Two paper about Weakly supervised learning

Presenter： Wentao Lei

Paper List	Time	Publish
WEAKLY-SUPERVISED BRAIN TUMOR CLASSIFICATION WITH GLOBAL DIAGNOSIS LABEL	2020	ISBI
WEAKLY SUPERVISED PROSTATE TMA CLASSIFICATION VIA GRAPH CONVOLUTIONAL NETWORKS	2020	ISBI

Content

- Background and motivation
- Problem formulation and setting
- Method
- Experiment

Weakly supervised learning

The cause of weak label：

1. Different level of annotation

Sample	Label
Image-wise	Video-wise
2D-Image	3D-Image

2. Different precision of annotation

Image-level labels vs Pixel-level labels

WEAKLY-SUPERVISED BRAIN TUMOR CLASSIFICATION WITH GLOBAL DIAGNOSIS LABEL

Yufan Zhou ⋆ *Zheshuo Li* ⋆ *Chunwei Ma* ⋆
Changyou Chen ⋆ Mingchen Gao ⋆ Hong Zhu † Jinhui Xu ⋆

⋆ Department of Computer Science and Engineering, University at Buffalo, SUNY
† School of Medical Information, Xuzhou Medical University

Background and motivation

Brain Tumor Classification on magnetic resonance images (MRI).

Only global diagnosis labels are given but not the slice-level labels.

We augment the label space by adding a label representing a healthy MRI slice

Basic setting:

Input: 3D MRI Image with global diagnosis label

Output: 8 types of tumors（+1 health in this paper）

Metrics: Accuracy and AUC

Weakly-supervised Learning

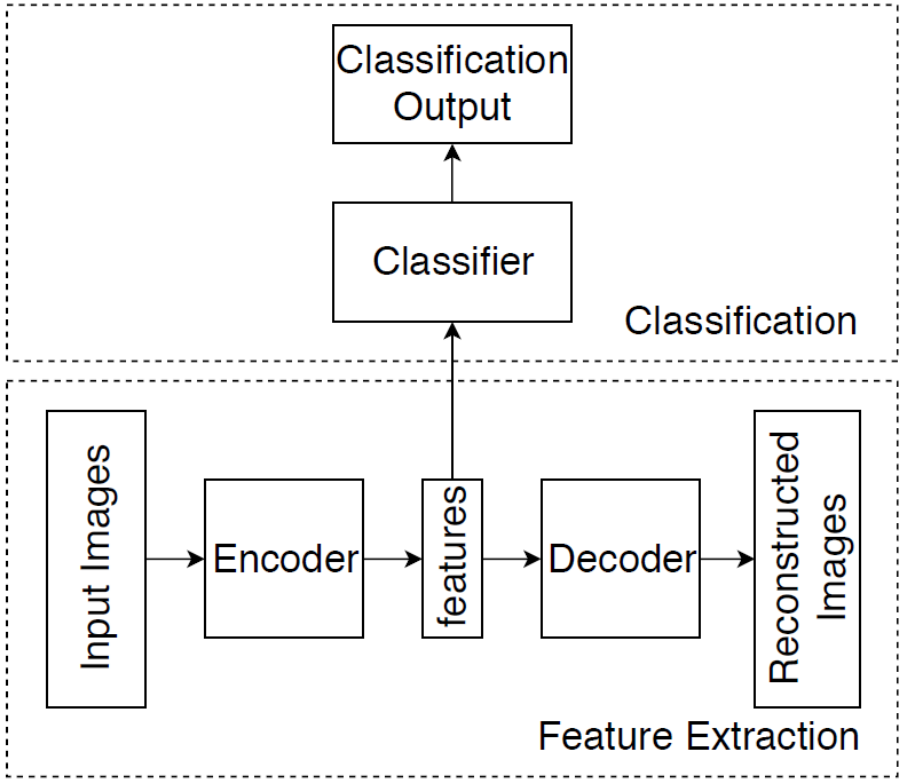
The waekly supervised objective function for MRI images:

$$\mathcal{L}(\mathbf{x}, \tilde{\mathbf{y}}, \mathbf{t}) \triangleq -\sum_{j=1}^{K} t_j \log \bar{y}_j + (\alpha \|\bar{y}_{1:K}\|^2 + \beta \bar{y}_{K+1}^2)$$

the objective function ensures that some slices are predicted as the ground true tumor type, but not all of them are expected to be categorized into the same class.

