Proceedings Paper:
Sims, J., Grabsch, H.I. orcid.org/0000-0001-9520-6228 and Magee, D. orcid.org/0000-0003-2170-3103 (2022) Using Hierarchically Connected Nodes and Multiple GNN Message Passing Steps to Increase the Contextual Information in Cell-Graph Classification. In: Imaging Systems for GI Endoscopy, and Graphs in Biomedical Image Analysis, ISGIE GRAIL 2022. ISGIE (Imaging Systems for GI Endoscopy) 2022, 18 Sep 2022, Singapore. Lecture Notes in Computer Science, 13754 . Springer Nature Switzerland , Cham, Switzerland , pp. 99-107. ISBN 9783031210822

https://doi.org/10.1007/978-3-031-21083-9_10

Using Hierarchically Connected Nodes and Multiple GNN Message Passing Steps to Increase the Contextual Information in Cell-Graph Classification

Joe Sims[2,3], Heike I. Grabsch[1,2], Derek Magee[3]

[1] Department of Pathology, GROW School for Oncology and Reproduction, Maastricht University Medical Center+, Maastricht, NL.
[2] Pathology and Data Analytics, Leeds Institute of Medical Research at St. James's, University of Leeds, Leeds, UK.
[3]School of Computing, University of Leeds, Leeds, UK.

Abstract. Graphs are useful in analysing histopathological images as they are able to represent neighbourhood interactions and spatial relationships. Typically graph nodes represent cells and the vertices are constructed by applying a nearest neighbor algorithm to cell's locations. When passing these graphs through one graph neural network (GNN) message passing step, each node can only utilise features from nodes within its immediate neighbourhood to make a classification. To overcome this, we introduce two levels of hierarchically connected nodes that we term "supernodes". These supernodes, used in conjunction with at least four GNN message passing steps, allow for cell node classifications to be influenced by a wider area, enabling the entire graph to learn tissue-level structures. The method is evaluated on a supervised task to classify individual cells as belonging to a specific tissue class. Results demonstrate that the inclusion of supernodes with multiple GNN message passing steps increases model accuracy.

Keywords: graph neural network, node classification, digital pathology.

1 Introduction

The phenotype and topological distribution of tissue components may influence cancer progression as well as patient prognosis and response to therapy [1, 2]. Convolutional neural networks (CNNs) have demonstrated to be effective at common computer vision tasks such as image classification and segmentation [3]. However, traditional CNNs only model local relations and are applied to data in a grid structure with fixed connectivity. When applying CNNs to patches from multi-gigabyte whole slide images (WSIs) this limits the model from learning wider representations and doesn't consider the interactions between entities within the tumour microenvironment. This paper tackles this issue by representing tissue as a graph structure which conserves spatial relations. We augment the graphs with multiple levels of hierarchy to increase the radius of spatial context that each node can utilise.

Graphs inherently capture relationships between entities making them appropriate for representing the tumour microenvironment. Graph neural networks (GNNs) are a variation of deep learning that accept graphs as input. These GNNs are able to capture different neighbourhood relations and accept irregular sized inputs. Furthermore, they have shown comparable accuracies to CNNs when performing disease classification [4, 5], and tissue segmentation [6]. One subclass of GNNs is a graph attention network (GAT) [7] which leverages masked self-attention layers to learn different weights for specific nodes within a neighbourhood of arbitrary size.

Propagating a graph through a GNN once is often referred to as one message passing step. In one GNN message passing step, information from nodes one hop away influence the learnt node embeddings. With every additional T GNN message passing step, information from nodes T hops away influence the learnt node embeddings [8]. However, there is currently no guidance regarding the number of message passing steps required for 'optimal' learning of node representations, specifically when applied to hierarchical graphs.

Several approaches have been made to develop adequate tissue-representations using graphs and GNNs. Zhou, Y., et al. (2019) [4] captured the tumour cell microenvironment using cell-graphs (graphs whose nodes represent cells), while others have used a number of clustering methods to represent tissue-level structures [1, 5, 6]. However, independent of the graph formation method, whilst only using one GNN message passing step, learnt node embeddings will be limited to the influence of its immediate neighbours. P. Pati, et al. (2021) [1] suggested a hierarchical graph structure which introduced connectivity between cells and larger, non-overlapping tissue regions. Although this method was successful at increasing context and capturing multi-scale information, the number of nodes representing the tissue level structures in conjunction with only two GNN message passing steps, limits the context that can be learnt. Furthermore, there was no suggested method for introducing additional layers of hierarchical connections.

