

Breast Cancer Imaging

An Analysis of Benign and Malignant Cancers

presented by
Bracken Sant

I. Introduction

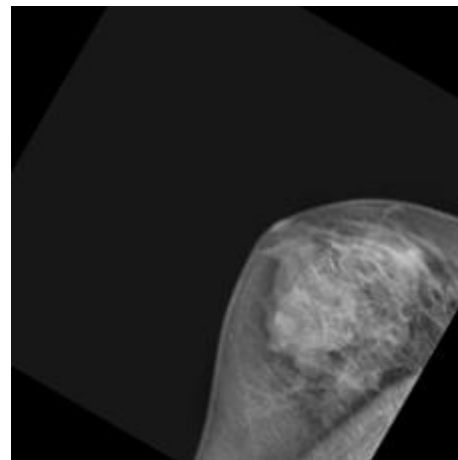
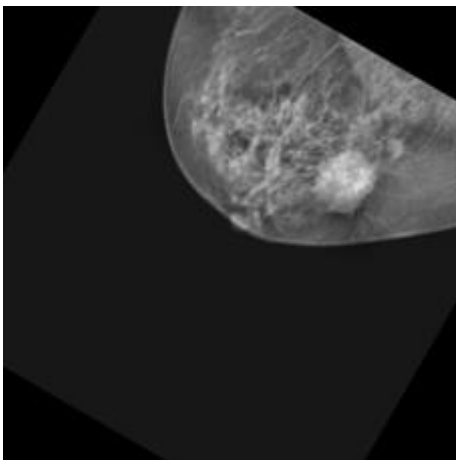
According to the National Breast Cancer Foundation's article "Breast Cancer Facts & Stats", "1 in 8 women in the United States will be diagnosed with breast cancer in her lifetime..." and "approximately 30% of all new female cancer diagnoses will be breast cancer." This is a very high percentage when compared to most other cancers for women, the exemption being skin cancer which is one of the most likely cancers to get for both men and women. However, According to Stony Brook Cancer Center's article "Different Kinds of Breast Lumps" out of those diagnosed with breast cancer, "Most breast lumps – 80% of those biopsied – are benign (non-cancerous)."

Mammography imaging is nearly half the price than both of the other options for breast cancer imaging (MRI and CT scans). With the out of pocket price being between \$200-\$300 for mammograms according to GoodRx, while the cost for MRI for breast cancer imaging being between \$400-\$600 and for CT breast cancer imaging being between \$350-\$500 according to The Women's Imaging Center's article "What is the Difference Between MRI and CT Scans?". There are a large percentage of tumors or lumps that aren't cancerous and don't need to be worried about or further inspected, so in an attempt to make diagnosing more affordable and to be more available I have made a convoluted neural network that predicts from mammography imaging whether the breast cancer is benign or malignant.

II. Data Processing

The only data processing that occurred was separating the data into Training, Validation, and Test data sets. I attempted to use augmentation to increase the overall accuracy along with the recall for the Malignant data set; however, adjusting the brightness, rotating the image, using sheer, zoom, flipping the image horizontally and vertically did not change the accuracy in a postive way. I believe that some of the reasons for this to be because the image is in black and white, the usable parts of the image are fairly small in the image, and there are multiple shots of each image at different angles because the images are from four-view mammography imaging.

The first image is an example of malignant and the second image is an example of benign.

