Classification of breast cancer histology images

Presented by-
Aniket Raj (201452027)
Rajeev Kumar Singh (201452062)

Under supervision of
Dr. Ashish Phophalia

May 2, 2018

Outline

Problem Statement

Classification of H&E stained breast histology microscopy images in

- Four classes
- Three classes

The dataset contains a total of 400 microscopy images in RGB color model, distributed as follows[1]:

- Normal : 100
- Benign : 100
- In-situ carcinoma : 100
- Invasive carcinoma : 100

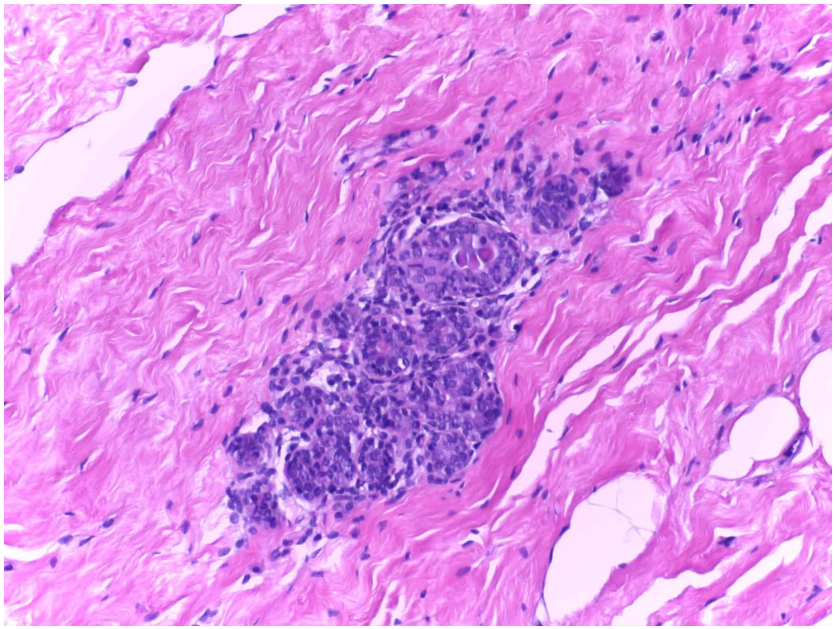


Figure: Normal

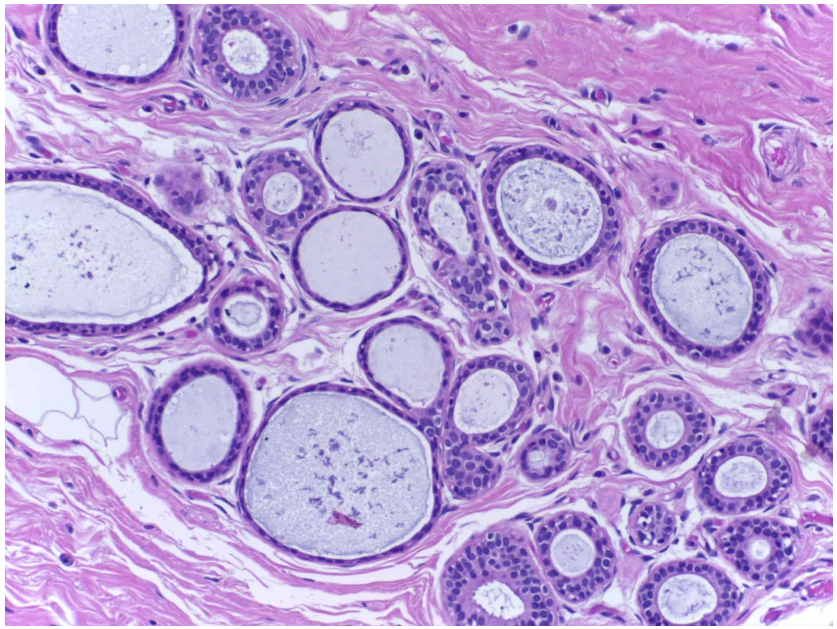


Figure: Benign

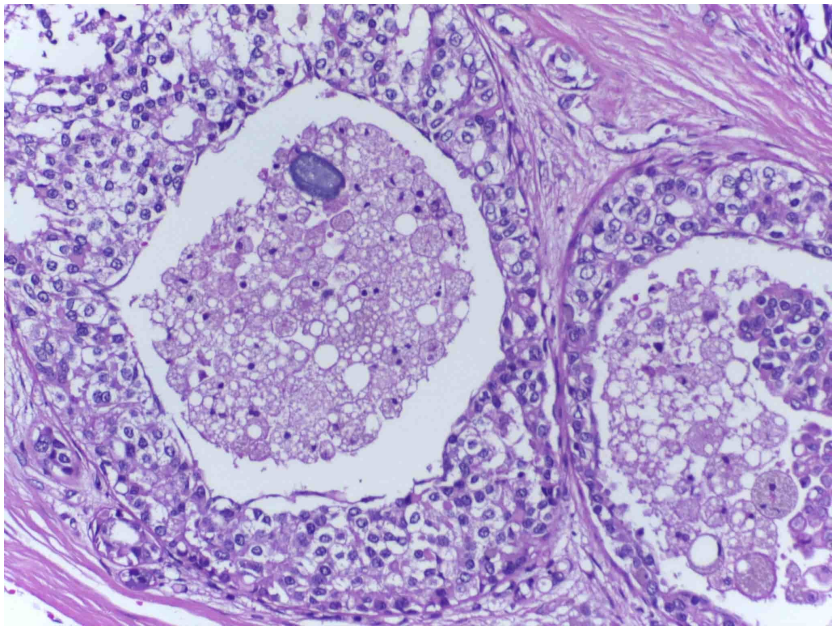


Figure: In situ

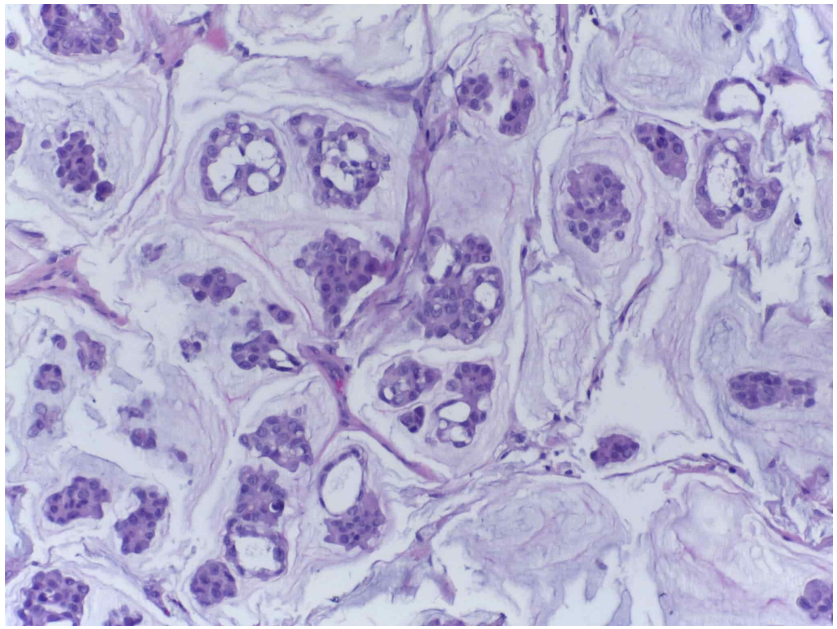


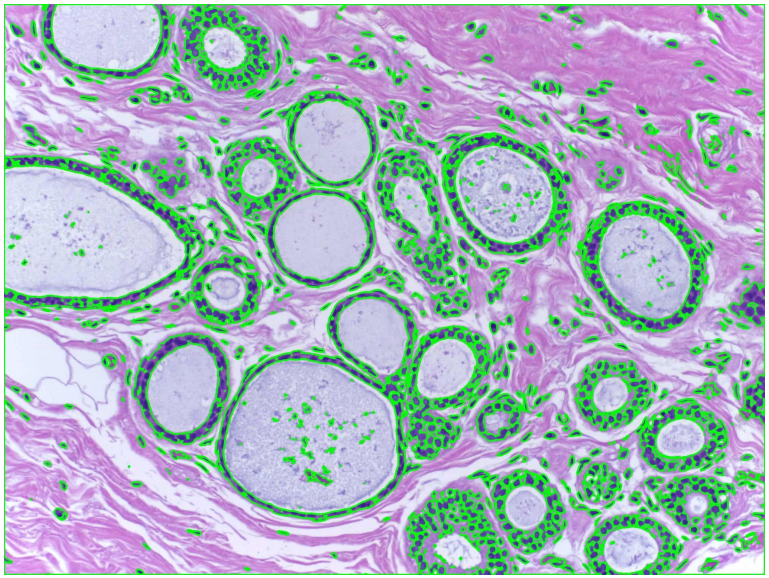
Figure: Invasive

Approach

- Conventional approach : Hand engineered features
- Modern : Feature extraction by CNN