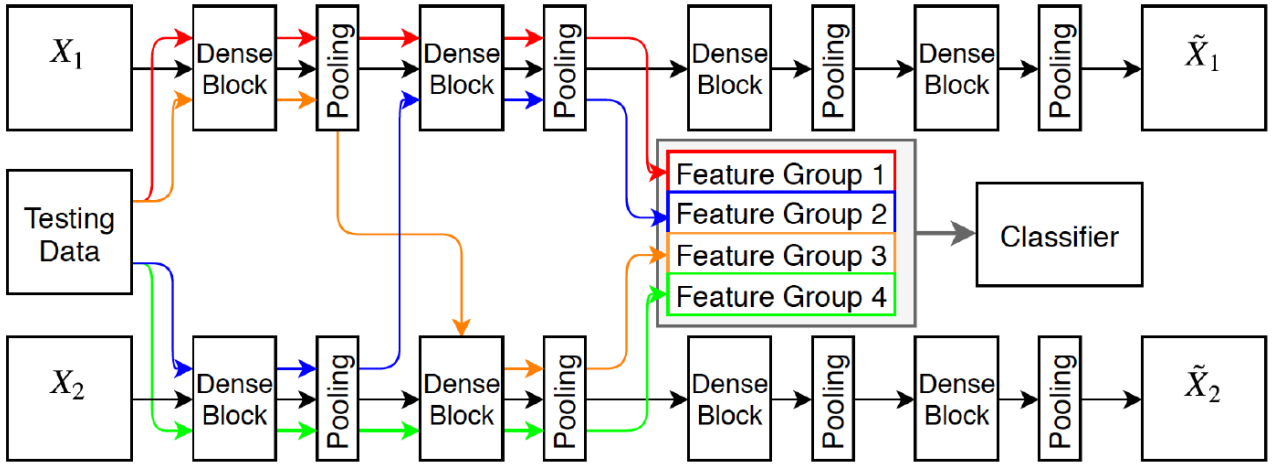
Feature Extraction via Autoencoder-Dropout

- An autoencoder is trained to extract features from the training data
- and then a classier is built based on the extracted features.



Multi-Autoencoders (MultiAE)

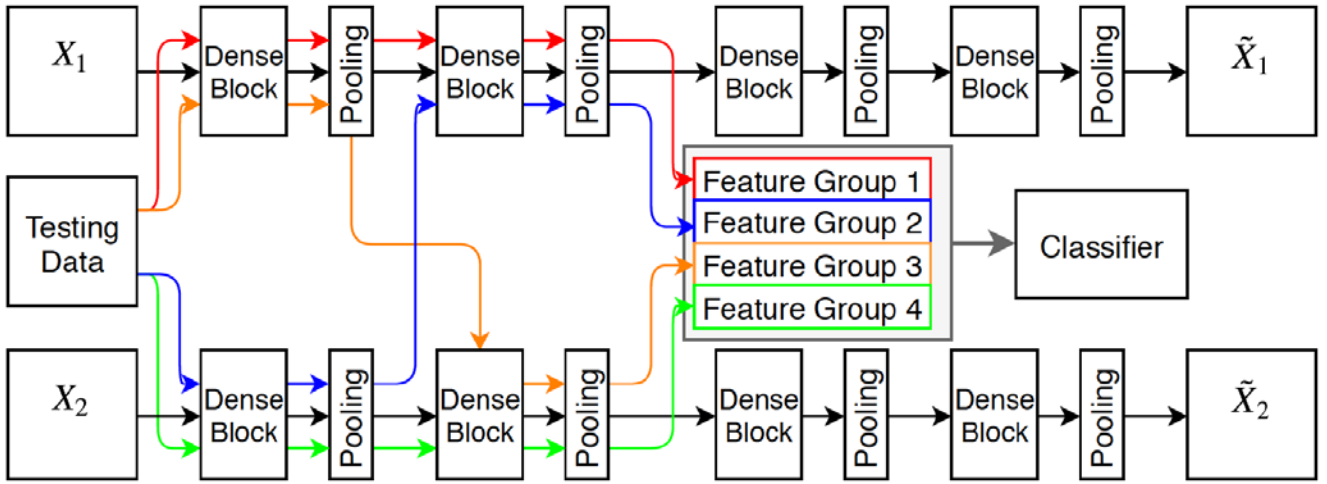
- Using multiple class-specic autoencoders to learn discriminative features, each corresponding to one class.
- Denote this model as MultiAE
- The features extracted from one autoencoder is called a feature group



