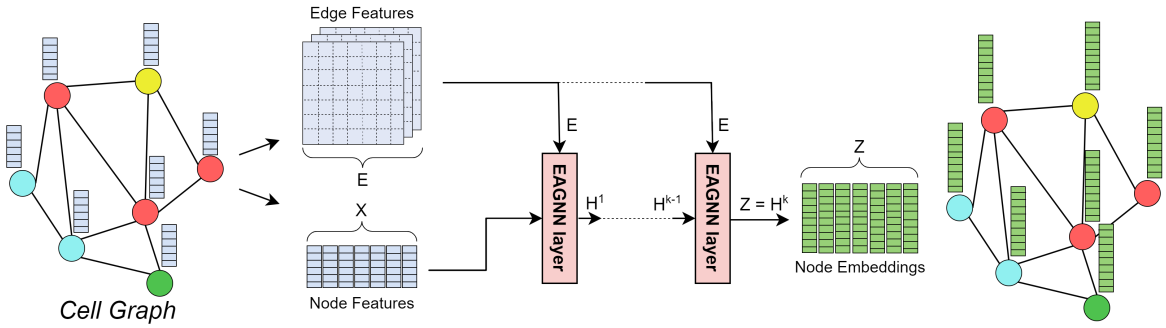


Figure 6: EAGNN architecture overview.

Model	Classifier	NF	EF	Precision	Recall	F1	ROC-AUC	Accuracy
$EAGNN_2$	TM-CNN	N_1	E_{1234}	0.8792	0.8529	0.8658	**0.9930**	0.9854
$EAGNN_3$	TM-CNN	N_1	E_{1234}	0.8873	0.8554	0.8711	0.9926	0.9860
$EAGNN_4$	TM-CNN	N_1	E_{1234}	0.8882	0.8563	0.8720	0.9927	0.9861
$EAGNN_5$	TM-CNN	N_1	E_{1234}	**0.9018**	0.8640	**0.8825**	0.9926	**0.9873**
$EAGNN_6$	TM-CNN	N_1	E_{1234}	0.8846	**0.8657**	0.8751	**0.9930**	0.9864

Table 3: Performance metrics of the various ($EAGNN_i$,TM-CNN) configurations on the validation set. The metrics are employed to ascertain the optimal number of aggregation layers (i) within the EAGNN backbone model. Node Features (NF) and Edge Features (EF) represent aggregated node and edge characteristics, respectively.