Medical Project

Detecting Cancer Metastasis on Gigapixel Pathology Images

By Jiawen Huang (jh4179)

CAMELYON16 challenge

https://camelyon16.grand-challenge.org/Data/

400 WSI (whole slide images) collected independently from two medical centers in the Netherlands.
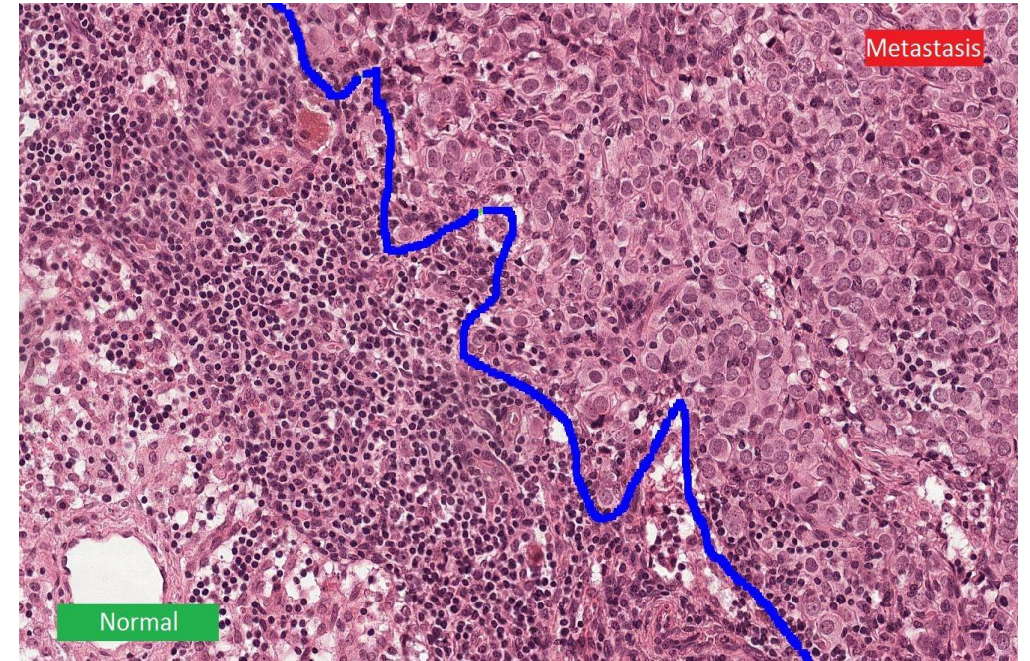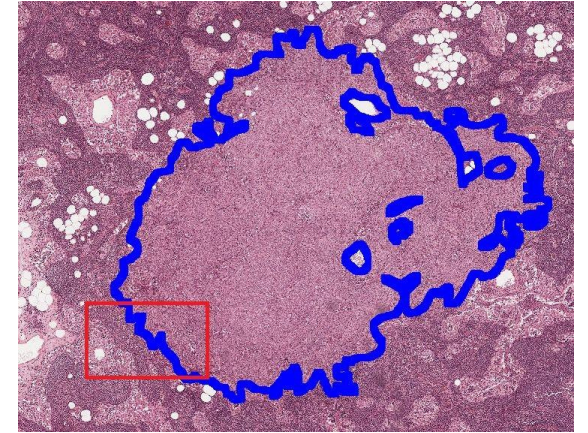
- About 600GB.