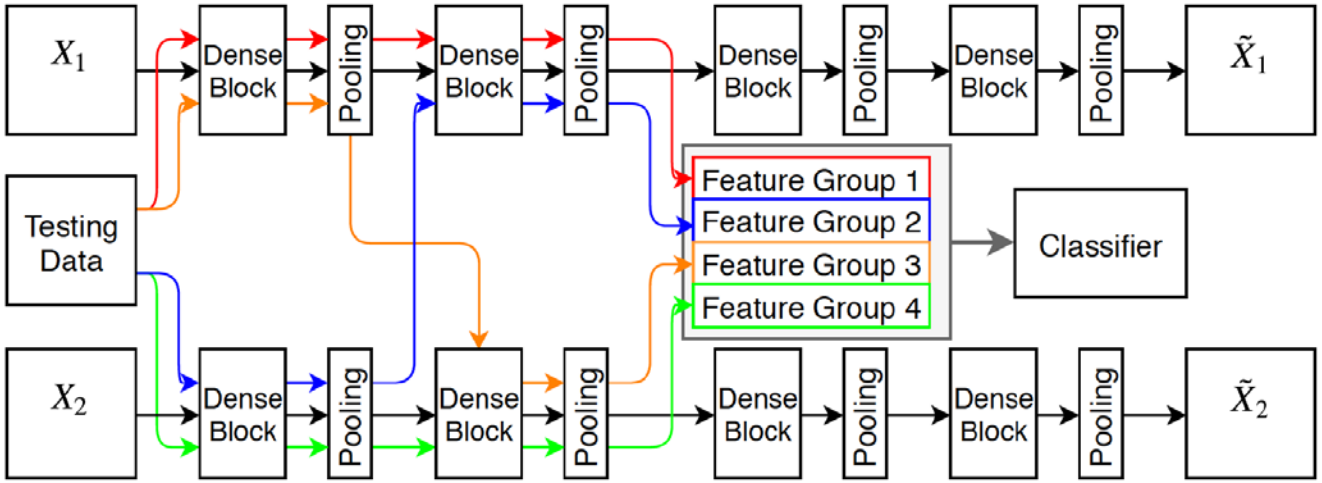
Random combination of autoencoders with dropout (MultiRAED)

We randomly pick paths of the dense blocks to learn features.

If we have m autoencoders with each having n dense blocks, there are m^n possible combinations for constructing the hierarchical feature representations



Feature Extraction via Autoencoder-Dropout



Every autoencoder consists of 3 dense blocks in encoder and decoder part respectively, and the classier consists of 2 dense blocks. All the dense blocks have 6 convolutional layers inside.

To prevent overtting, dropout with dropout rate 0.1

We randomly select 27 feature groups (out of $8^3 = 512$ possible feature groups) for each MR volume due to the memory limitation.

Experimental setting

- Dataset：It contains 8 tumor types with 295 patients. training set (72%), validation set (14%) and a test set (14%)
- 7-fold crossValidation
- Test on a public dataset which contains 3 tumor types and about 200 patients

Experimental Results and Discussion

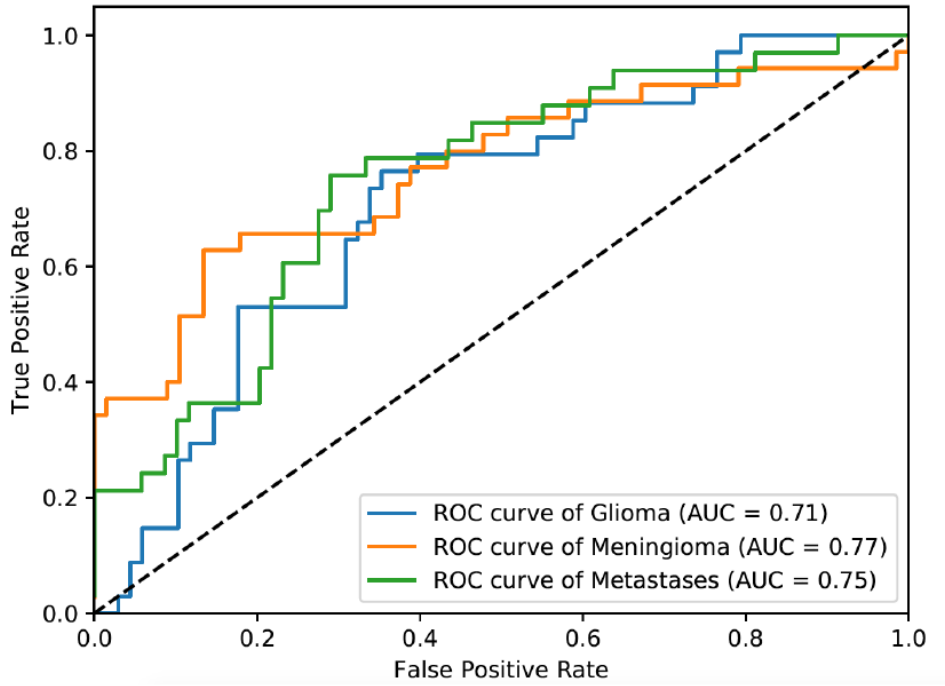


Fig. 3. ROC curve of 3-type tumor classification on our proprietary dataset, the figure is plotted based on one class vs. all other classes comparison.

