Self-Supervised Learning with Graph Neural Networks for Region of Interest Retrieval in Histopathology

Yiğit Özen[1], Selim Aksoy[1], Kemal Kösemehmetoğlu[2], Sevgen Önder[2], Ayşegül Üner[2]

[1]Department of Computer Engineering, Bilkent University, Ankara, Turkey

[2]Department of Pathology, Hacettepe University, Ankara, Turkey

`yigit.ozen@bilkent.edu.tr, saksoy@cs.bilkent.edu.tr`

Abstract

Deep learning has achieved successful performance in representation learning and content-based retrieval. However, there are two major challenges. First, supervised training of deep networks requires large amount of labeled data which is limited in the medical field. Second, the clinical practice in histopathology uses regions of interest (ROI) with arbitrary shapes and sizes. We propose a generic method that utilizes graph neural networks (GNN) and self-supervised learning. GNN enables representing arbitrarily-shaped ROIs as graphs and encoding contextual information. Self-supervised training improves quality of learned representations without requiring labeled data.

1. Introduction

- Advances in slide scanning technologies have enabled the whole diagnostic process of cancer done in digital form.
- Pathological image analysis systems have a great potential in aiding this process.
- Content-based histopathology image retrieval systems can provide auxiliary information to pathologists by identifying regions that have similar content.

2. Challenges

- Effective retrieval requires effective representations.
- Learning representations through deep neural networks requires large amount of labeled data.
- Data annotation is time-consuming and data sharing is difficult in histopathology.
- Diagnostically relevant regions of interest (ROI) have variable size and shape.
- Patch-based methods cannot exploit context information.

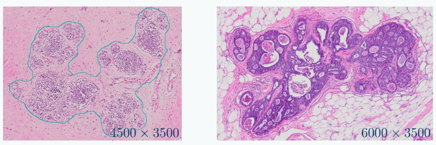

Figure 1: *ROIs with varying shapes and sizes.*

3. Proposed Methodology

Overview:

- Model ROIs as graphs with fixed-sized patches as vertices.
- Encode patches using a convolutional neural network.
- Learn ROI representations using a graph neural network.

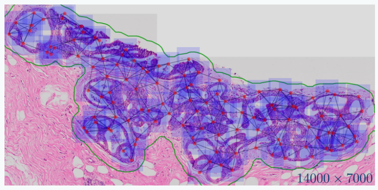

Figure 2: *Example ROI represented as a graph of patches.*

Content-based Retrieval:

- Train a graph neural network (GNN) using a self-supervised contrastive loss without any need for labels.
- Establish the retrieval database with the representations extracted by the GNN.
- Perform retrieval using Euclidean distance between the representations.

Self-Supervised Learning:

1. Sample a mini-batch of M ROI graphs.
2. Augment each graph twice by randomly dropping vertices, resulting in $2M$ graphs in the batch.
3. Encode each graph using the GNN.
4. Apply a nonlinear transformation using an MLP projection.
5. Compute the loss function for a positive pair (i, j) as

$$l_{i,j} = -\log \frac{\exp(sim(z_i, z_j)/\tau)}{\Sigma_{k=1}^{2M} \mathbb{I}_{[k \neq i]} \exp(sim(z_i, z_k))/\tau}$$

where sim is the cosine similarity, z is the output of the projection, and τ is the temperature parameter.

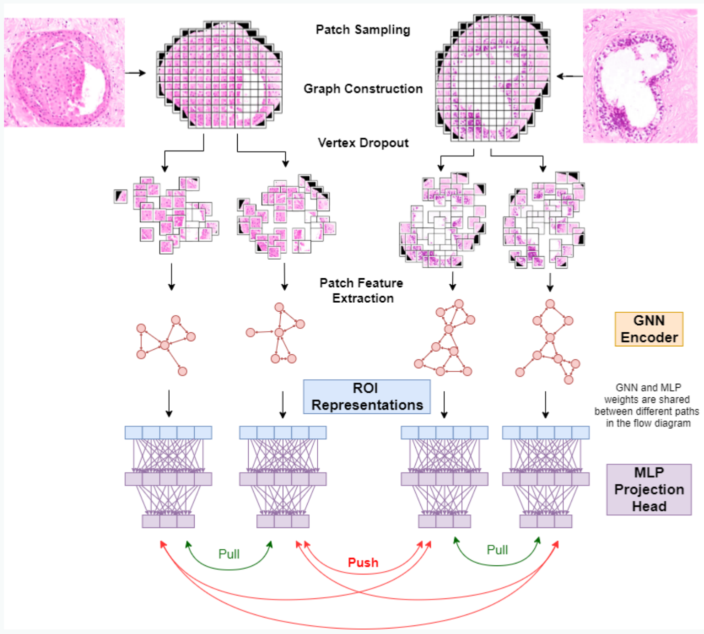

Figure 3: *The self-supervised learning pipeline.*

ICPR 2020, International Conference on Pattern Recognition, January 10–15, 2021, Milan, Italy

