Breast Cancer Detection on Pathology Images

Applied Deep Learning

Name: Smarth Gupta

UNI: sg3868

Motivation

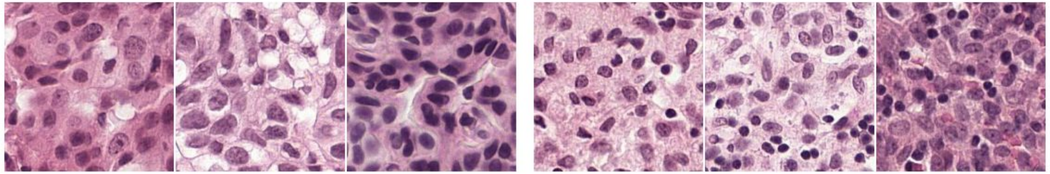


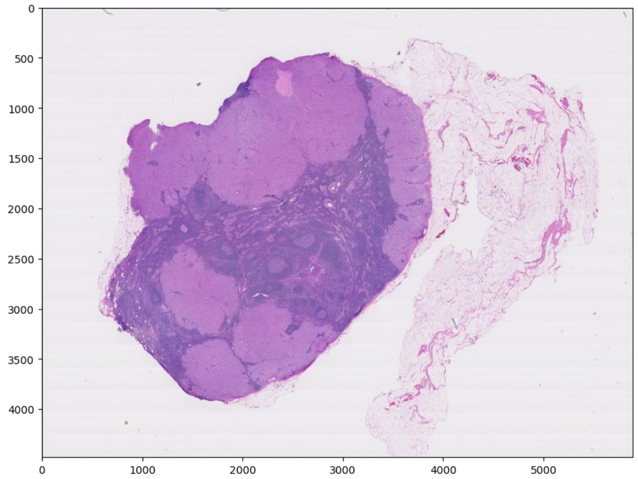
Fig. 1. Left: three tumor patches and **right**: three challenging normal patches.

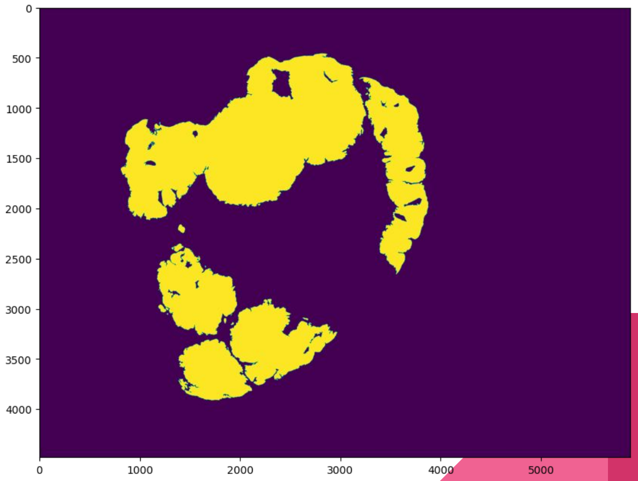
- Each year, more than 230 ,000 breast cancer patients in the U.S. hinge on whether the cancer has metastasized
- Metastasis detection is performed by pathologists reviewing large expanses of biological tissues.
- This process is labor intensive and error-prone.
- We can use various SOTA image segmentation/classification techniques to tackle this challenge.
- In this project we aim to implement a small scale implementation of the model proposed in the paper

"Detecting Cancer Metastases on Gigapixel Pathology Images"

Dataset

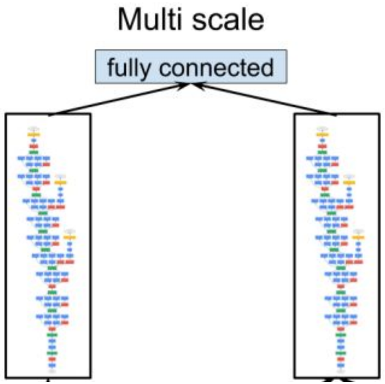
We use the CAMELYON 16 Multi Giga Pixel slides data. A set of 22 slides were provided to us. Each slide has a corresponding mask which annotates the tumor region. We can access around 9 magnification levels.

