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*". 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. In the experiment, L is set as 2.

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

$\mathbf{X}^{(0)} \leftarrow \mathbf{X}_i^{\mathrm{T}}$;
$\mathbf{A}^{(0)} \leftarrow \mathbf{A}_i$;
for $l = 0$ *to* $L - 1$ **do**
  $\mathbf{Z}^{(l)} \leftarrow \mathcal{G}_{embed}^{(l)}(\mathbf{A}^{(l)}, \mathbf{X}^{(l)}) \in \mathbb{R}^{n_l \times d}$;
  $\mathbf{S}^{(l)} \leftarrow softmax_r\left(\mathcal{G}_{pool}^{(l)}(\mathbf{A}^{(l)}, \mathbf{X}^{(l)})\right) \in \mathbb{R}^{n_{l+1} \times n_l}$;
  $\mathbf{X}^{(l+1)} \leftarrow \mathbf{S}^{(l)\mathrm{T}}\mathbf{Z}^{(l)} \in \mathbb{R}^{n_{l+1} \times d}$;
  $\mathbf{A}^{(l+1)} \leftarrow \mathbf{S}^{(l)\mathrm{T}}\mathbf{A}^{(l)}\mathbf{S}^{(l)} \in \mathbb{R}^{n_{l+1} \times n_{l+1}}$;
end
$\mathbf{W}_f, \mathbf{b}_f \leftarrow$ the trainable weighting matrix and bias for the fully connected layer;
$\mathbf{g} \leftarrow ReLU(Maxpool_r(\mathbf{X}^{(L)})\mathbf{W}_f + \mathbf{b}_f)$;
return $\mathbf{g}$;

Algorithm 2: The computation flowchart of GCN $\mathbf{Z} \leftarrow \mathcal{G}(\mathbf{A}, \mathbf{X})$. In the experiment, K is set as 3.

Input:
$\mathbf{X} \leftarrow$ The features.
$\mathbf{A} \leftarrow$ The adjacency matrix.
$\tilde{\mathbf{A}} \leftarrow \mathbf{A} + \mathbf{I}$;
$\tilde{\mathbf{D}} \leftarrow diag(\sum_j \tilde{\mathbf{A}}_{1j}, \sum_j \tilde{\mathbf{A}}_{2j}, ..., \sum_j \tilde{\mathbf{A}}_{nj})$,;
$\mathbf{H}^{(0)} \leftarrow \mathbf{X}$;
for $k = 1$ *to* K **do**
    $\mathbf{W}^{(k)} \leftarrow$ the trainable weighting matrix for the k-th step;
    $\mathbf{H}^{(k)} = ReLU(\tilde{\mathbf{D}}^{-\frac{1}{2}}\tilde{\mathbf{A}}\tilde{\mathbf{D}}^{-\frac{1}{2}}\mathbf{H}^{(k-1)}\mathbf{W}^{(k)})$;
end
$\mathbf{Z} \leftarrow \mathbf{H}^{(K)}$;
return $\mathbf{Z}$;