Self-Supervised Learning with Graph Neural Networks for Region of Interest Retrieval in Histopathology

Yiğit Özen[1], Selim Aksoy[1], Kemal Kösemehmetoğlu[2], Sevgen Önder[2], Ayşegül Üner[2]

[1]Department of Computer Engineering, Bilkent University, Ankara, Turkey
[2]Department of Pathology, Hacettepe University, Ankara, Turkey

yigit.ozen@bilkent.edu.tr, saksoy@cs.bilkent.edu.tr

3. Proposed Methodology (continued)

Graph Neural Network:

- Common elements:
  - Neighborhood aggregation: encodes context.
  - Local pooling: learns hierarchical features.
  - Global pooling: aggregates vertex features into graph representation.
- GCN: A number of stacked graph convolution layers, followed by global pooling.
- DiffPool: In addition to vertex encoding, learns vertex cluster assignments for local pooling.
- GraphConv: Top-k local pooling by dropping vertices based on a learned score.

4. Data Set

Table 1: *Class distribution of slides and ROIs in training, validation, and test sets. Note that a slide can contain multiple ROIs corresponding to different diagnostic labels, resulting in a multi-label setting for each slide.*

		Benign	Atypia	In Situ	Invasive	Total
Slide	Training Set	30	16	16	13	39
	Validation Set	15	7	8	6	18
	Test Set	16	8	9	6	21
	Total	61	31	33	25	78
ROI	Training Set	226	55	154	102	537
	Validation Set	109	25	56	50	240
	Test Set	105	30	69	49	253
	Total	440	110	279	201	1030

5. Results

Table 2: *ROI retrieval results for different methods and training settings.*

Method	Architecture	MAP@10	MAP@25
Supervised	DiffPool	0.62	0.59
	GraphConv	0.73	0.64
	GCN	0.80	0.76
Self-Supervised	DiffPool	0.78	0.70
	GraphConv	0.82	0.75
	GCN	**0.86**	**0.80**

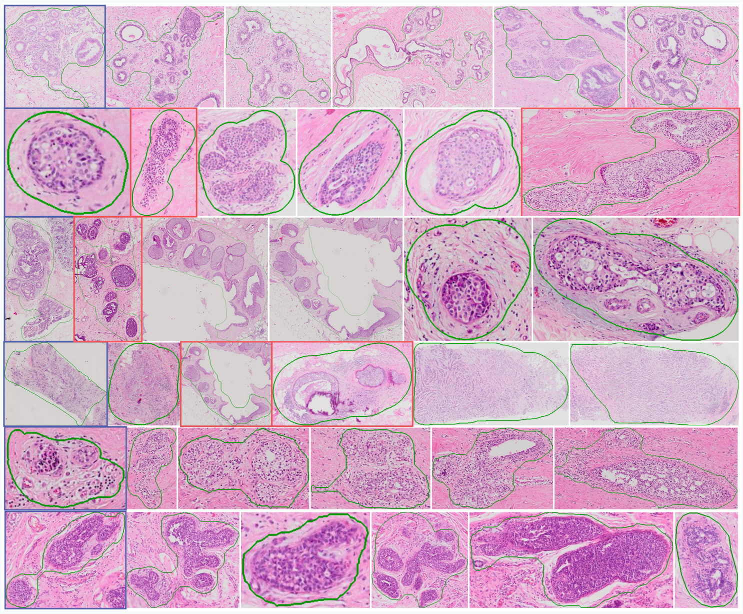

Figure 4: *Examples of ROI retrieval using the best model, showing the top 5 retrieved items for each query in separate rows. Query images are marked in blue. Among the retrieved images, the irrelevant ones are marked in red. The ROI boundaries are marked in green. Query ROI classes from top to bottom are: benign, in situ, in situ, invasive, atypia, and benign.*

6. Conclusions

- ROI images have vastly different shapes and sizes.
- Labeling ROIs is time-consuming.
- We proposed a framework to generate representations for variable-sized ROIs without label-based supervised learning.
- We demonstrated the representation power of the proposed approach in a retrieval scenario with comparative experiments on a breast histopathology data set.

Acknowledgment

This work was supported in part by the Scientific and Technological Research Council of Turkey (grant 117E172).