I select 6 images as my training images.
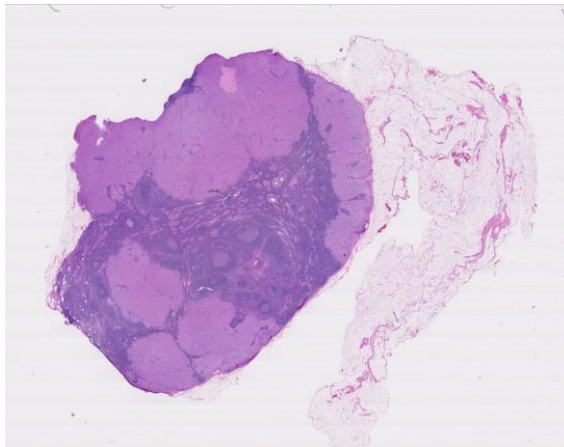- 075, 078, 081, 084, 091, 096

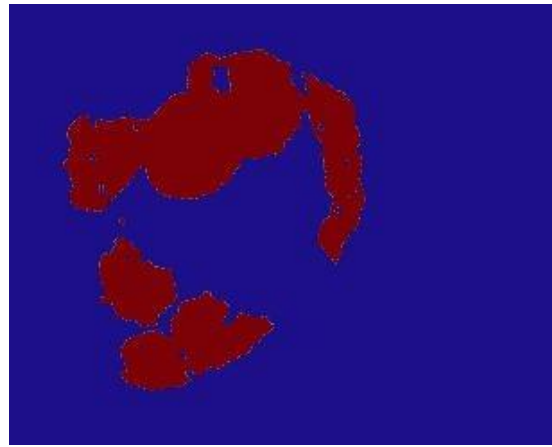And 3 images as my testing images.
- 016, 101, 110
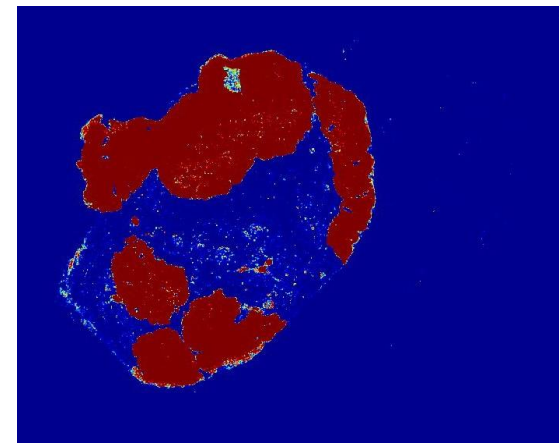
Project: Develop a tool to assist physicians

- Given a collection of training data, develop a model that outputs a heatmap showing regions of a biopsy image likely to contain cancer.

- Emphasis on **assist**. Not replace.