Table 1. Classification accuracy on public dataset

2D-SingleAE	2D-MultiAE	2D-MultiRAED	[6]	[7]
89.62% ±3.22%	90.87% ±1.63%	**91.80%** **±2.80%**	91.28%	86.56%

Table 2. Accuracy on proprietary brain tumor dataset

Models	8-type	3-type
3D-DenseNet	38.61% ± 7.88%	55.00% ± 5.93%
3D-MultiRAED	48.06% ± 13.67%	63.33% ± 4.71%
Weakly	47.23% ± 6.23%	67.05% ± 7.09%
Weakly-MultiRAED	56.33% ± 4.89%	73.65% ± 3.65%
Supervised-MultiRAED	**57.13% ± 1.92%**	**73.95% ± 8.94%**

WEAKLY SUPERVISED PROSTATE TMA CLASSIFICATION VIA GRAPH CONVOLUTIONAL NETWORKS

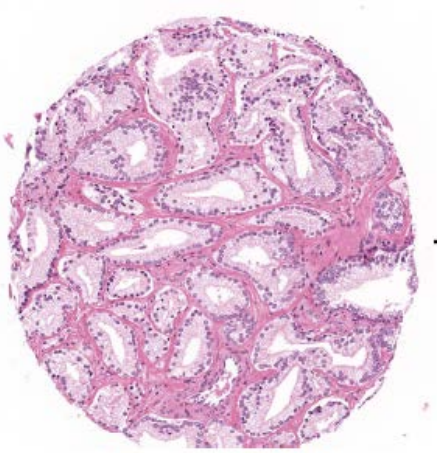
Jingwen Wang[1], *Richard J. Chen*[1], *Ming Y. Lu*[1], *Alexander Baras*[2], *Faisal Mahmood*[1]

[1]Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA
[2]Department of Pathology, Johns Hopkins School of Medicine, Baltimore, MD
jwang111@bwh.harvard.edu faisalmahmood@bwh.harvard.edu

Background and motivation

Challenges:
Previous Deep learning objective Gleason grading requires manual pixel-level annotation. (a pixel-level annotation mask suggesting the Gleason scores[1-10], Gleason scores>=6 means high risk)

We only have the image-level labels instead of the pixel level labels.



Prostate TMA

Basic setting:

Input: 2D Image with pixel-level annotation

Output: high-risk and low risk (binary classification)