In P. Pati, et al., (2021) [1] and Anklin, V., et al. [6] node features were obtained by passing forward image patches through a pre-trained CNN to produce features that are abstract and exposed to bias arising from variability in colour and scanner-specific attributes across pathology slides [9]. Whereas, Zhou, Y., et al. (2019) [4] extracted morphological cell features which are independent of these biases.

In this paper, we propose a method that applies the concept of hierarchical graph formation to cell-graphs to increase the contextual information when learning tissue-level representations. We introduce two sets of sparsely distributed, regularly spaced nodes termed "supernodes" which form edge connections in a hierarchical manner. To produce an optimal result from these nodes, we can demonstrate that a GNN model's architecture should be composed of multiple (at least four) GNN message passing steps. The main contributions of this paper are:

- A novel method for creating hierarchical graphs that increases contextual information without being limited by the size of tissue regions or exposed to bias arising from variation across histopathology slides;

- The use of multiple, (at least four) independently-weighted GNN message passing steps to utilise the hierarchically connected supernodes within the graph;

- An evaluation of the proposed methods in a node classification task to segment tissue regions in 54 HE-stained cores containing tumour regions from patients with stage II/IIIb gastric cancer.

2 Method

2.1 Data

The dataset used in this paper was composed of 2 haematoxylin-eosin (HE) stained tissue-microarrays (TMAs) containing 54 3mm diameter tissue cores sampled from tumour regions from patients with stage II/IIIb gastric cancer. Using the HeteroGenius MIM Cell-Analysis Add-On (HeteroGenius, Leeds, UK) which is a U-NET-based cell detector and classifier trained on over 50,000 annotated cells, the centroid position of every cell nucleus within the cores were detected along with 14 other features. These included size (μm), elongation, mean intensity, standard deviation of intensity, angle and the cell's probability of being one of the following cell types: tumour, lymphocyte, granulocyte, plasma cell, fibroblast, muscle, endothelium, normal epithelium, and other. Within the two TMAs this resulted in ~2.6 million cells detected with their 14 corresponding features.

The specific task this method was applied to was node classification to identify large tissue structures. Four output classes were identified with a pathologist. These classes were: cancer, muscle, stroma and follicle (aggregates of lymphocytes). Regions were manually annotated by a pathologist to define the ground truths. Roughly ~2.5% (~67,000) of the total cells were labelled. For each core, the cells were randomly assigned to a train-test split of 80% and 20%, respectively.

2.2 Cell-Graph Formation

A graph is defined as $G = (V, E)$ where V are a set of nodes (vertices) with corresponding features and E are the edges connecting two nodes. These have corresponding features that represent their interaction. To form the cell-graph, each cell was represented as a node. Assuming that cell interactions occur between adjacent cells, edge connections were formed by applying the k-d tree nearest-neighbor algorithm to all cell coordinates in the individual TMA cores, with k-neighbours=5 and the number of leaf nodes set to n=3.

2.3 Supernodes

To increase the context of a cell node's learnt embedding, we introduce two levels of hierarchically connected supernodes. To define the first level of supernodes (L1), a square grid of edge length 200μm is superimposed onto the TMA core. The locations

of the grid's vertices define the locations of the L1 supernodes. Edges are formed with every cell node and other L1 supernodes within a 200μm radius. The L1 supernode features are calculated as the average of the cell node features whom it shares an edge with.

For the second level of supernodes (L2), a square grid of edge length 400μm, aligned with the same coordinate space as the L1 grid, is superimposed onto the TMA core. The locations of this grid's vertices define the locations of the L2 supernodes. Edges are formed with every L1 supernode within a 200μm radius. **Fig. 1** provides a visualization of the locations of the two levels of supernodes and one set of hierarchically connected nodes.

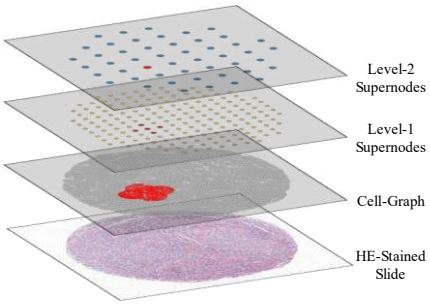

Fig. 1. A visualization of the supernodes' positions and an example of one L2 supernode and its hierarchical connections (connected nodes shown in red). The HE-stained slide is included for reference.

