CAMELYON17: METU-VISION TEAM

PN-STAGE CLASSIFICATION OF BREAST CANCER PATIENTS FROM WHOLE SLIDE IMAGES OF LYMPH NODES USING UPSAMPLING U-NET AND DECISION FUSION

Melih Çetin[1], Eren Halici[1], Mustafa Ümit Öner[1], Ugur Halici[1, 2]

METU VISION
Computer Vision and Intelligent Systems Research Laboratory,
[1]Department of Electrical and Electronics Engineering,
[2] Neuroscience and Neurotechnology, GSNAS
Middle East Technical University

ABSTRACT

The aim of the Camelyon17 challenge is to evaluate algorithms for automated detection and classification of breast cancer metastases in whole-slide images of histological lymph node sections. This paper explains the algorithm developed by the METU VISION team. In the proposed algorithm, a modified version of U-Net, called Up-sampling U-Net is used for segmentation. Camelyon16 data is used for training Up-sampling U-Net. Post processing is done for mask correction and then these masks are used for determining slide classes. The decisions are fused through a binary decision tree, which is optimized considering quadratic weighted Cohen's Kappa value calculated with respect to slide classes in Camelyon17 training set. pN stages of patients are decided according to the rules defined on slide classes. Artificial patients are created to validate the results.

Index Terms— metastasis detection, U-Net, decision fusion.

1. INTRODUCTION

The task in Camelyon17 is to determine a pN-stage for every patient in the test dataset. Five slides are provided for each patient (each slide corresponds to exactly one node).

The class of a single lymph node (a single slide) is decided according to the metastasis sizes as follows:

- Macro (Macro-metastasis): ≥2.0 mm
- Micro (Micro-metastasis): ≥ 200 μm (or more than 200 cells), AND < 2.0 mm.
- Isolated Tumor Cells (ITC): < 200 μm (or less than 200 cells)
- Negative (N0): Neither metastasis nor ITC

The pN stages of the patients are defined as follows:

- pN0: No ITC or micro-metastases or macro-metastases
- pN0(i+): Only ITCs found
- PN1mi: Micro-metastases found but no macro-metastases
- PN1: Metastases found in 1-3 lymph nodes (i.e. slides), at least one is a macro metastases
- PN2: Metastases found in 4-9 lymph nodes, at least one is a macro metastasis

Camelyon17 training data contains 100 labeled patients with 500 labeled slides. Test data contains 100 patients with 500 slides.

2. METU VISION ALGORITHM

The block schema of the algorithm submitted by METU-VISION team is given in Figure 1. The algorithm follows the pre-processing, segmentation, post-processing, single expert decision, decision fusion and pN stage decision steps. The details of each step are explained in the following sections.

2.1. Pre-processing

Considering the computing capability we have (several PC's and 3 workstations, two having Nvidia TITAN-X GPU cards and the other having Nvidia M5000), it is decided to use Level 2 images for training the U-Nets and extracting the masks in Level 4. However the quality of Level 2 images in WSI's was poor, therefore Level 0 images were down-sampled with anti-aliasing to Level 2 images.