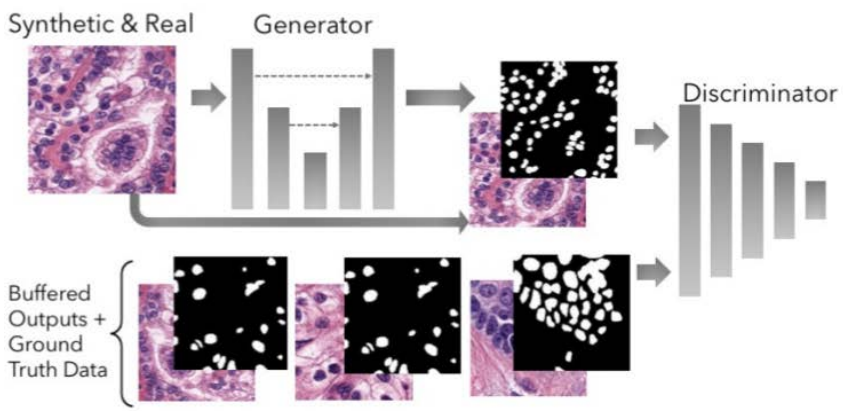
Metrics: Accuracy and AUC

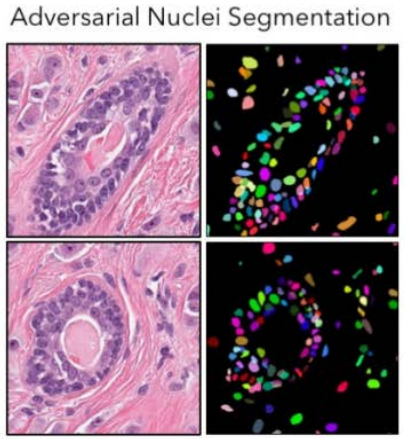
TMA Graph Construction

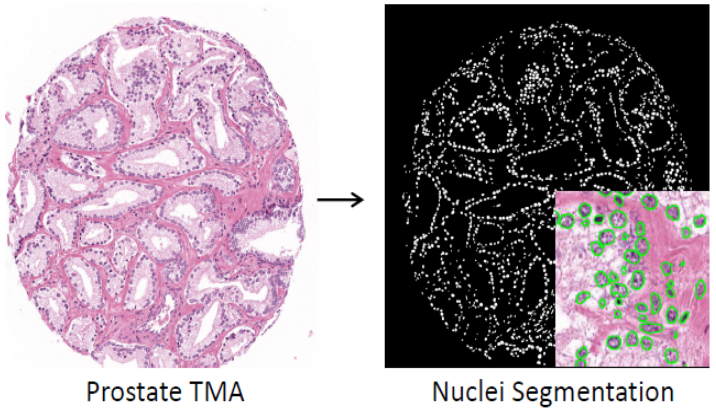
Nuclei Segmentation:

In order detect the nuclei robustly and clearly, we utilize conditional GANs (cGANs) We train our model for segmentation on the Multi-Organ Nuclei Segmentation dataset

We train our model for segmentation on the Multi-Organ Nuclei Segmentation dataset. This dataset, which consists of 21,623 nuclei in 30 images, comes from 18 different hospitals and include diverse nuclear appearances from a variety of organs like liver, prostate, bladder, etc







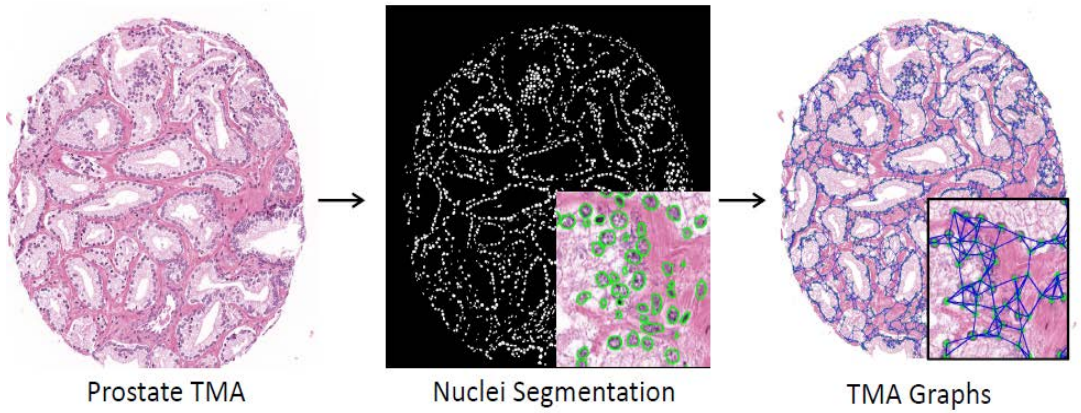
TMA Graph Construction

Nuclei Connection

To capture the architectural structure between neighboring nuclei.

Assume that nuclei that are close to each other are more likely to have interactions

Use K-nearest neighbors (KNN) algorithm, in which each nucleus is connected to its top 5 nearest neighboring nucleus if they are within a certain Euclidean distance (100 pixels in paper).