Introducing the supernodes significantly increases the number of edge connections. Across the 54 TMA cores, **Table 1** shows the average number of cells, L1 supernodes, and L2 supernodes, as well as the number of cumulative edges as the supernodes are introduced. The number of edges includes shown in **Table 1** includes self-loops.

Table 1. For each node type, the number of nodes and edges are the average over the 54 TMA cores. The number of edges include self-loops.

Node Type	Number of Nodes	Number of Edges
Cell Nodes Only	30,205	85,899
Cell Nodes + L1	30,373	175,888
Cell Nodes + L1 + L2	30,416	176,224

2.4 Model Architecture

The method in this work uses a GAT with one attention head that contains a 2-layer feed forward neural network within one GNN message passing step. The feed forward network is shown in Equation (1) where h_i is the set of input node features for the i-th node, h_i' is the corresponding transformed output, $\boldsymbol{W}_k$ are learnable sets of weights, and

tanh is the hyperbolic tangent activation function. The activation function *tanh* was used over *LeakyReLU* due to its superior performance during experimentation.

The input node features had size 17 which consisted of the 14 output features from the HeteroGenius MIM cell-analysis tool along with a one-hot encoding of the node's supernode status. Cell nodes were encoded as [1,0,0], L1 supernodes as [0,1,0], and L2 supernodes as [0,0,1].

$$\vec{h}_i' = tanh\left(\boldsymbol{W}_3 tanh\left(\boldsymbol{W}_2 tanh\left(\boldsymbol{W}_1 \vec{h}_i\right)\right)\right) \tag{1}$$

The input also included 3 edge features. These were: the difference in x-coordinates; the difference in y-coordinates; and the Euclidean distance between the source and target node. For all nodes, self-loops were included with edge features equal to 0.

The attention mechanism within the network was modified to account for edge features. In Equation (2), the edge features e_{ij} corresponding to the edge connecting the i-th node and its j-th neighbor are concatenated with the transformed node features [10], where a is the attention weight vector and α_{ij} is the multi-head attention coefficient.

$$\alpha_{ij} = \frac{exp\left(LeakyRelu\left(\vec{a}^T[\vec{h}_{i'} \| \vec{h}_{j'} \| \vec{e}_{ij}]\right)\right)}{\sum_{k \in \mathcal{N}_i}\left(exp\left(LeakyRelu\left(\vec{a}^T[\vec{h}_{i'} \| \vec{h}_{j'} \| \vec{e}_{ij}]\right)\right)\right)} \tag{2}$$

Mentioned in Xu, K., et al. (2018) [11] and evident in Equation (2), one GNN message passing step allows information from immediately connected neighbours (one hop away) to influence the i-th node's representation. To embed information from nodes T hops away, information would have to pass sequentially through T GNN message passing steps, where each message passing step is comprised of a new set of weights. This makes it unfeasible to embed distant information by stacking tens of message passing steps end-to-end as the number of parameters would either be too large to train, making it subject to computational limits or be subject to vanishing and exploding gradients. Furthermore, nodes that are tens of hops away would have negligible influence on node's learnt representations compared with nodes fewer hops away. However, by using two levels of supernodes in conjunction with multiple (four) GNN message passing steps, information is able to travel from one cell node up to a L2 supernode and back down to another cell node that exists a maximum distance of 800µm away. **Fig. 2.** demonstrates this concept of information travelling from one cell node to a L1 and L2 supernode, then travelling back down to a separate L1 supernode and cell node, using four GNN message passing steps.

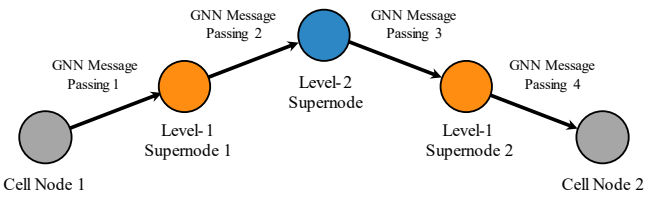

Fig. 2. Above shows how information from Cell Node 1 can be influence the learnt node embedding of Cell Node 2 by being passing information through 4 GNN message passing steps.

Although the GAT model allows for the implementation of multiple attention heads and the inclusion of a multi-layer feed forward network, both of which can increase the accuracy of the learnt representation, this would not achieve the same function as multiple GNN message passing steps. Increasing the number of attention heads and layers in the feed forward network would continue to aggregate information one hop away and not allow for the influence of nodes multiple hops away.