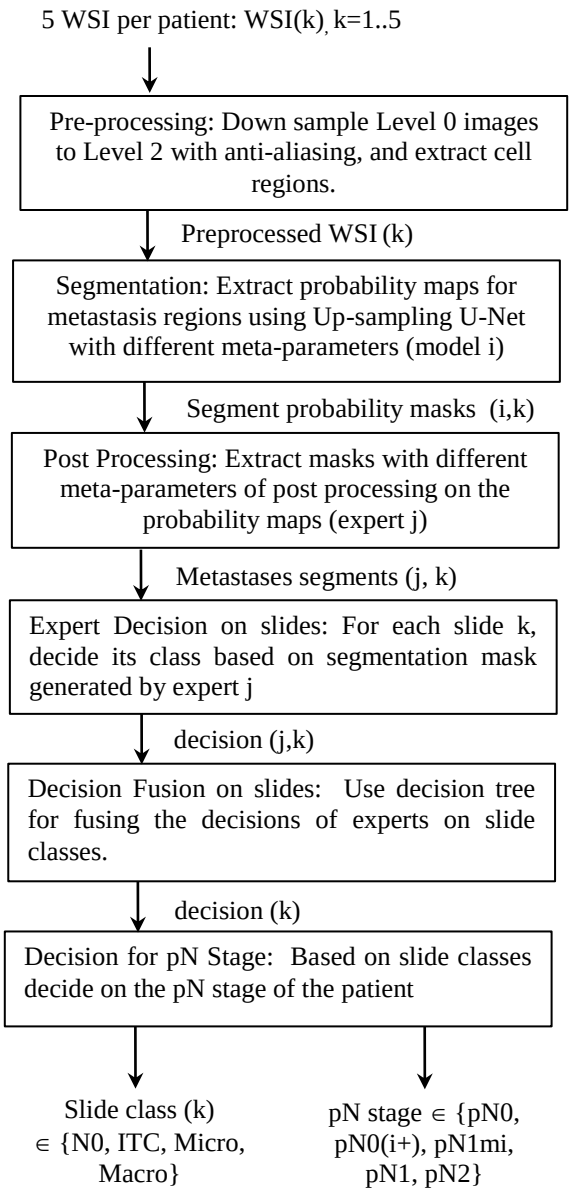

Figure 1. Block schema of METU-VISION algorithm

2.2. Segmentation

U-Net is a convolutional network for biomedical image segmentation proposed in [1]. The original U-Net structure is given in Figure 2.a. However, in this study, instead of using up-conv operation defined in the original U-net, we modified it as follows and called the resulting network. Up-sampling U-Net (Figure 2.b):

- Up-conv step is not used. Instead, up-sampling is applied on each feature map separately and then a convolution operation is applied in order to reduce the number feature maps into a quarter. Concatenation of feature maps (from the symmetric part of U-shape in the same layer) is applied after this convolution layer. This is the major modification.
- Probability maps for metastasis regions (instead of binary masks) are generated as outputs.

Due to memory limitations, instead of complete Level 2 image, patches of size NxN on cell regions are used as inputs to the U-net. The outputs of the U-net are combined back to obtain the complete probability mask of the Level 2 image.

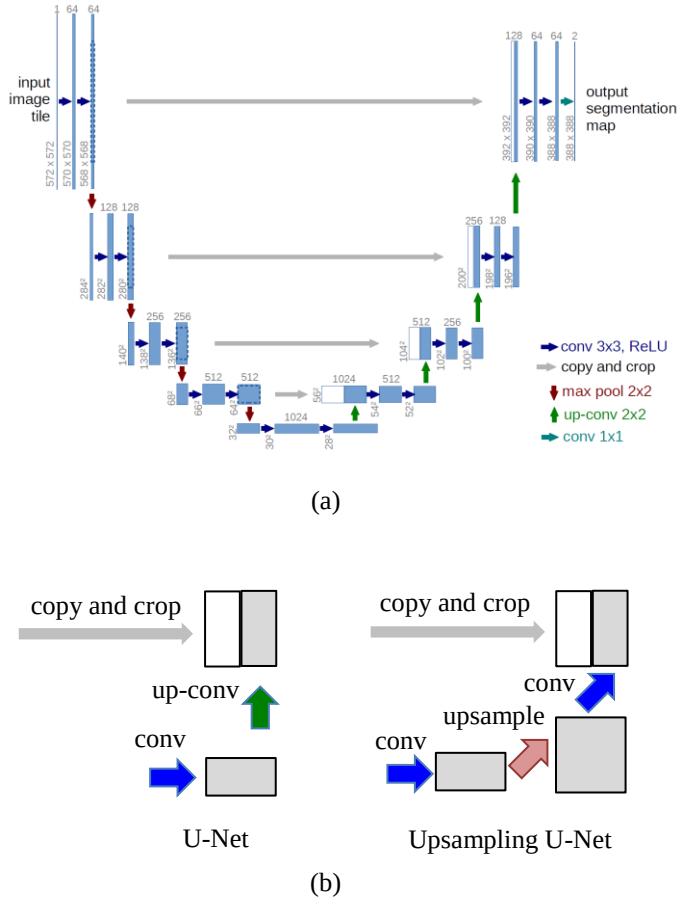

Figure 2. a) Original U-Net [1], b) Up-conv operation in U-Net is replaced by upsample and then conv operations Upsampling U-Net

Up-sampling U-Net is trained mainly using the Camelyon16 slides having label as tumor (20 of them are reserved for validation of segmentation step) and the slides with masks in the Camelyon17 (10 of them are reserved for validation). These slides are separated into NxN patches. The number of patches obtained from these slides depends on N.