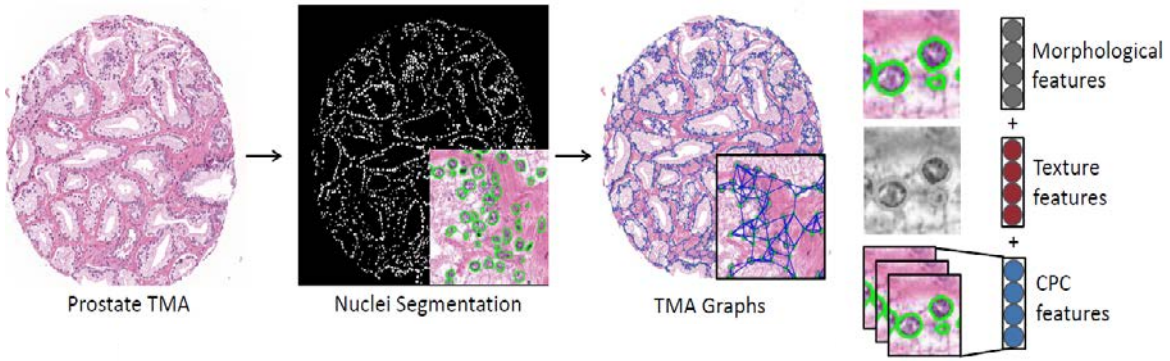
TMA Graph Construction

Nuclei Feature Extraction

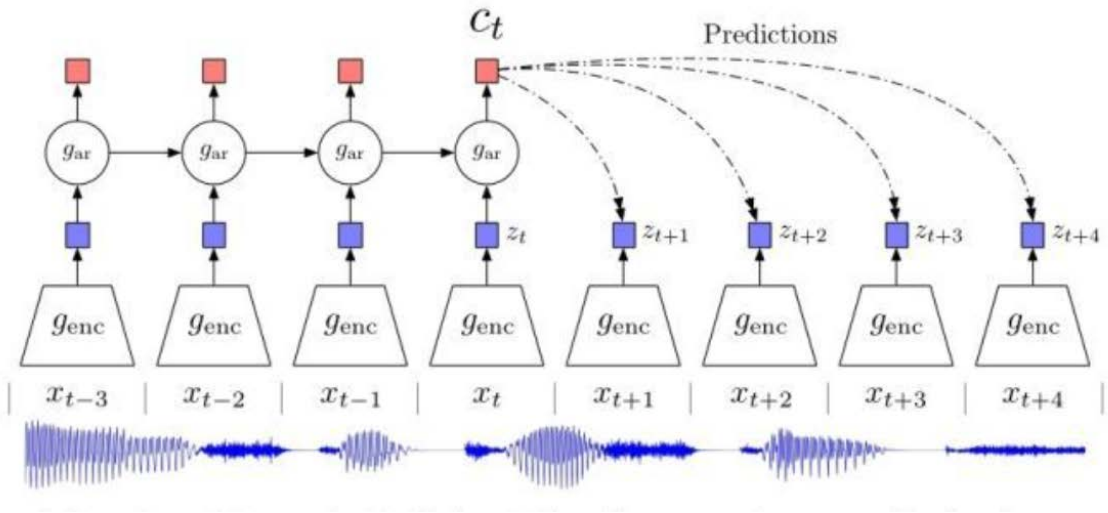
- Morphological Feature (area, roundness etc.)
- Texture Features (dissimilarity, homogeneity, energy and ASM based on the GLCM.)
- Contrastive Predictive Coding (CPC) Features

Based on the nuclei masks obtained using a cGAN, we then extract both morphological and texture features for each nucleus, along with features extracted from CPC-based self-supervise learning

We generate 8 morphological features, 4 texture features from GLCMs and 1024 features from CPC. Then, we concatenate them together. Finally we form a feature matrix V $\in R^{N_i \times F}$ where N_i is the number of nuclei in the graph and F is the number of features (1036 in our method)



CPC-based self-supervise learning



CPC utilizes a contrastive loss, through which the mutual information shared between the context c^t, the present, and future observations z_{t+k}; (k > 0) can be maximized

By minimizing the CPC objective in the latent space, we implicitly encourage learning of such shared high-level abstractions specific to prostate tissue, which one is unlikely to effectively capture by using features extracted through naive transfer learning from ImageNet