2.5 Experiential Evaluation

To assess the influence of supernodes on the learnt node representations and evaluate the effect of using multiple GNN message passing steps, two experiments were carried out. One was to assess how the addition of L1 supernodes and L2 supernodes influence the overall accuracy. The other was to assess how the number of GNN message passing steps influence the accuracy of the model when using two levels of supernodes. During training the loss function used was a masked, weighted MSE loss. Each model was trained for 10,000 epochs with a learning rate of 1e-3.

3 Results

3.1 The Inclusion of Supernodes

Firstly, all models used in this experiment had 4 GNN message passing steps. The node feature sizes from the input (17) to output (4) were 17, 64, 64, 64 and 4. For comparison, three models were trained and tested on the data containing different levels of hierarchy. Specifically the three sets of data were: cell nodes with no supernodes; cell nodes with L1 supernodes; and cell nodes with both L1 and L2 supernodes. No further cross validation was carried out. **Table 2** shows the accuracy of the three models on the train and test sets.

Table 2. The model accuracy when trained and tested with and without the existence of supernodes. 'Node Type' represents the additional presence of the L1 and L2 supernodes.

Node Type	Train Accuracy	Test Accuracy
Cell Nodes Only	73.66	73.15
Cell Nodes + L1	90.69	90.50
Cell Nodes + L1 + L2	**93.80**	**93.40**

It is clear in **Table 2** that with the addition of supernodes, there was a significant increase in accuracy. **Fig. 3** provides a visual demonstration of how the model's output was affected when including the different levels of supernodes. **Fig. 3a)** shows a clear presence of localized noise, and the classifications appear to be more locally clustered. Whereas, in **Fig. 3c)** larger scale structures such as tissue regions are evident, showing an increase in context in the learnt node embeddings.

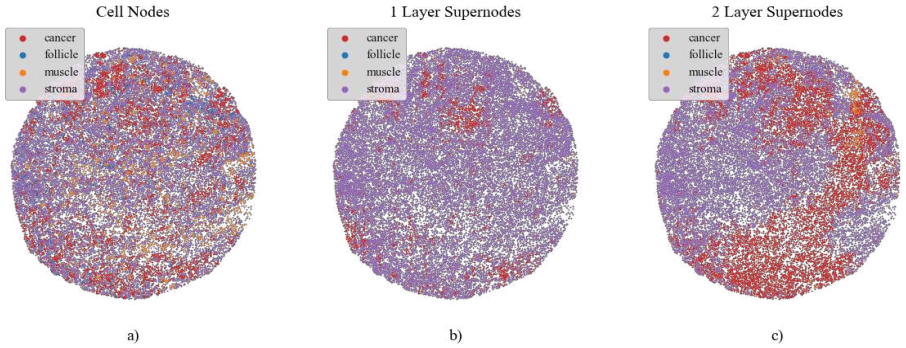

Fig. 3. An example of the model's outputs when trained on a) only the cell-graph, b) the cell-graphs with the inclusion of L1 supernodes, and c) the cell-graph with the inclusion of L1 and L2 supernodes.

3.2 Multiple GNN Message Passing Steps

The second experiment was to determine how the number of GNN message passing steps influences the model's performance when used in conjunction with 2 levels of supernodes. Four models were trained for comparison. For each model, the number of input features were 17 and output features were 4. For models composed of >1 GNN message passing step, the intermediate node features were 64. These are shown in **Table 3**.

Table 3. The accuracy of the model composed of n message passing steps where n = 1,2,3,4.

Number of Message Passing Steps	Node Features	Train Accuracy	Test Accuracy
1	17, 4	72.77	73.15
2	17, 64, 4	89.38	88.12
3	17, 64, 64, 4	90.31	90.50
4	17, 64, 64, 64, 4	**93.80**	**93.40**

The results in **Table 3** suggest that increasing the number of GNN message passing steps increases the accuracy when applied to hierarchical graphs containing two levels of supernodes.

4 Discussion

P. Pati, et al. (2021) [1] introduced applying hierarchically connected nodes to histopathology slides. They determined nodes as the centres of tissue regions that were connected to cells lying within those tissue regions. This was demonstrated to be effective at increasing context in a graph classification task. However, no method was proposed for further increasing the context through more hierarchical connections. As a result, the learnt node representations were limited by the size of tissue regions and the number

of GNN message passing steps (two). Furthermore, the node features were extracted from passing patches of HE-stained slides through a pretrained network. These features were abstract and uninterpretable with no clinical context.