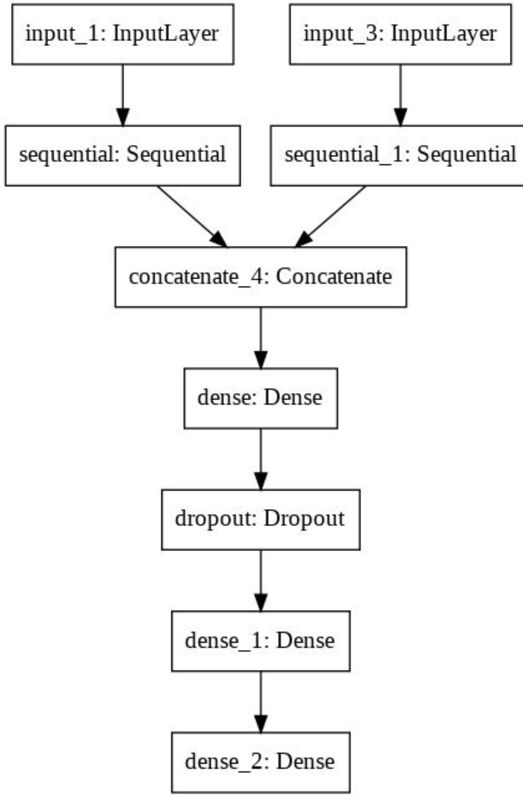




Methodology

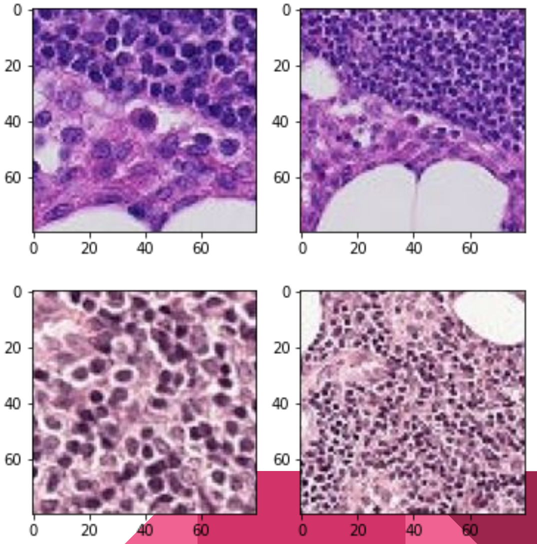
- **Architectures used:** Inceptionv3 (Multi-scale) *fine-tuned for layers > 150*
- **Initial Weights:** Image Net
- **Sliding Window:** 80x80
- **Center:** 50x50
- **Loss used:** Binary Cross entropy loss
- **Magnifications:** (2, 3) and (3, 4)





Data Augmentation

We have used various data augmentation techniques to complement our training data and make our models more robust.



- Random Orthogonal Rotations
- Random Horizontal and Vertical Flips
- Brightness with a maximum delta of 64/255
- Saturation with a maximum delta of 0.25
- Hue with a maximum delta of 0.04
- Contrast with a maximum delta of 0.75

Train Test Split

```
train_list = ['tumor_110', 'tumor_031', 'tumor_035', 'tumor_019', 'tumor_057', 'tumor_096',
'tumor_005', 'tumor_081', 'tumor_012', 'tumor_023', 'tumor_094', 'tumor_016', 'tumor_084',
'tumor_001', 'tumor_059', 'tumor_101', 'tumor_078']
val_list = ['tumor_075', 'tumor_064']
test_list = ['tumor_091', 'tumor_002']
```

