TIGER Challenge

Description of the Developed Algorithm

Team VUNO

TEAM

- Team: VUNO
- Members:
  - Sungduk Cho
  - Dong-Hee Kim
  - Hyungjoon Jang
  - Chanmin Park
  - Kyungdoc Kim
- Algorithm: TIGER Algorithm LB2 (Version 55c49c9e-4216-4142-b1c8-f5d85781add3)
- Email: kyungdoc.kim@vuno.co

DATA

- Preprocessing

Task	MPP	Patch Size	Others
Segmentation	1.0	512x512	• Adjust means of saturation & value (luminance) on tissue pixels if they exceed preset numbers to alleviate color variance between individual slides. (Mean saturation > 50, mean value > 150) • Sample slightly wider (> 256μm) regions than the given ROI bounds with null paddings on label masks to guide edge pixels with sufficient neighboring pixel information.
Detection	0.5	> 300x300 images → 160x160 patches < 300x300 images → patches (no crop)	
TILs	4.0	Segmentation: WSI	• Read tissue mask at level 3 to check mask size and classify task.
	0.5	Detection: 2048x2048	

- Used datasets

Dataset	Detail
WSIROIS	• Train segmentation and detection models.
WSITILS	• Use as pseudo-label to train segmentation model. • Use to estimate the constant factor to predict TILs score.
WSIBULK	• Not used.

- Data split

Task	Detail
Segmentation	• Divide into 5-fold. • Split by the slide to gain meaningful validation results (all the patches generated from the same slide are assigned to same fold).
Detection	
TILs	• Sample few WSI to estimate the constant factor used in TILs score prediction.

METHOD

- Model

Task		Architecture	Detail
Segmentation		U-Net with EfficientNet-B2/B0 encoder	• ImageNet-Noisy Student pretrained • SCSE attention module as decoder • Ensemble 5 models
Detection		YOLOv5 m	• COCO pretrained • Ensemble 7 models
TILs	Segmentation	U-Net with EfficientNet-B0 encoder	• Train segmentation model on mpp=4 (instead of mpp=0.5). • ImageNet-Noisy Student pretrained • SCSE attention module as decoder • Ensemble 5 models
	Detection	Same as detection model	

- Training procedure

Task		Detail

Segmentation	Common factors	• Reduce classes into 3: stroma, tumor, rest.
		• Optimize 5-fold cross validation pipelines independently by observing ts_dice (official evaluation metric) over validation set.
		• Select each optimal epoch from 5-fold cross validation pipelines.
		• Augmentations:
		  ○ Random scale (scale limit = 0.05)
		  ○ Elastic transform (alpha=20, sigma = 10, alpha affine=10)
		  ○ Rotate
		  ○ Flip
		  ○ Color jitter (brightness=0.1, contrast=0.1, saturation=0.1, hue=0.08)
		  ○ Jpeg compression (quality lower=20, quality upper=90)
		  ○ Normalize
		• CosineAnnealingWarmRestarts LR scheduling
		• Layer-wise learning rate
		  ○ Encoder base LR=1e-3

