Breast Cancer Detection Through High Resolution Mammograms

Nhung Le, Nithish Addepalli, Ravi Choudhary

NYU

nhl256@nyu.edu, bva212@nyu.edu, rc3620@nyu.edu

May 9, 2019

Overview

Motivation

- Breast cancer is the second most common cancer in women worldwide. The five year survival rates for stage 0 or stage 1 breast cancers are close to 100%, but the rates go down dramatically for later stages: 93% for stage II, 72% for stage III and 22% for stage IV [Eric Antoine Scuccimarra, 2018].
- Mammogram screening is one of the most common diagnosis method for breast cancer.
- the prediction algorithms provided radiologists with useful information for breast cancer diagnosis.

Related Work

1. Use of CNN based deep learning methods to classify mammographic breast density on 22,000 mammogram images to observe an AUC of 0.9421.[Aly A. Mohamed (2018)]
2. Use of high-resolution Multi-view Deep convolutional Neural networks by Professor Krzysztof J. Geras on 886,000 mammogram scans achieving performance comparable to a committee of radiologists when presented with the same data.[Krzysztof J. Geras et. al. (2017)]
3. Various methods and versions of deep convolutional neural networks implemented on the ROI extracted from CBIS-DDSM dataset achieving accuracy of 95% and recall of 90% on test dataset. [Eric Antoine Scuccimarra, 2018]

Data Source and Description

The data set CBIS-DDSM [1](Curated Breast Imaging Subset of Digital Database for Screening Mammography) is an updated and standardized version of the DDSM, which has 2,620 scanned film mammography studies. The images went through a thorough preparation process including removal of questionable cases, image decompression, image processing, image cropping, and mass segmentation.

Metric	Value	Train	Val	Test
# Patients	1,607	1,097	349	161
# Scans	3,103	2,158	645	300

Data Prepossessing

- Sampling done based on patient id to avoid any overlap that would result in leakage.
- Train / Validation / Test: ratio 70:20:10
- Different techniques were used for batching the images of different sizes (e.g., training data with batch size $= 1$, we used the images without any resizing. For training with batch size > 1, images were either padded to the same size $7500 * 5500$ i.e., maximum length and width of images or they were resized using transforms
- Image augmentation by flipping the scans about the horizontal/vertical axis randomly
- Image normalization

	Train	Val	Test
# Benign	1,172	385	182
# Malignant	986	260	118

Loss Functions

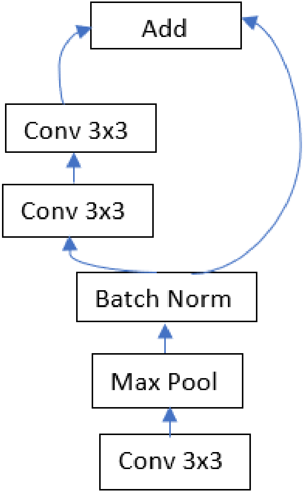
Cross-Entropy Loss:

$$L_i = -f_{y_i} + log \sum_{j=1}^{n} e^{f_j}$$

where f_j is the j-th element of the vector of class score f, and i is i-th instance.

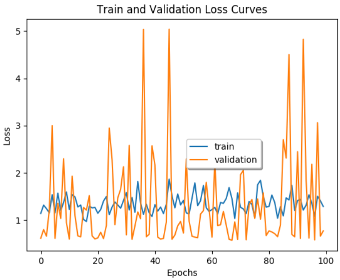
Models

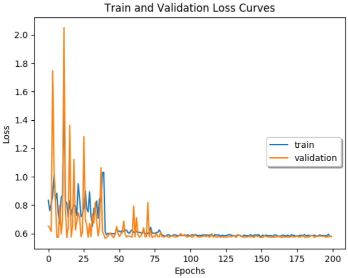
- Inception ResNet v2
- ResNet 18
- ResNet 34
- ResNet 50
- Customized Convolutional Neural Network



Hyper-parameters Selection

- Learning Rate
- Mini-batch Size
- Optimization: Adam, SGD, Root Mean Square Propagion
- Activation: ReLU, Leaky ReLU, ELU





Evaluation Metrics

Accuracy

The prediction is accuracy if the true specialty category is in the top 5 specialty categories highest predicted probability.

ROC - Receiver Operating Characteristic

ROC is a probability curve of True Positive Rate against False Positive Rate for different cut-off point.

