Supplementary Materials
The members of the GoldenPass team were affiliated with Graduate School of Artificial Intelligence in Korea Advanced Institute of Science and Technology (KAIST) and Knowledge of AI Lab in NCSOFT company. The members of the MediTrain team were affiliated with the Medical Science Research Center in Ansan Hospital of Korea University. Lastly, the members of the DRM team were from the School of Electrical Engineering in KAIST.
**S1 Table.** Algorithm descriptions and hyper-parameters

| Team       | GoldenPass                      | MediTrain                      | DRM                   |
|------------|---------------------------------|--------------------------------|-----------------------|
| Architecture | FPN+ResNet-50 encoder (model) | FPN+VGG-16 encoder (model) | Inception v4, SVM (model) |
|            | BCE+Jaccard (loss)              | BCE (loss)                     | BCEWithLogitsLoss (loss) |
| Input size (batch size) | 256×256×3 (16) | 768×1,920×3 (16) | 299×299×3 (16) |
| Slide layer level | Level 4 for AMC                | Level 4                        | Level 4 |
|             | Level 3 for SNUBH               |                                |          |
| Optimization (learning rate) | Adam (lr=1e-3)                | RMSProp (lr=4e-4)              | Adam (lr=3e-3) |
| Decay rate | Step decay when validation loss does not improve | None | Step decrease by 1/3 for every 4 epochs |
| Augmentation real-time | Random horizontal and vertical flip | Random horizontal and vertical flip | Random horizontal and vertical flip |
|            | Random rotation                 | Random rotation (±45°)          | Random rotation (±90°, 180°) |
| Pre-processing | ROI extraction to generate patches for training | K-fold cross-validation: 5-fold | Thresholding tumor (> 15%) |
|             | Rescale based on the magnification level of original slide | Fine-tuning on VGG-16 with ImageNet pre-trained weights | L2 norm |
|             | Stain normalization: Vahadane’s method | Down-sampling for resizing to specific dimension with | |
|             | Otsu thresholding to discriminate the cell areas | Lanczos interpolation | |
| Post-processing | Resizing the heatmap with different scales and thresholds | Averaging the output predictions from each fold | Generation of heatmap from patch classifier |
|             |                                 |                                 | Extraction of major axis, minor axis from heatmap map |
|             |                                 |                                 | |
| Inference for confidence | Max probability value from the heatmap | CNN output confidence | SVM classification from the extracted features |
| Output threshold | Youden’s J Statistic | Youden’s J Statistic | Youden’s J Statistic |

BCE, binary cross entropy; CNN, convolutional neural networks; DRM, DeepRunningMachine; FPN, feature pyramid network; ROI, region of interest; SVM, support vector machine.