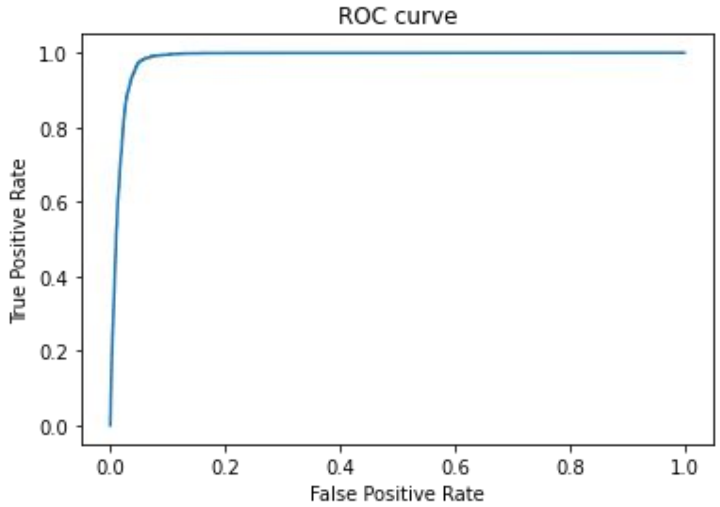
- The Split was chosen randomly at first but was later modified because the size of tumor was very skewed.
- Some slides had very large tumor patches while others had very less.
- We randomly balanced the slide distribution looking at the tumor size and tissues across the splits.

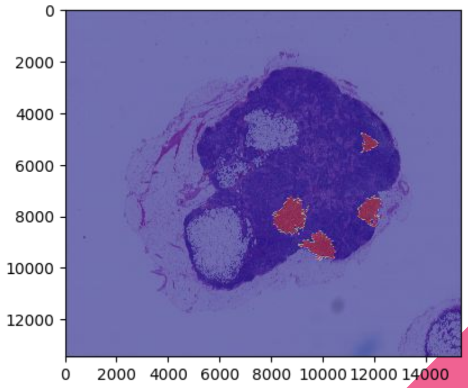
Results (Test Slide)

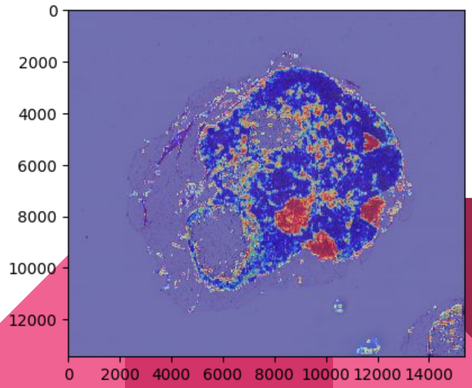
Using the magnification 2 and 3 slides.

- We see that all the tumor regions were identified correctly.
- We could generate a very high Recall (Which is important in the context of medical prognosis)

```
1) AUC:0.98528

2) Threshold:0.52767

3) Sensitivity:0.97877

4) Specificity:0.94747

5) Recall:0.97809

6) Precision:0.21504
```