		○ Decoder / Head base LR=2e-3 • Maximum 300 epochs • *Loss = (0.09 * CrossEntropyLoss) + (0.91 * DiceLoss(tumor/stroma only))* • Lamb optimizer • Stochastic weight averaging with 10 epochs and LR=2e-4 from 120 epochs
	Training base* models	• Feed WSIROIS dataset. • Mean ensembles with 5 different cross validated models' probability outputs.
	Training self-pseudo-curriculum models	• Generate pseudo-labels by estimating WSIROIS dataset with the previous base* models. ○ Use 2 types of pseudo-labels. ▪ Sole • No additional rules applied. ▪ Consensus • Compare pseudo-labels to the given GT. • Prune the mismatched area and assign ignore index. • With probability outputs of the pseudo-labels, compose a self-pseudo-curriculum by applying 5 different reliability thresholds to confidence score maps as [90%, 80%, 70%, 60%, 50%]. ○ Assign ignore index to the pixel whose confidence score is below the corresponding reliability threshold. • Feed self-pseudo-curriculum dataset with curriculum epochs setup as among ([0, 10, 30, 60, 100], [0, 10, 20, 40, 80], [0, 10, 20, 30, 40]). • Mean ensembles with 5 different cross validated models' probability outputs.
	Training semi-supervised models (applied on mpp4*** models only)	• Sample several patches per slide from peritumor area. • Generate pseudo-labels by estimating WSITILS dataset with the aforementioned self-pseudo-curriculum** models. (This time, only 'Sole' type pseudo-labels are generated.) • Take the same procedures (2~4) of the aforementioned self-pseudo-curriculum** model training.
Detection	Common factors	• 5-fold cross validation • Augmentations: HSV, shear, flip, mosaic, mixup, Blur, CLAHE • Linear LR scheduling • AdamW, Adam and SGD optimizers • Bayesian hyperparameter optimization with early stopping and maximum 500 epochs • Monitor mAP, FROC, Recall at Precision 0.95 while training.
	Training Semi-supervised learning	• Generate pseudo-label by combining segmentation and detection results in stroma and rest region.

		• Provide pseudo-labeled data from segmentation images of WSIROIS. • Then, train 160x160 patch image cropped with pretrained model with training set.

TILs		• Extract patches with high peritumoral stroma area:
		  ○ Run segmentation on WSI at mpp=4 w/ segmentation model trained on mpp=4.
		  ○ Get tumor boundary points from segmentation mask.
		  ○ Sample 50 peritumoral patches where center point is tumor boundary point
		  ○ Dilate tumor region and overlap with stroma region to get peritumoral stroma region. Width of peritumoral stroma band is around 100μm^2.
		  ○ Get top 20 patches with high peritumoral stroma area.
		• Get detected lymphocytes cells to peritumoral stroma area ratio for each patch:
		  ○ Run detection on each patch at mpp=0.5.
		  ○ *average of patch ratios = (Σ (detected lymphocyte cells / peritumoral stroma area)) / number of patches*
		  ○ *accumulated ratio = (Σ detected lymphocyte cells) / (Σ peritumoral stroma area)*
		• Calculate pseudo-TILs score:
		  ○ *pseudo-TILs score = (average of patch ratios + accumulated ratio) / 2*
		• Compute pseudo-TILs score for WSITILS data.
		• Based on pseudo-TILs scores, estimate constant value (i.e., a in y = ax) using TILs score annotation.

- Post-processing steps

Task	Detail
Segmentation	• Suppress tumor class probability outputs over the region where tumor class vs. rest class are briskly competing against each other (among pixels whose confidence score is lower than 80% after applying temperature=1.5). ⇒ Reduce false positives of tumor predictions. • Suppress tumor predictions in small size. (< 256μm^2) This helped reduce false positives of tumor predictions. • Apply different tissue mask, which is distinct from the given tissue mask, selectively on tumor class prediction. ⇒ Reduce false positives of tumor predictions.
Detection	• Implement Center-NMS based on center point distance (Existing NMS works based on IOU). • The point of our postprocess scheme is to increase the confidence of definite points and lower the confidence of ambiguous points. • During Center-NMS, classify points as number of detectors found that the point and adjust the confidence by it. (i.e., decrease points' confidence when number of detectors detect that point are less than 2.) • To focus on the stroma area, delete or lower confidence which in tumor or rest class area by combining the segmentation result.
Task selection (L1 or L2)	• To determine which pipeline to run (either L1 or L2), check if the tissue mask size is within the L1 mask size limit or not.

- TIL Score
  - Run segmentation and detection then compute pseudo-TILs score based on number of lymphocyte cells to peritumoral stroma area ratio. (Refer to equations in TILs training procedure).
    - average of patch ratios = (Σ (detected lymphocyte cells / peritumoral stroma area)) / number of patches
    - accumulated ratio = (Σ detected lymphocyte cells) / (Σ peritumoral stroma area)
    - pseudo-TILs score = (average of patch ratios + accumulated ratio) / 2
  - Multiply the found constant factor for scaling.
  - TILs score = constant factor × pseudo-TILs score
  - Clip TILs score from 1 to 95.