H2G-Net: A multi-resolution refinement approach for semantic segmentation of gigapixel histopathological images

André Pedersen[1,2], Erik Smistad[3,4], Tor V. Rise[1,5], Vibeke G. Dale[1,5], Henrik S. Pettersen[1,5], Tor-Arne S. Nordmo[6], David Bouget[4], Ingerid Reinertsen[3,4], and Marit Valla[1,2,5,7]

[1] IKOM, NTNU, [2] Clinic of Surgery, St. Olavs hospital, [3] ISB, NTNU [4] SINTEF, NTNU, [5] Dept. of Pathology, St. Olavs hospital, [6] IFI, UiT, [7] Clinic of Laboratory Medicine, St. Olavs hospital

Objectives and contributions

Goal: Accurate and rapid segmentation of breast cancer in histopathological images.

Clinical motivation

Existing solutions are:

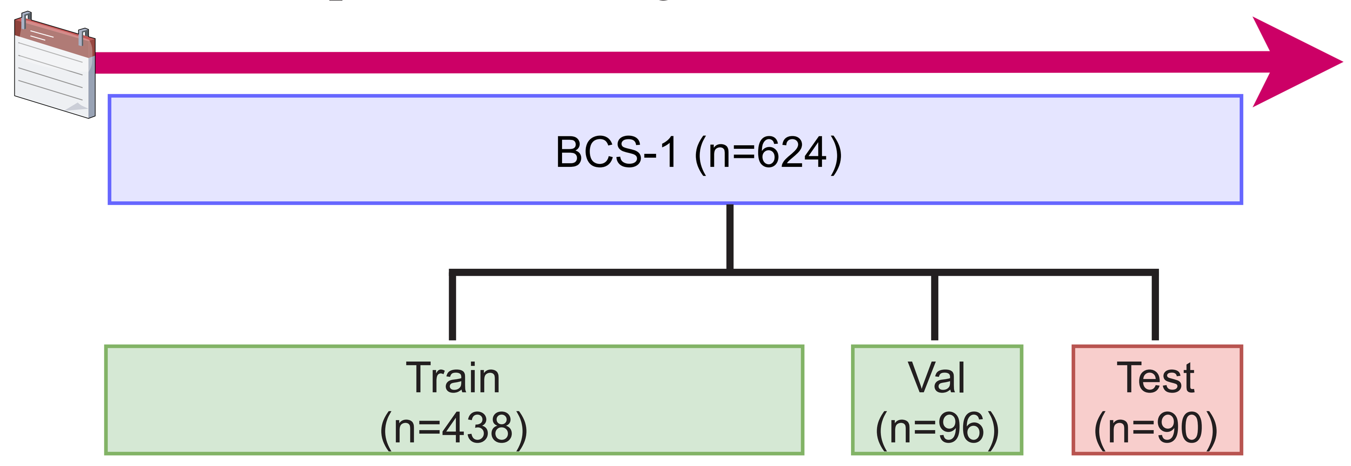
- Too computationally expensive or slow for routine usage.
- Not open-access or user-friendly.
- Do not efficiently use high and low-resolution information.

Key Contributions:

- A simple, rapid, and efficient model for automatic breast cancer segmentation.
- A two-stage, multi-scale convolutional neural network, H2G-Net.
- A novel framework for memory efficient, hierarchical-balanced, random patch sampling.