Biopsy image

**Ground truth
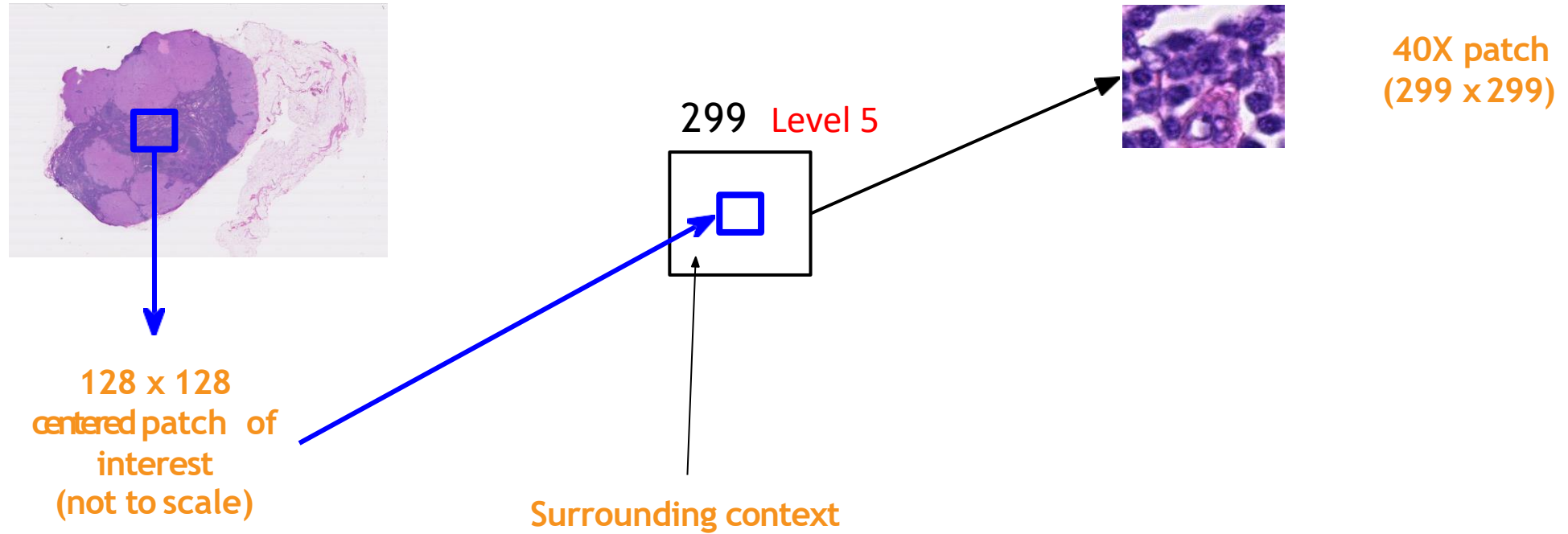(from pathologist)**

Model predictions

Approach

[Detecting Cancer Metastases on Gigapixel Pathology Images](#), 2017

Using Multi Scale Images to classify if a region contains cancer

-- Level 5, 6, 7

Single Scale Model

299 Level 5

40X patch
(299 x 299)

128 x 128
centered patch of
interest
(not to scale)

Surrounding context

Multi Scale Model

299   Level 6

40X patch
(299 x 299)

20X patch
(downsampled
to 299 x 299)

128 x 128 patch
of interest
(not to scale)

Surrounding context

Multi Scale Model

128 x 128 patch of interest (not to scale)

299   Level 7

Surrounding context

40X patch (299 x 299)

20X patch (downsampled to 299 x 299)

10X patch (downsampled to 299 x 299)

...