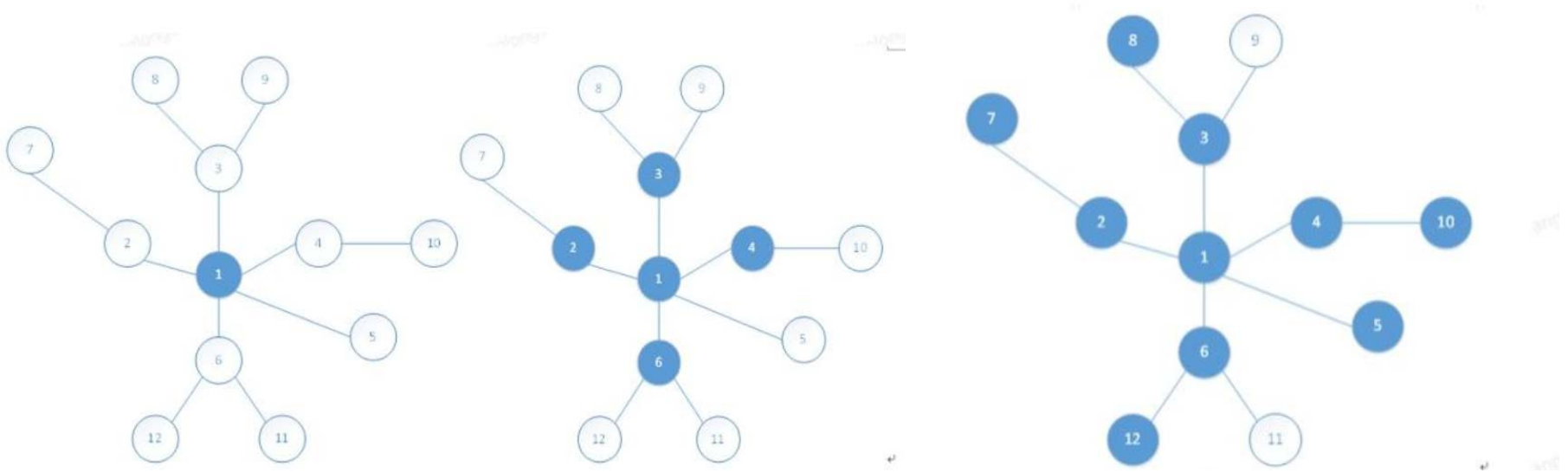
Graph Convolution Networks

The convolution and pooling operations can be defined as follows:

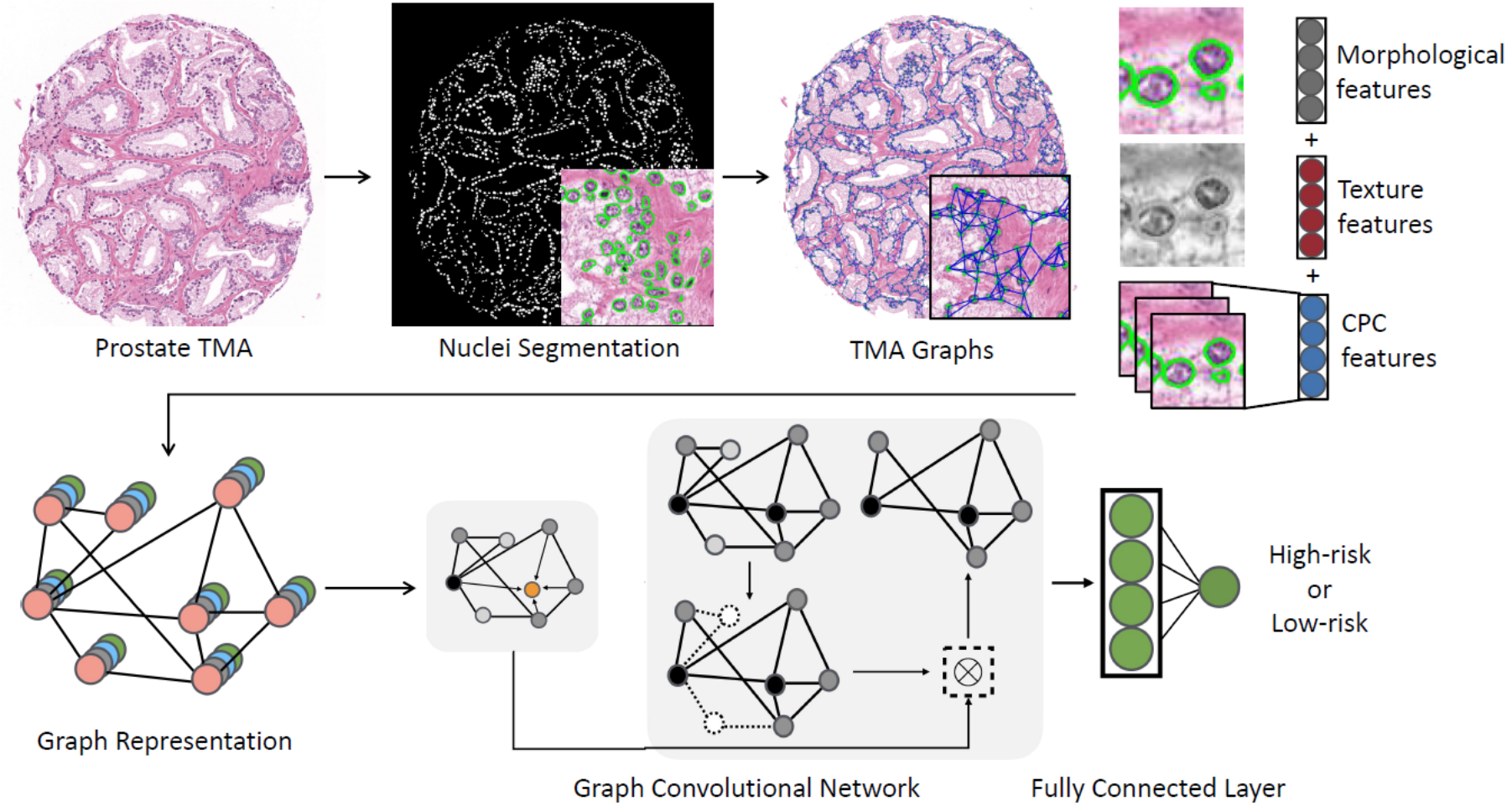
$$a_v^{(k)} = \mathbf{MAX}\left(\left\{\mathbf{ReLU}\left(W \cdot h_u^{(k-1)}\right), \forall u \in \mathcal{N}(v)\right\}\right)$$