Two models, with different U-Net meta-parameters are considered (see Table 1).

Table 1. Meta-parameters for the models

Model	Patch size (N pixels)	Up-down Layer count	Batch size	# of epochs	Augmentation	# of training patches	# of validation patches
m1	512	5	16	21	8	621k	74k
m2	252	4	16	78	1	478k	257k
m3	(p(m1)+p(m2))/2, pixel by pixel average						
m4	T(p(m1)) AND T(p(m2)), pixel by pixel AND						

2.4. Post-processing

In post processing the following steps are applied:

- Level 2 probability maps are down sampled to Level 4 masks, for faster post-processing.
- Thresholding: Probability maps are thresholded with T=0.5 in order to obtain binary masks. Mixed models are also generated by pixel-wise averaging (m3) the probability maps or by ANDding (m4) binary masks of m1 and m2 (see Table 2).
- Morphologic opening operations with structuring elements r1 $\in$ {0, 5, 10} are applied in order to remove small regions, where r1 = 0 means no opening operation is applied.
- Closing operations with r2 $\in$ {0, 5, 10, 15} are applied in order to combine close regions into a single region, where r2=0 means no closing operation is applied. If r1=0, then r2=0.

2.4. Single Expert Decision

Each combination of model (m#), opening radius (r1) and closing radius (r2) corresponds to a different expert. Experts are named with the convention e-m#-r1-r2.

By each expert e-m#-r1-r2, the corresponding Level 4 mask is considered. The major axis length of each connected component in the mask is extracted and the maximum of them is used in order to decide on the slide class.

2.5. Decision Fusion

For decision fusion a two-stage binary decision tree, which we call as 2S-DT, is used (Figures 3). In the first stage, expert e1 separates the samples into classes {N, P}, where N=N0 $\cup$ ITC and P=Micro$\cup$Macro. In the second stage, the samples classified as N in the first stage are separated into N0 and ITC classes by expert e2. The samples classified as either ITC, Micro or Macro by e2 are labeled as ITC. Similarly, the samples classified as P in the first stage are separated into Micro and Macro classes by expert e3. The samples classified as either N0, ITC or Micro by e3 are labeled as Micro.

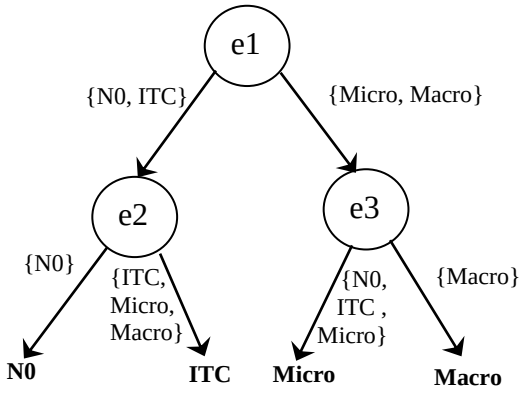

Figure 3. Binary 2 Stage Decision Tree,

The experts e1, e2 and e3 in 2S-DT are determined by a brute force search on the Camelyon17 training set. The expert combination maximizing the slide class based Cohen's Kappa ($\boldsymbol{\kappa_s}$) value on Camelyon17 training set is chosen for generating the results on the Camelyon17 test set.