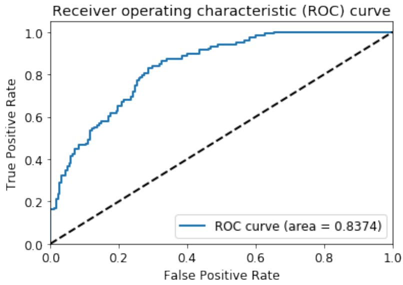
AUC - Area Under Curve

AUC represents degree or measure of separability. It tells how much model is capable of distinguishing between classes. The higher the AUC, the better the model is at predicting 0s as 0s and 1s as 1s.

Performance Evaluation

Architecture	Transfer Learning	2 Classes		4 Classes	
		AUC	Accuracy	*AUC	Accuracy
ResNet18 (full image)	-	-	-	0.63	0.36
Custom CNN with PReLU	-	-	-	0.64	0.35
ResNet18	Imagenet	0.71	0.65	0.53	0.32
ResNet34	Imagenet	0.67	0.65	-	-
Custom CNN	Chest X-Ray	0.64	0.56	0.73	0.43
ResNet50	Imagenet	0.82	0.72	-	-
Inception Resnet V2	Imagenet	0.84	0.74	0.85	0.60

Performance on 2 classes



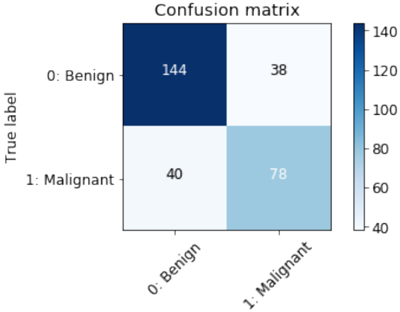
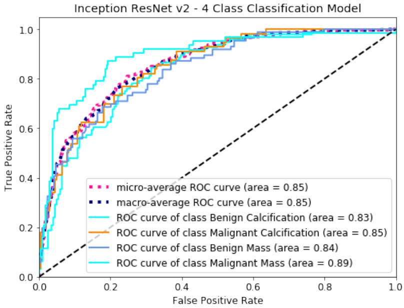


Figure: Inception ResNet v2 2 class classification model AUC plot and Confusion Matrix

Performance on 4 classes



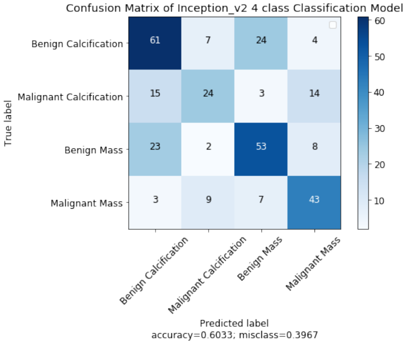


Figure: Inception ResNet v2 4 class classification model AUC plot and Confusion Matrix

Heatmap

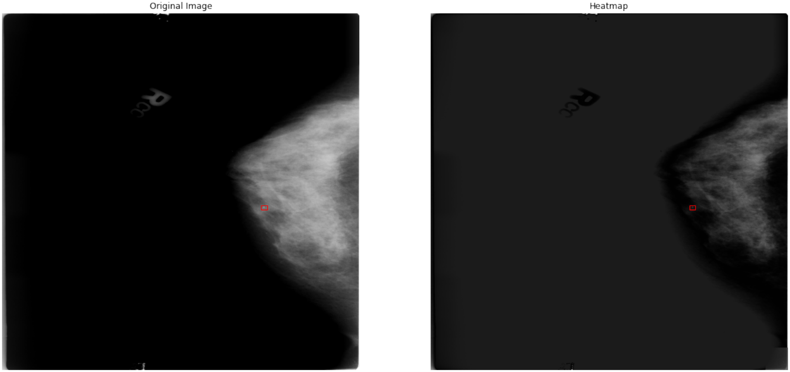


Figure: Heatmap for a lesion

References

Eric Antoine Scuccimarra(2018)
ConvNets for Detecting Abnormalities in DDSM Mammograms
https://medium.com/@ericscuccimarra/convnets-for-classifying-ddsm-mammograms-1739e0fe8028.

Krzysztof J. Geras et. al. (2017)
High-Resolution Breast Cancer Screening with Multi-View Deep Convolutional Neural Networks
https://arxiv.org/pdf/1703.07047.pdf.

Aly A. Mohamed et. al. (2018)
Aly A. Mohamed, Wendie A. Berg, Hong Peng, Yahong Luo, Rachel C. Jankowitz, Shandong Wu
A deep learning method for classifying mammographic breast density categories
https://www.ncbi.nlm.nih.gov/pubmed/29159811.

Rebecca Sawyer Lee et. al. (2017)
A curated mammography data set for use in computer-aided detection and diagnosis research
https://www.nature.com/articles/sdata2017177

Thank You