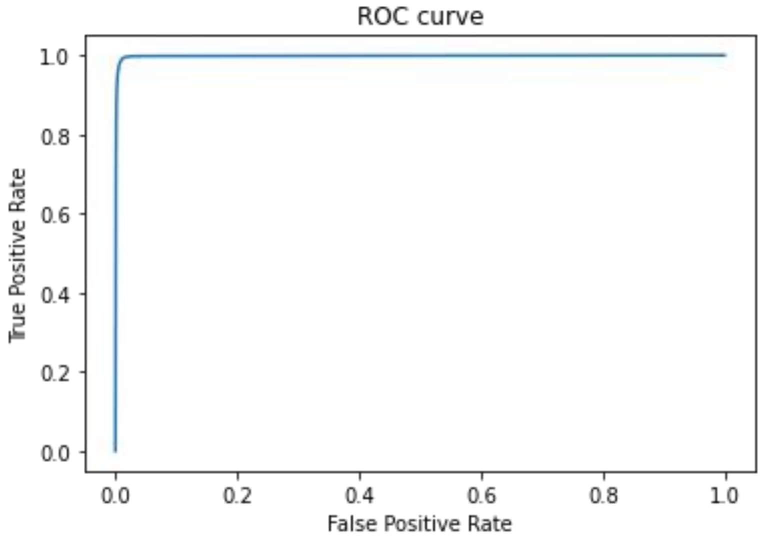




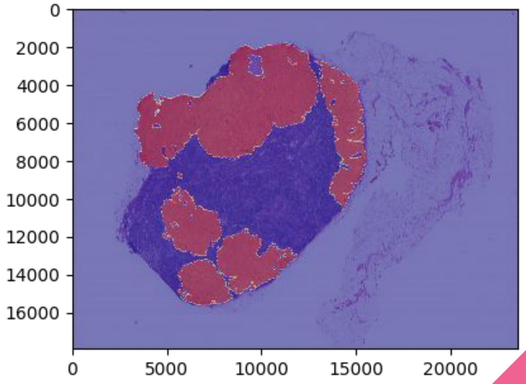


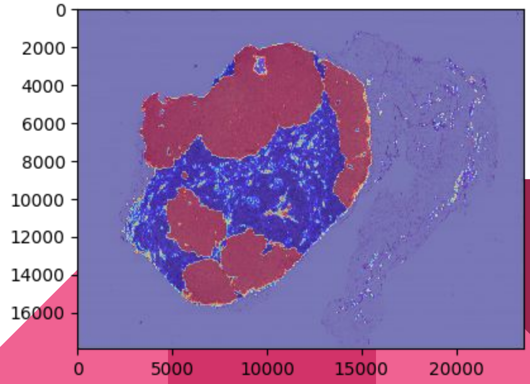
Results (Train Slide)

Using the magnification 2 and 3 slides.




```
1) AUC:0.99807

2) Threshold:0.79503

3) Sensitivity:0.99294

4) Specificity:0.98734

5) Recall:0.99775

6) Precision:0.88057
```



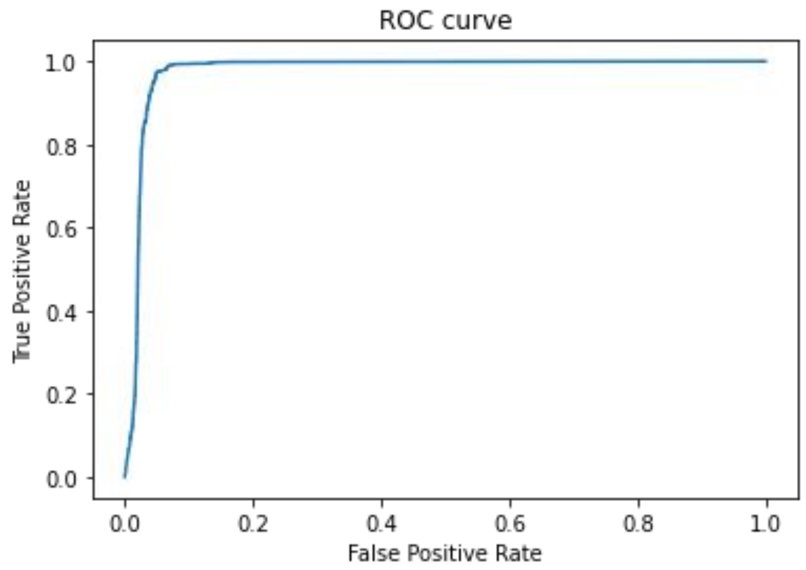


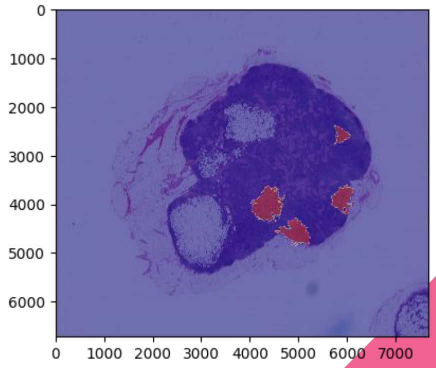
Results (Test Slide)

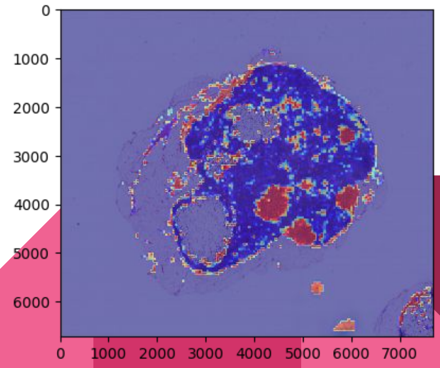
Using the magnification 3 and 4 slides.