W. Lu, et al. (2020) [5] defined nodes as the centres of cell clusters. Similarly to the methods presented in this work, the features of these nodes were an average of the cell node features within the cluster. This allowed for tissue-level structures to be represented but could not capture higher resolution information from a cellular level. Whilst this presented an efficient method of learning graph representations, there was no proposed method for developing hierarchical connections nor was there a suggested number of GNN message passing steps to optimize the potentially learnt node embeddings.

The method proposed in this paper allows for the representation of granular information from the cells and provides a method for creating indefinite levels of hierarchically connected nodes that we termed "supernodes". These supernodes have shown to increase context in a node classification task, allowing for tissue-level structures to be learnt on a cellular level. We demonstrate that to optimize the outcomes from using two levels of supernodes, multiple (at least four) GNN message passing steps are required.

When comparing the number of message passing steps, it is likely that the increase in performance seen in **Table 2** was a result of including more learnable parameters. However, we demonstrated that with the inclusion of 2 levels of hierarchically connected supernodes, this number of GNN message passing steps enabled the graph to learn from an increased context. In addition, the model accuracy is high (>90%) and only considered 17 input node features, all of which held clinical relevance.

One limitation of this work lies in the labelled output classes. With the tumour-microenvironment being complex and heterogeneous [8, 12], in many scenarios there are more than four tissue classes. Likewise, these tissue regions don't consistently maintain hard boundaries which makes the assignment of a single class subjective and inconsistent. However, this was not inherent to the model but to the labelling system. A further limitation lies within the quantity of annotated data (~2.5% of all cell nodes). This is significantly lower than that proposed in Gao, J.P., et al., (2018) [8] who demonstrated the effect of incomplete labels on a graph-based segmentation task, with the lowest percentage of annotated data covering 5% of the total pixels within the tissue.

5 Conclusion

In this paper, we have presented a novel method for increasing the contextual information when performing node classification on cell-graphs containing tumour regions. We introduced the concept of supernodes that can be connected hierarchically. From comparing the accuracy of models trained on cell nodes alone, cell nodes with one level of supernodes, and cell nodes with two levels of supernodes, we can conclude that the inclusion of supernodes increases the contextual information learnt by cell nodes. Through a separate comparison, the accuracy was compared between four models composed of 1-4 GNN message passing steps. The model with 4 GNN message passing steps achieved the highest performance implying that this architecture is required to utilise two levels of supernodes.

References

1. Pati, P., et al., *Hierarchical graph representations in digital pathology.* Med Image Anal, 2022. **75**: p. 102264.
2. Mi, H., et al., *Digital Pathology Analysis Quantifies Spatial Heterogeneity of CD3, CD4, CD8, CD20, and FoxP3 Immune Markers in Triple-Negative Breast Cancer.* Front Physiol, 2020. **11**: p. 583333.
3. Janowczyk, A. and A. Madabhushi, *Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases.* J Pathol Inform, 2016. **7**: p. 29.
4. Zhou, Y., et al. *CGC-Net: Cell Graph Convolutional Network for Grading of Colorectal Cancer Histology Images*. in *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*. 2019.
5. Lu, W., et al. *Capturing Cellular Topology in Multi-Gigapixel Pathology Images*. in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 2020.
6. Anklin, V., et al. *Learning Whole-Slide Segmentation from Inexact and Incomplete Labels using Tissue Graphs*. 2021. arXiv:2103.03129.
7. Veličković, P., et al. *Graph Attention Networks*. 2017. arXiv:1710.10903.
8. Gao, J.P., et al., *Tumor heterogeneity of gastric cancer: From the perspective of tumor-initiating cell.* World J Gastroenterol, 2018. **24**(24): p. 2567-2581.
9. Anghel, A., et al., *A High-Performance System for Robust Stain Normalization of Whole-Slide Images in Histopathology.* Front Med (Lausanne), 2019. **6**: p. 193.
10. Chen, J. and H. Chen *Edge-Featured Graph Attention Network*. 2021. arXiv:2101.07671.
11. Xu, K., et al. *Representation Learning on Graphs with Jumping Knowledge Networks*. 2018. arXiv:1806.03536.
12. Junttila, M.R. and F.J. de Sauvage, *Influence of tumour micro-environment heterogeneity on therapeutic response.* Nature, 2013. **501**(7467): p. 346-54.