Approach

- Active Contour Model : ACM describes the boundaries of shapes in an image. It utilizes knowledge, such as image intensity distribution information, boundary shape information, and texture information to obtain accurate results for object boundaries image analysis.



- Connected Component : Connected-component labeling is used in computer vision to detect connected regions in binary digital images. It scans an image and groups its pixels into components based on pixel connectivity.

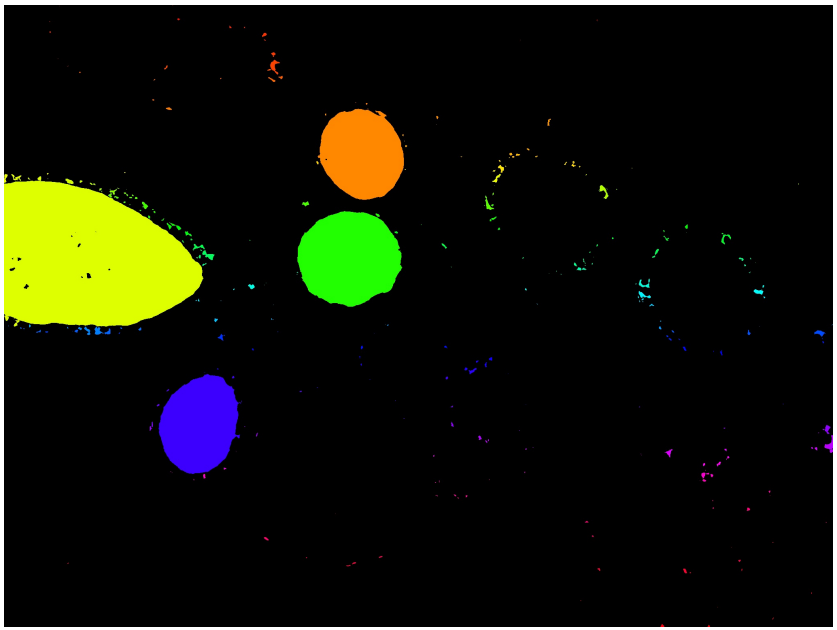


Figure: Connected Component

- Blob Detection
  - Laplacian of Gaussian
  - Difference of Gaussian

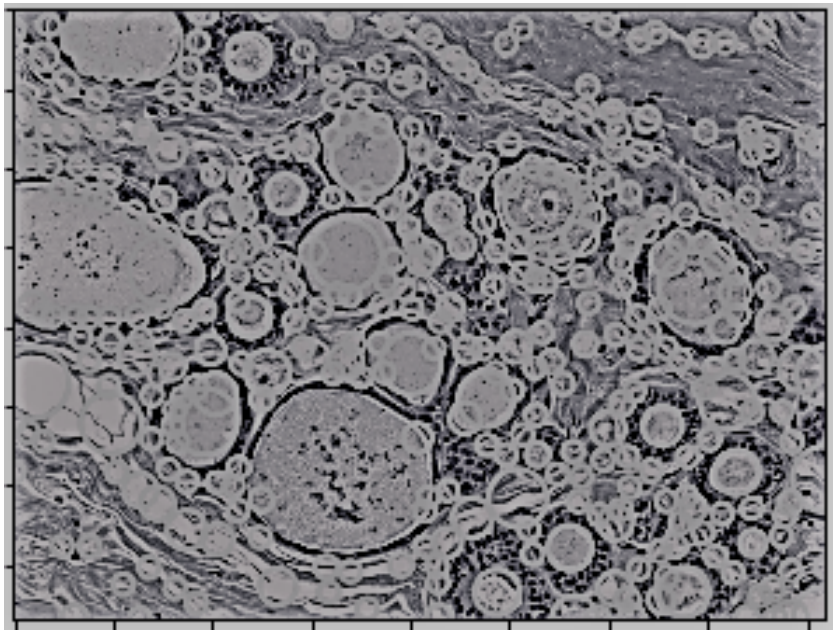