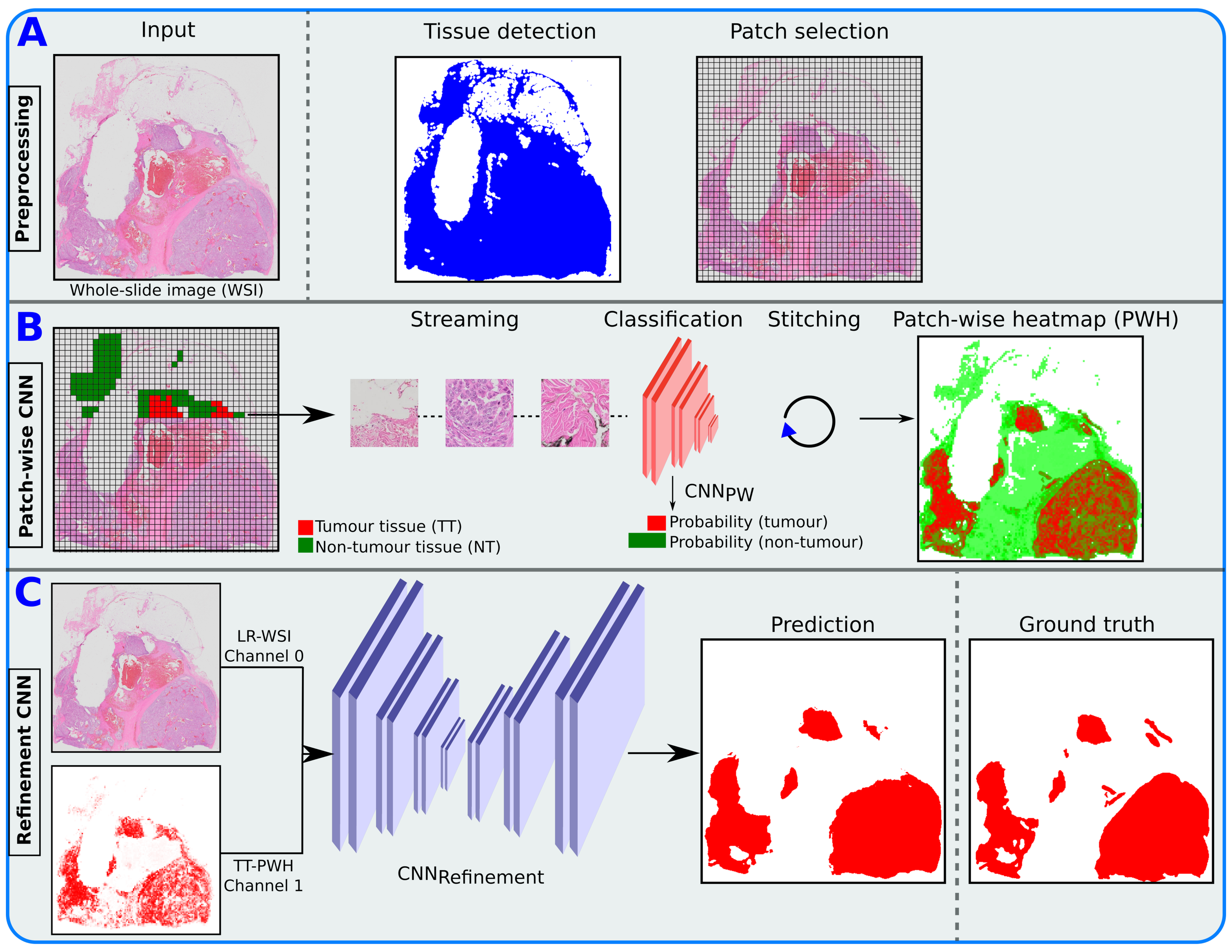
Dataset

- Ground truth: 624 whole slide images (WSIs) of breast cancer marked manually by pathologists.
- Norwegian breast cancer cohort, BCS-1
- WSIs were split into training, validation, and test sets.