Input images
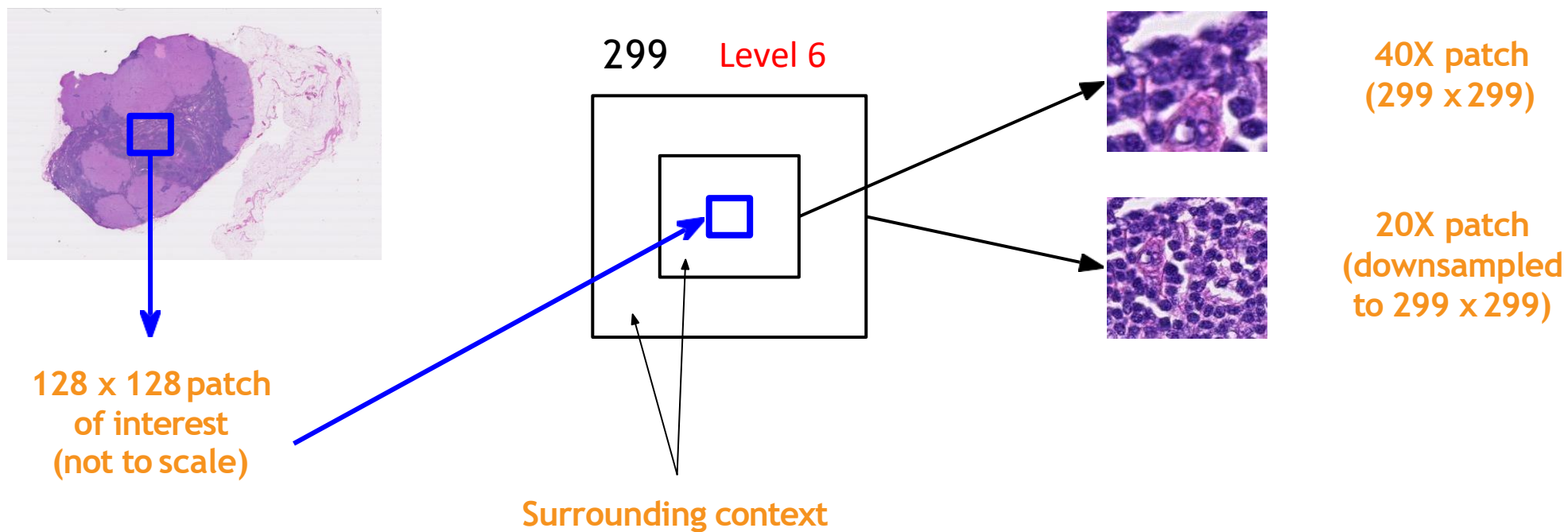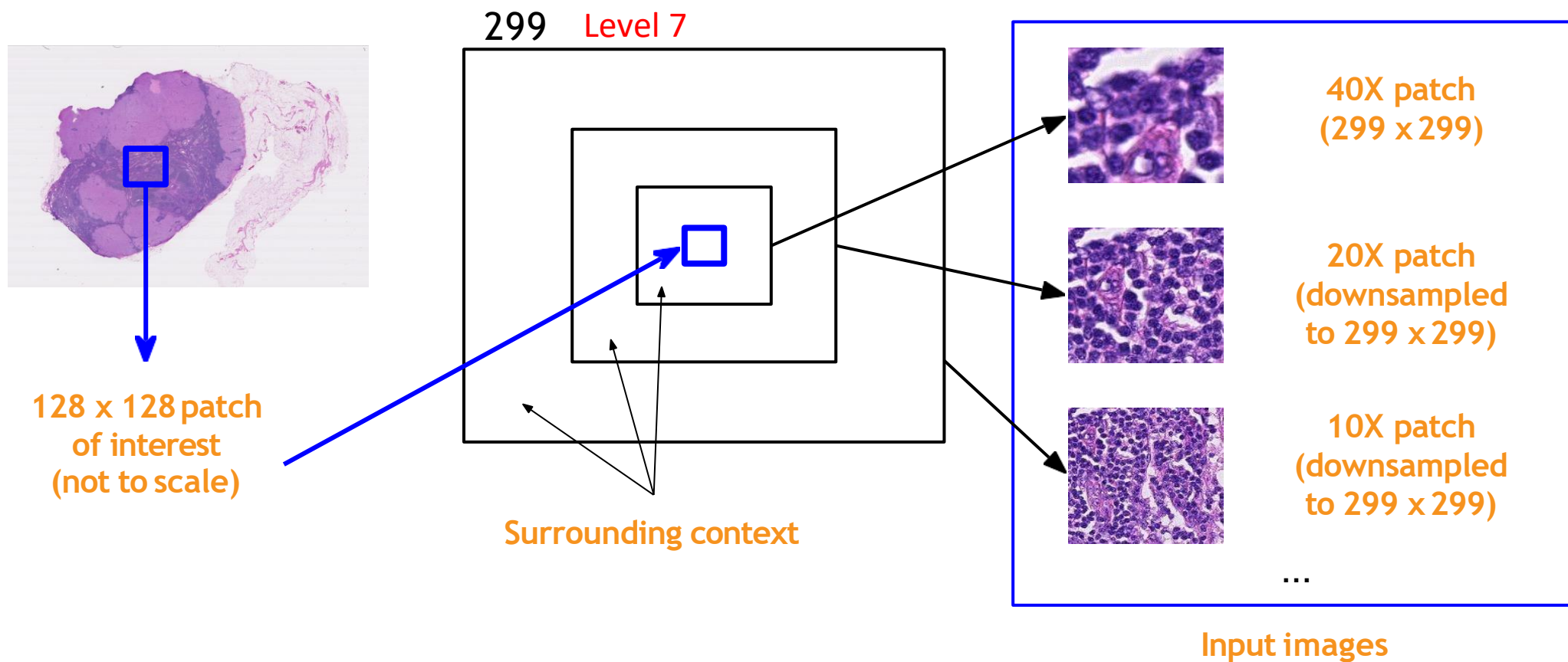
Key Experiments