Figure: LOG

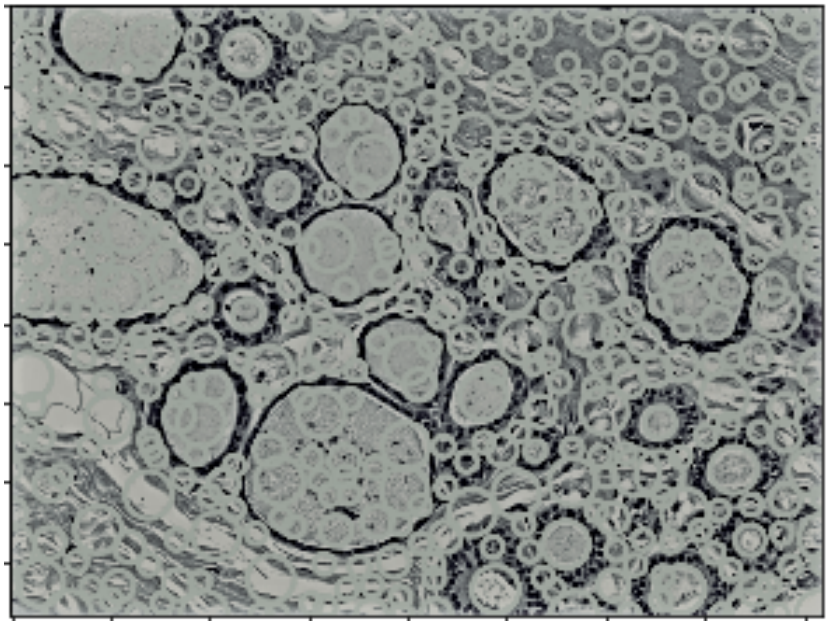


Figure: DOG

- Transfer Learning : It can be defined as learning and gaining knowledge from one problem and use the stored knowledge in some other problem.

- Densenet

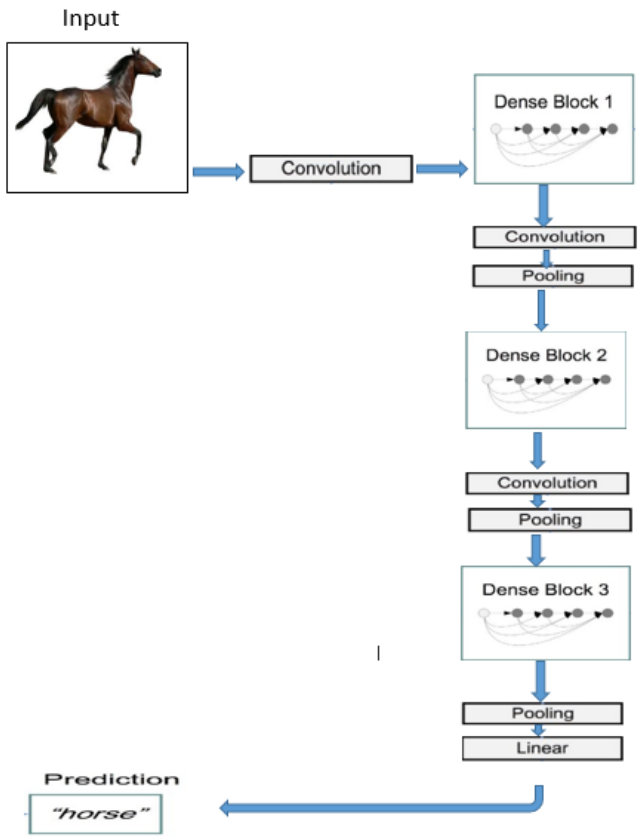


Figure: Full densenet example with 3 blocks from source paper

- VGG16

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 **LRN**	conv3-64 **conv3-64**	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 **conv3-128**	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 **conv1-256**	conv3-256 conv3-256 **conv3-256**	conv3-256 conv3-256 conv3-256 **conv3-256**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