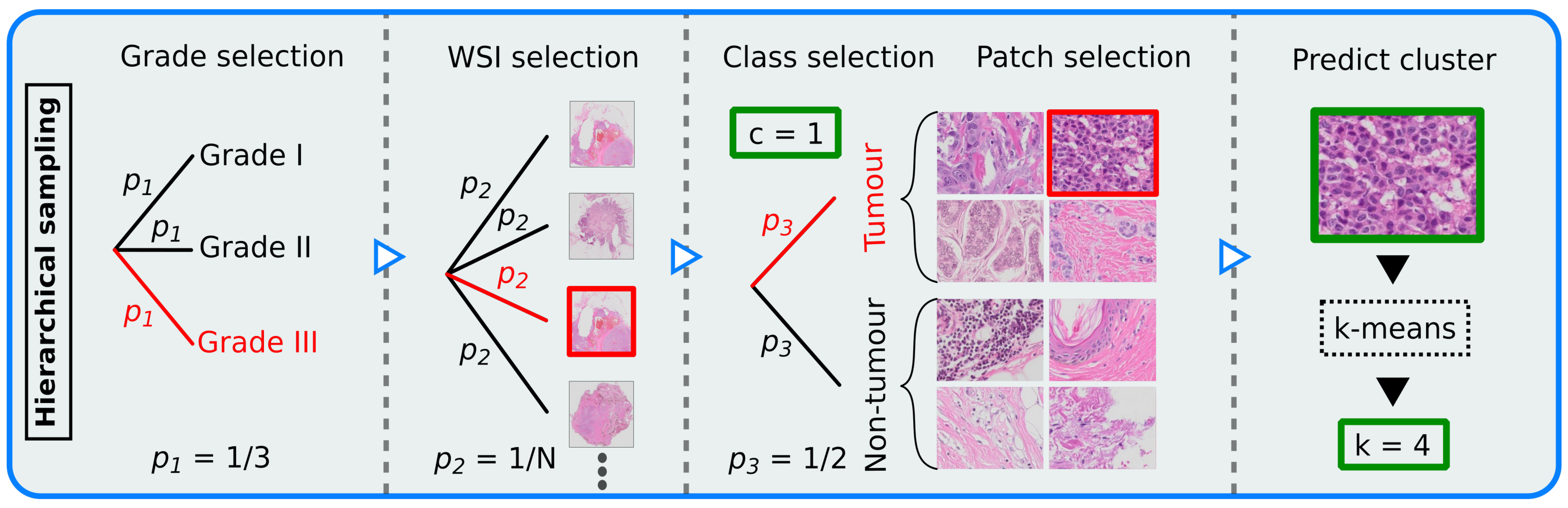


Methodology

Inference pipeline of H2G-Net:



Hierarchical sampling scheme:



Loss function for detection stage:

$$\mathcal{L}_{\text{CWCE}} = -\frac{1}{K_b}\sum_{k=1}^{K_b}\sum_{c=1}^{C}\sum_{i=1}^{B} 1(q_{i,k} = k) y_{i,c} log(p_{i,c})$$

- $\mathcal{L}_{CWCE}$: Cluster-weighted categorical cross-entropy loss
- K_b: Cluster frequency for a given mini batch of size B
- y: Current ground truth class of a total number of classes C
- p, q: Class and cluster prediction for a given class c and cluster k

Experiments & Results

Quantitative Results on BCS-1 test set:

	Designs	Recall	Precision	DSC
(1)	**Otsu**	0.990 ± 0.027	0.534 ± 0.200	0.669 ± 0.179
(2)	**UNet-LR**	0.931 ± 0.113	0.851 ± 0.165	0.874 ± 0.128
(3)	**Inc-PW**	0.881 ± 0.118	0.909 ± 0.099	0.887 ± 0.089
(4)	**Mob-PW**	0.879 ± 0.123	0.907 ± 0.100	0.885 ± 0.094
(5)	**Mob-KM-PW**	0.853 ± 0.124	0.909 ± 0.097	0.872 ± 0.092
(6)	**Mob-PW-AGUNet**	$\mathbf{0.954 \pm 0.066}$	0.909 ± 0.097	0.927 ± 0.072
(7)	**Mob-PW-DAGUNet**	0.942 ± 0.075	0.922 ± 0.091	0.928 ± 0.072
(8)	**Mob-PW-DoubleUNet**	0.949 ± 0.073	0.919 ± 0.093	0.929 ± 0.074
(9)	**H2G-Net**	0.944 ± 0.074	$\mathbf{0.929 \pm 0.088}$	$\mathbf{0.933 \pm 0.069}$