Imbalanced Data

Data Augmentation

Imbalanced data

Goal: Increase proportion of positive samples, but also keep as much
   original distribution as possible

- Tissue > 40%
  - Positive Sample -- Keep
  - Negative Sample – Drop 85%

- Tissue < 40%
  - Positive Sample -- Drop (To avoid false positive)
  - Negative Sample – Drop 98% (can train model better)

Data augmentation

- Vertical and Horizontal Flip
- Rotate -- 90 degree
- Brightness -- 20% to 100%

Model Comparison -- Accuracy (Validation)

- Single Scale Model
  - Inception V3 -- 95%
  - Simple Model – 70%

- Multi Scale Model
  - Inception V3 -- 66%
  - Simple Model – 92%

Model Comparison -- Test on Slide 101 (Threshold: 0.5)

Model	Accuracy	Precision	Recall	
Single-scale Inception V3	0.22	0.22	1	Tend to classify all as positive
Multi-scale Simple model	0.64	0.41	0.85	Better

Multi-scale Simple Model Result -- Test Slide 101

Mask and Slide Prediction and Slide Prediction and Mask

Total patches used to test: 225
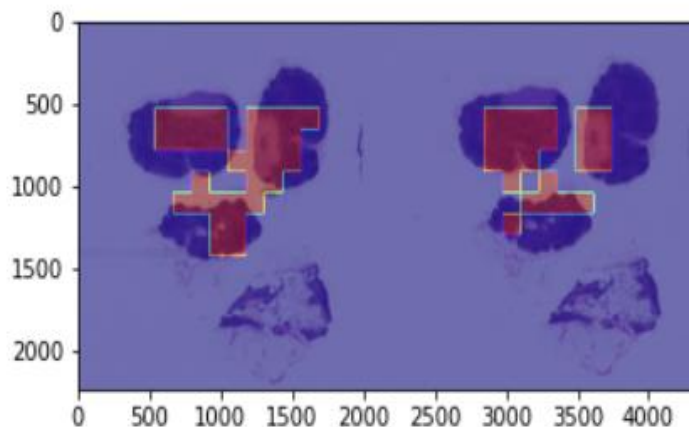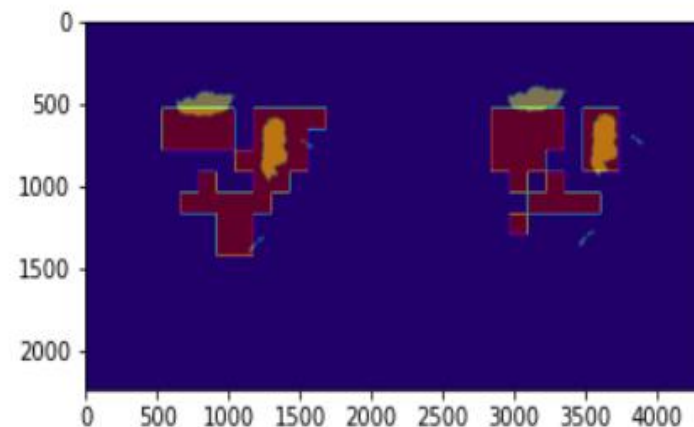
23 42 32 4
Accuracy: 0.6435643564356436
Precision: 0.41818181818181815
Recall: 0.8518518518518519

Further Work

1. Why can not multi-scale input improve inception V3 model?

Potential Solutions:
- If the prediction on the Level 5 image is positive, we include level 6 and 7 images as reference ; otherwise we classify it directly as negative to avoid false positive

- Try other kinds of image models, like VGG

Further Work

2. Why does the multi-scale simple model work not so good for slide with less cancer tissues?

Potential Solutions:
- Higher resolution images as input

- Fine-tune training data or model to detect minor difference of patches