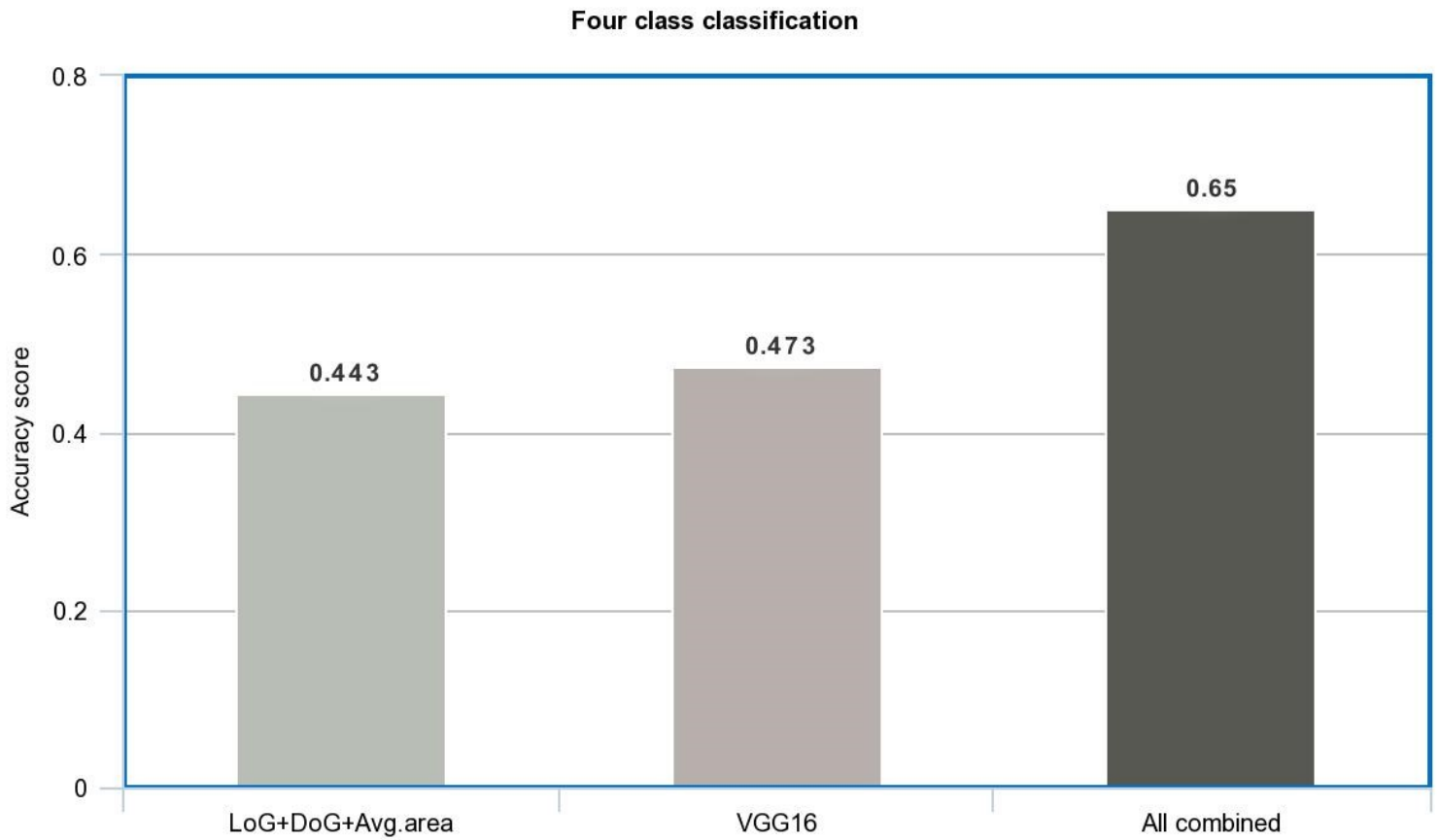
Figure: ConvNet configuration from source paper [3]