2.6. Decision on pN Stage of Cancer

Once the classes for each slide of a patient are decided, the pN stage is decided as:

pN0: **#(**S $\in$ N0)=5, the number of slides in class N0 is 5
pN0(i+): **#(**S $\in$ ITC) $\geq$1 AND **#(**S $\in$ Micro$\cup$Macro)=0
PN1mi: **#(**S $\in$ Micro) $\geq$1 and **#(**S $\in$ Macro)=0
PN1: **#(**S $\in$ Macro) $\geq$1 and 1$\leq$**#(**S $\in$ Micro$\cup$Macro) $\leq$3
PN2: **#(**S $\in$ Macro) $\geq$1 and 4$\leq$**#(**S $\in$ Micro$\cup$Macro) $\leq$5

3. RESULTS

Slide (class) based and pN stage based quadratic weighted Cohen's Kappa values (κ_s and κ_p respectively) obtained by different experts on camelyon17 train set are presented in Table 3 and Table 4, respectively. Best κ_s (0.770) is obtained by the expert e-m3-10-15 and κ_p (0.780), which is obtained by the experts e-m4-10-5/10/15.

The experts e1, e2 and e3 in 2S-DT are determined by a brute force search on the Camelyon17 training set. The expert combinations maximizing κ_s value, which is 0.798, are presented in Table 5. For these combinations κ_p values are observed to be the same, κ_p= 0.796. Furthermore, an artificial data set having 500 artificial patients is created by selecting slides from Camelyon17 training set randomly. On

this artificial data set, the pN stage based Kappa value is found as $\kappa_{p\text{-}a}$= 0.787

Table 3. Slide class based Kappa **(κ_s)** values obtained by single experts on Camelyon17 training set (Expert naming convention is e-m#-r1-r2)

r1-r2	Kappa (κ_s)			
	m1	**m2**	**m3**	**m4**
0, 0	0.315	0.350	0.543	0.682
5, 0	0.364	0.467	0.715	0.750
5, 5	0.354	0.437	0.706	0.752
5, 10	0.345	0.420	0.706	0.747
5, 15	0.342	0.407	0.697	0.748
10, 0	0.474	0.685	0.766	0.763
10, 5	0.475	0.683	**0.770**	0.767
10, 10	0.471	0.671	0.763	0.765
10, 15	0.473	0.671	0.764	0.768

Table 4. pN stage based Kappa **(κ_p)** values obtained by single experts on Camelyon17 training set (Expert naming convention is e-m#-r1-r2)

r1-r2	Kappa (κ_p)			
	m1	**m2**	**m3**	**m4**
0, 0	0.564	0.608	0.664	0.728
5, 0	0.598	0.650	0.673	0.750
5, 5	0.570	0.659	0.662	0.757
5, 10	0.564	0.611	0.664	0.737
5, 15	0.537	0.604	0.664	0.726
10, 0	0.634	0.711	0.728	0.774
10, 5	0.634	0.711	0.737	**0.780**
10, 10	0.591	0.689	0.715	**0.780**
10, 15	0.591	0.689	0.715	**0.780**

Table 5. 2S-DT decision fusion Kappa values for of experts on Camelyon17 training set and on an artificial patient set.

2S-DT Experts		Kappa				
		Alone		Decison fusion		
		κ_s	κ_p	κ_s	κ_p	$\kappa_{p\text{-}a}$
e1	e-m4-10-5	0.770	0.737	0.798	0.796	0.787
e2	e-m3-10-0	0.766	0.728			
	e-m3-10-5	0.770	0.737			
	e-m3-10-10	0.763	0.715			
	e-m3-10-15	0.764	0.715			
e3	e-m2-5-0	0.420	0.611			

4. CONCLUSSIONS

For single expert decisions, while the best κ_s value (κ_s=0.770) was obtained by the expert e-m3-10-15, the best κ_p value (κ_p=0.780) was obtained by the experts e-m4-10-5/10/15. It is observed that fusion of decisions of experts e1= e-m4-10-5, e2=e-m3-10-x and e3=e-m2-5-0 improves the best κ_s value to κ_s =0.798. Furthermore, it is observed that κ_p value for this combination also improved, which is κ_p=0.796. Among the experts e-m3-10-x, e2 is set to e2=e-m3-10-10 for METU VISION submission to Camelyon 17 challenge.

For the Camelyon17 test set, the slide classes and patient stages are determined and submitted using the decision fusions of these experts.

11. REFERENCES

[1] O. Ronneberger, P. Fischer, and T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation, MICCAI 2015