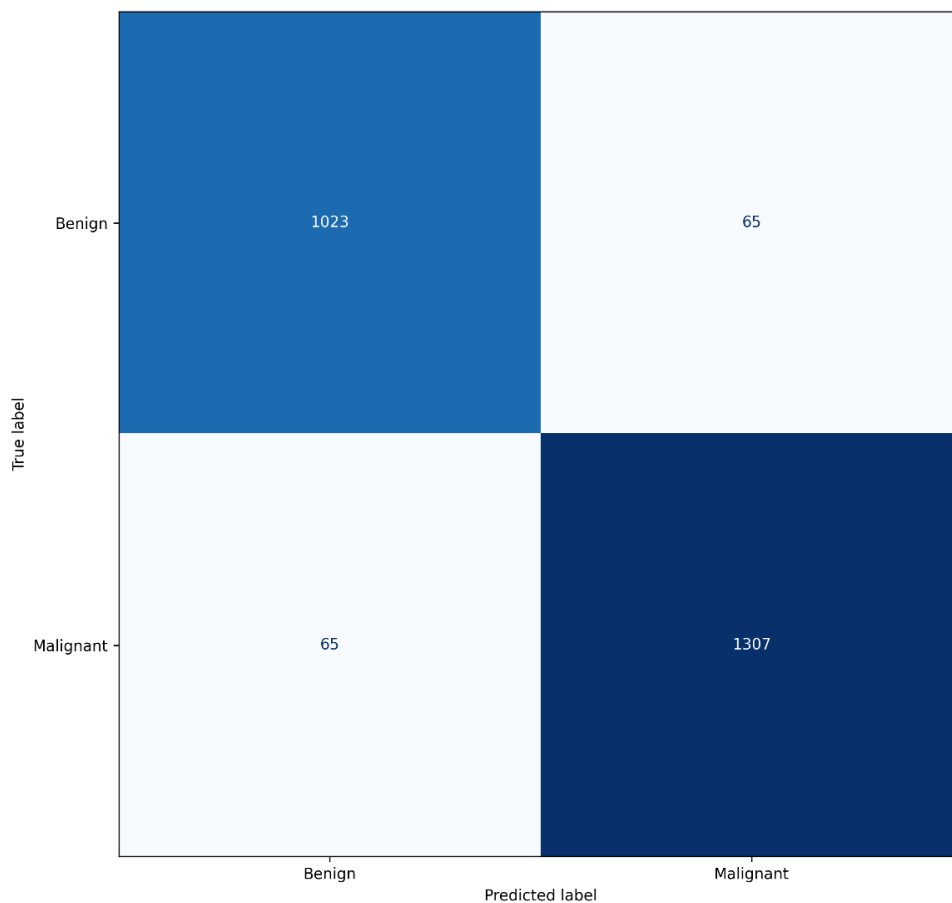


III. Machine Learning Model

The machine learning model that I picked was a convolutional neural network. I decided to make my own model from the keras library. I used 4 different layers the first 3 being double convolutional with a max pool and dropout being between the layers and the last with a flatten, dense, dropout, and then final dense to 2 outputs. I used differing kernal sizes to set different window sizes, and used kernal constraint and bias constraint to help skew the results to predict Malignant; however, it only ended up evening out the results of misassigned Benign and Malignant. I used alternating swish and leaky_relu activations for the layers.

IV. Interpretation of Results

The metrics that I chose to look at for the model were accuracy and recall or sensitivity for the Malignant classification. The accuracy is good for an overall look at how well the model is doing, the recall for the Malignant classification is to check type 2 errors or if we predict that the cancer is Benign when it is Malignant. The model had an accuracy of 95% and a recall of 95%.



V. Comparison of Results

A study with its percentages was shown on the National Library of Medicine and these are the given results that they found for CT and Mammogram results: “CT sensitivity, specificity, and accuracy of CT for breast cancer detection was 84.21%, 99.3%, and 98.68% compared to 78.95%, 93.78%, and 93.16% for four-view mammography.” Compared with the sensitivity of both CT and mammography is lower than the sensitivity / recall of the model and the accuracy is higher than the mammography but lower than the accuracy of the CT scan.

The accuracy if it only predicted Malignant would be 55.8%.

VI. Ethical Implications

The only ethical implications would be if it is ethical to use breast cancer imaging. However, all of the images used are from public databases and in order for them to be on them they would have had to consent for their images to be used for research.

VII. Python Notebooks & Work Cited

Below is the Github Gist link to the notebook used during this case study:

<https://gist.github.com/bracken576/f9e49ab675a96d597efc29a5480abde5#file-breast-cancer-imaging-cnn-model-ipynb>

Below are the links for websites used to gather information:

<https://data.mendeley.com/datasets/ywsbh3ndr8/5>

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9695285/>

<https://cancer.stonybrookmedicine.edu/breast-cancer-team/patients/bse/breastlumps>

<https://www.nationalbreastcancer.org/breast-cancer-facts/>

<https://www.goodrx.com/conditions/breast-cancer/how-much-mammogram-cost>

<https://thewomensimagingcenter.com/news/mri-mri-breast-mri-shoulder-mri-knee-mri-spine-mri-cost-mri-low-cost-denver-imaging-denver-womens-imaging-center>

<https://pubs.rsna.org/doi/10.1148/rg.2021200145>