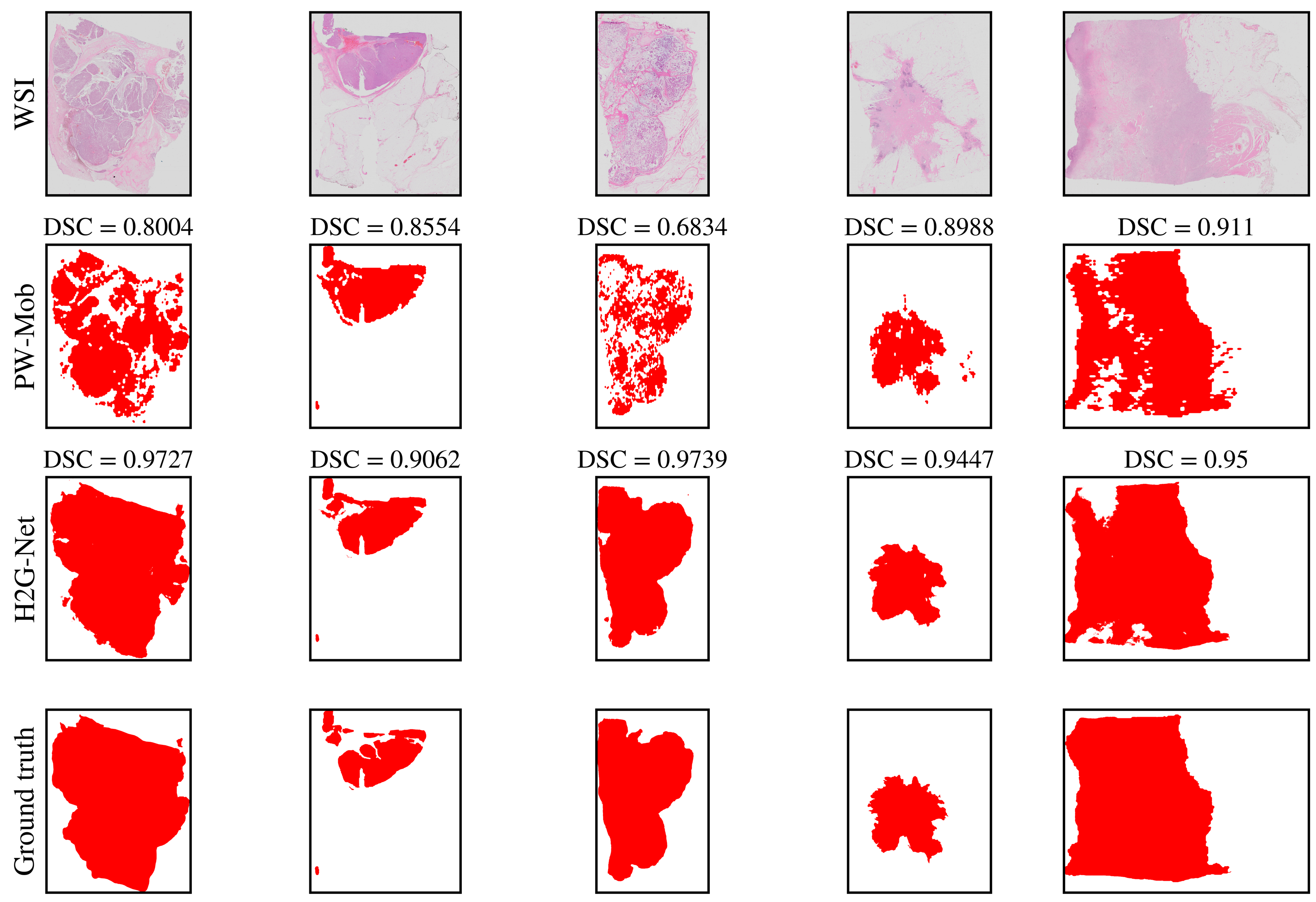
- An ablation study was conducted to find the most impactful components of the pipeline. The numbers represent performance metrics (the higher the better).
- LR: low-resolution, PW: patch-wise, Inc: InceptionV3, Mob: MobileNetV2, DSC: dice similarity coefficient

Runtime measurements (in seconds) using H2G-Net:

	Patch-wise	Refinement	Total
OpenVINO	57.32 ± 0.20	0.75 ± 0.01	58.07 ± 0.20
TensorRT	39.88 ± 0.62	0.38 ± 0.00	40.26 ± 0.62

- Runtime measurements were conducted on a representative $\times 400$ WSI, using both the OpenVINO (CPU) and TensorRT (GPU) inference engines.

Qualitative Results on BCS-1 test set:



Published article:

Paper & Code & Model

This work was supported by The Liaison Committee for Education, Research and Innovation in Central Norway; The Joint Research Committee between St. Olavs hospital and the Faculty of Medicine and Health Sciences, NTNU (FFU); The Cancer Foundation, St. Olavs hospital, Trondheim University Hospital; and The Clinic of Laboratory Medicine, St. Olavs hospital, Trondheim University Hospital.