$$h_v^{(k)} = W \cdot \left[h_v^{(k-1)}, a_v^{(k)}\right]$$

The feature vector for node v is denoted by $h_v^{(k)}$, The information from neighbor at the next iteration is represented by $a_{(k)}$



Framework


Prostate TMA
Nuclei Segmentation
TMA Graphs
Morphological features
+
Texture features
+
CPC features
Graph Representation
Graph Convolutional Network
Fully Connected Layer
High-risk
or
Low-risk

EXPERIMENTS AND RESULTS

- Image dataset is from 5 prostate TMAs, each containing 200-300 spots. There are 886 images in total

- The experiment is on 5 classbalanced splits where each splits contain 3498 crops for training, 388 crops for validation and 432 crops for testing.

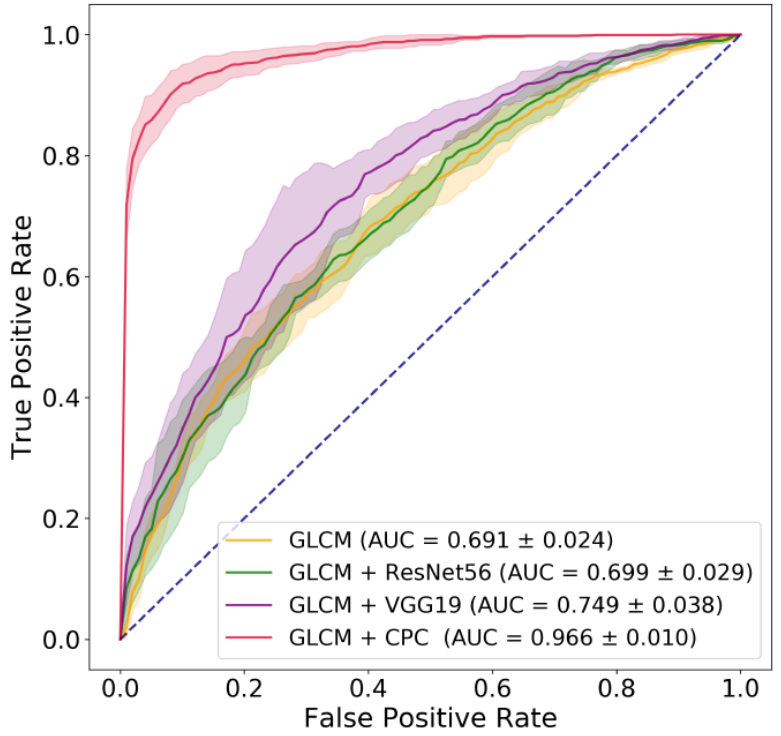


Table 1. A comparative analysis of prostate TMA classification using various models and features.

Model	**Accuracy** ↑	**AUC**↑
GCN + GLCM features	0.6299 ± 0.0391	0.6909 ± 0.0240
GCN + GLCM + Transfer Learning (ResNet56)	0.6412 ± 0.0181	0.6987 ± 0.0294
GCN + GLCM + Transfer Learning (VGG19) [17]	0.7194 ± 0.1192	0.7486 ± 0.0377
GCN + GLCM + CPC features (Proposed)	**0.8995 ± 0.0222**	**0.9659 ± 0.0096**

(Transfer learning means using the features generated through simple transfer learning from ImageNet)

Conclusion

If the level of annotation is higher than what the task requires，(i.e. the annotation is not accrate or precise) , we may need weakly supervised learning.

The solution is to make the model more robust to this weak label or generate a stronger feature (when there is no precise label like pixel-wise label)

Thanks for listening