- We see that all the tumor regions were identified correctly.
- We could generate a very high Recall (Which is important in the context of medical prognosis)

```
1) AUC:0.97659

2) Threshold:0.48429

3) Sensitivity:0.97478

4) Specificity:0.95004

8) Recall:0.97277

9) Precision:0.22275
```



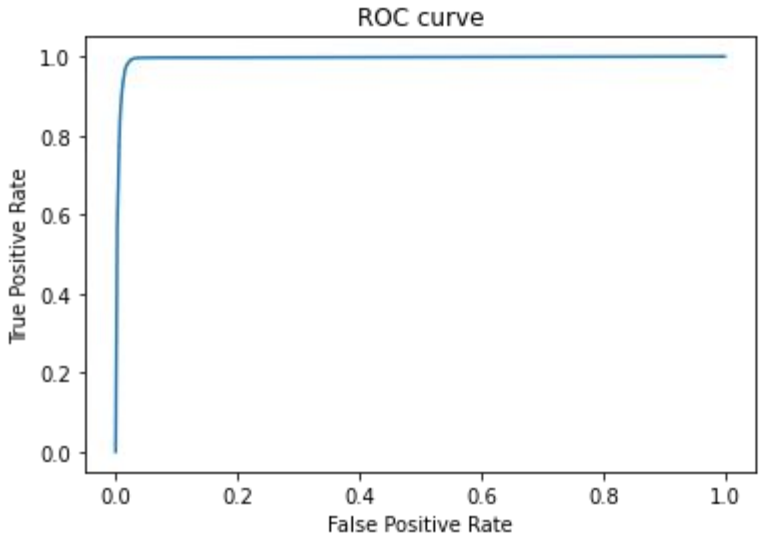


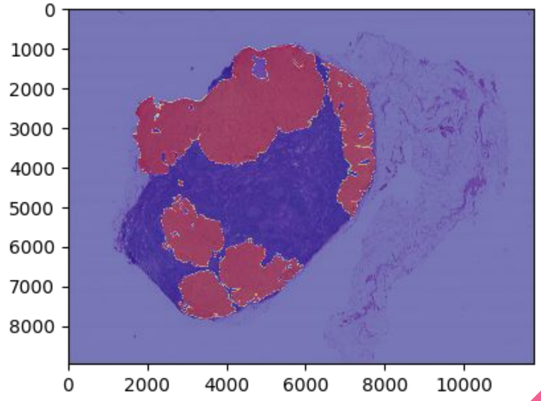


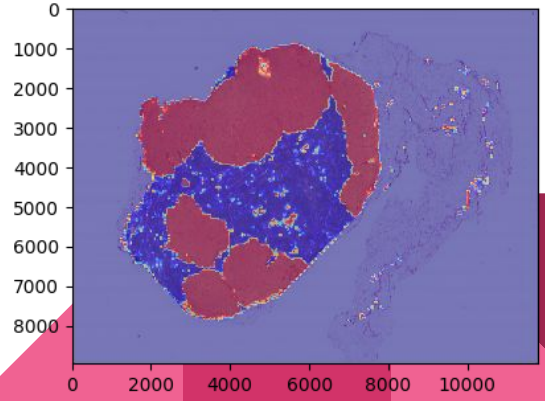
Results (Train Slide)

Using the magnification 3 and 4 slides.

1) AUC:0.99457
2) Threshold:0.72445
3) Sensitivity:0.99161
4) Specificity:0.97514
5) Recall:0.99155
6) Precision:0.88057







Conclusion

- We notice better results with higher magnification images (2, 3) as compared to lower magnification (3, 4).
- We were able to train a model with high Recall
- Transfer Learning with Fine-Tuning was effective in producing good results with less computationally intense training
- The model seems to make less accurate predictions on boundaries
- More data can be utilized to make predictions better.

Practical Consideration

- Didn't consider magnification level lower than 2 because of memory constraints in Google Colab and greater training times due to a large dataset

- Truncated the training and validation data to a smaller sample size, to allow handling memory constraints

- Fine-tuned only the top layers of inception model

Future Improvements

- Use lower magnification Images, by getting access to better GPU and high RAM machines

- Use prediction averaging by computing the prediction on each possible slide orientation to improve accuracy and introduce rotational invariance.

- Use better foreground and background separation techniques to improve performance on boundaries.

Thank You