## Supplemental Material

## Yushan Zheng

Beijing Advanced Innovation Center for Biomedical Engineering, Beihang University, Beijing 100191, China. yszheng@buaa.edu.com

This is the supplemental material for paper "Tracing Diagnosis Paths on Histopathology WSIs for Diagnostically Relevant Case Recommendation, MIC-CAI 2020". It includes the algorithm of the ROI feature extraction network.

**Algorithm 1:** The computation flowchart of the ROI feature extraction network  $\mathbf{g} = \mathcal{F}_{graph}(\mathbf{A}, \mathbf{X})$ , where  $\mathcal{G}_{embed}^{(l)}$  and  $\mathcal{G}_{pool}^{(l)}$  are GCN modules formulated in Algorithm2.

```
formulated in Algorithm2.

Input:
\mathbf{A} \in \mathbb{R}^{n_p \times n_p} \leftarrow \text{The adjacency matrix of the } n_p \text{ patches in the given ROI.}
\mathbf{X} \in \mathbb{R}^{d \times n_p} \leftarrow \text{The patch features for the ROI.}
L \leftarrow \text{The number of Diffpool modules.}
n_l \leftarrow \text{The number of graph vetexes after the } l\text{-th pooling. Specifically, } n_0 = n_p
and n_{l+1} < n_l.

1 \mathbf{X}^{(0)} \leftarrow \mathbf{X}_i^T;
2 \mathbf{A}^{(0)} \leftarrow \mathbf{A}_i;
3 \mathbf{Z}^{(0)} \leftarrow \mathcal{G}_{embed}^{(0)}(\mathbf{A}^{(0)}, \mathbf{X}^{(0)}) \in \mathbb{R}^{n_0 \times d};
4 for l = 0 to L - 1 do
5 \mathbf{S}^{(l)} \leftarrow softmax_r \left(\mathcal{G}_{pool}^{(l)}(\mathbf{A}^{(l)}, \mathbf{X}^{(l)})\right) \in \mathbb{R}^{n_{l+1} \times n_l};
6 \mathbf{X}^{(l+1)} \leftarrow \mathbf{S}^{(l)T}\mathbf{Z}^{(l)} \in \mathbb{R}^{n_{l+1} \times d};
7 \mathbf{A}^{(l+1)} \leftarrow \mathbf{S}^{(l)T}\mathbf{A}^{(l)}\mathbf{S}^{(l)} \in \mathbb{R}^{n_{l+1} \times n_{l+1}};
8 \mathbf{Z}^{(l)} \leftarrow \mathcal{G}_{embed}^{(l)}(\mathbf{A}^{(l)}, \mathbf{X}^{(l)}) \in \mathbb{R}^{n_l \times d};
9 end
10 \mathbf{g} \leftarrow Maxpool_r(\mathbf{X}^{(L)});
```

11 return **g**;

```
Algorithm 2: The computation flowchart of GCN \mathbf{Z} \leftarrow \mathcal{G}(\mathbf{A}, \mathbf{X})

Input:

\mathbf{X} \leftarrow The features.

\mathbf{A} \leftarrow The adjacency matrix.

1 \tilde{\mathbf{A}} \leftarrow \mathbf{A} + \mathbf{I};

2 \tilde{\mathbf{D}} \leftarrow diag(\sum_{j} \tilde{\mathbf{A}}_{1j}, \sum_{j} \tilde{\mathbf{A}}_{2j}, ..., \sum_{j} \tilde{\mathbf{A}}_{nj}),;

3 \mathbf{H}^{(0)} \leftarrow \mathbf{X};

4 for k = 1 to K do

5 \mathbf{W}^{(k)} \leftarrow the trainable weighting matrix for the k-th step;

6 \mathbf{H}^{(k)} = ReLU(\tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}} \mathbf{H}^{(k-1)} \mathbf{W}^{(k)});

7 end

8 \mathbf{Z} \leftarrow \mathbf{H}^{(K)};

9 return \mathbf{Z};
```