Experimental Setup

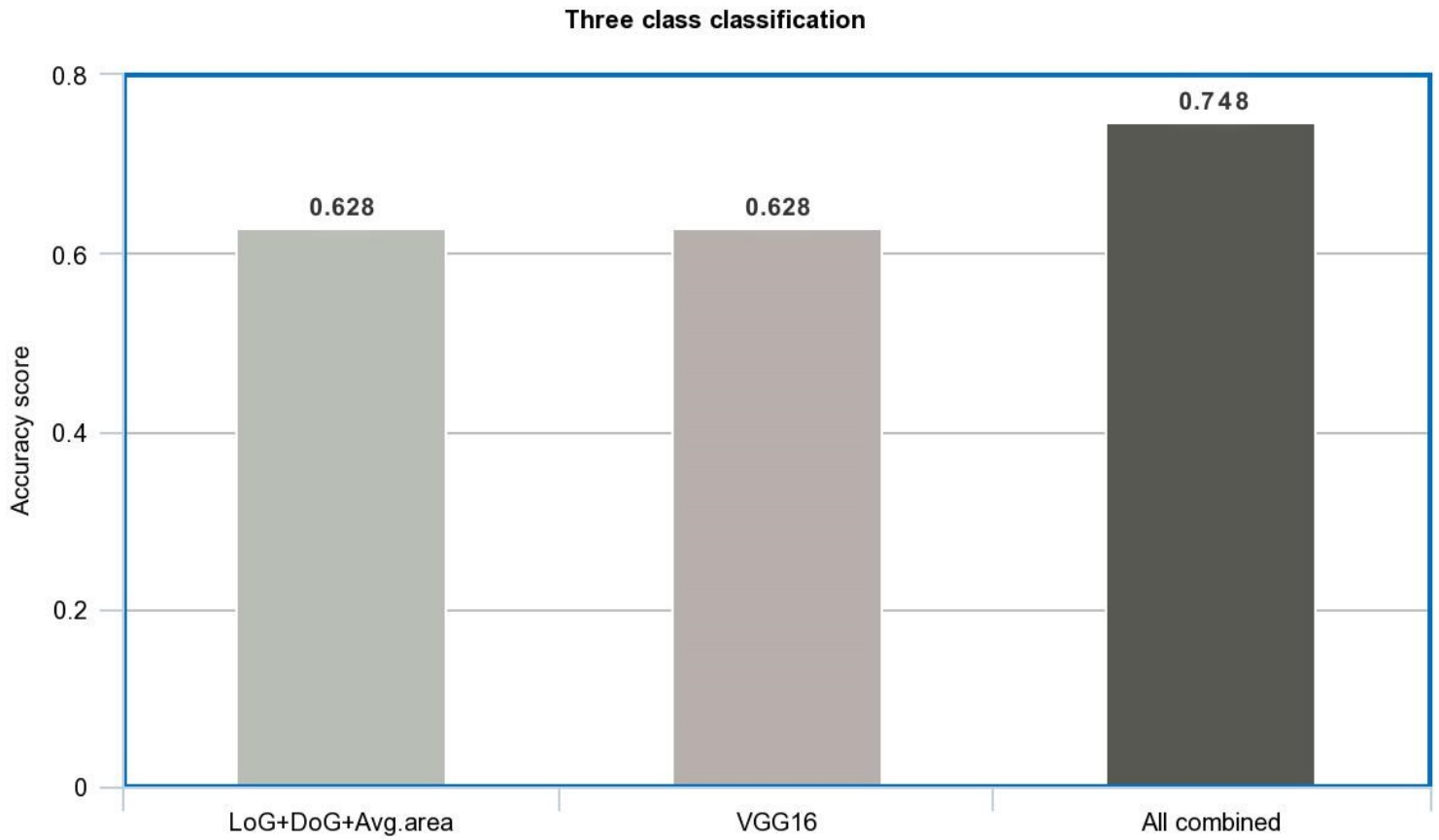
- Training data : 70% of the total data available is used as training data.
- Testing data : Remaining 30% of the data is used as test data.
- The experiment is repeated five times with different split of training and test data.
- A SVM classifier with linear kernal is trained for the classification problem.

Result

- Four Class Classification



- Three Class Classification



Conclusion

- The pretrained model is not fine-tuned due to the smaller size of available training data.
- We achieved the test accuracy of 65% for four class classification using a small dataset.
- The number of features used for training SVM classifier is small which drastically reduces the training time.

References

1. Dateset from ICIAR 2018 Grand challenge on Breast Cancer Histology images. https://iciar2018-challenge.grand-challenge.org/home/
2. Gao Huang,Zhuang Liu,Laurens van der Maaten. Densely Connected Convolutional Networks. CoRRabs/1608.06993 2016 http://arxiv.org/abs/1608.06993
3. Karen Simonyan Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. CoRRabs/1409.1556 2014 http://arxiv.org/abs/1409.1556