Breast Cancer Detection using Deep Learning

CIND860: Capstone Project

Project by: Ann Sam
ann.sam@torontomu.ca
Student #501160843

Supervisor: Dr. Ceni Baboglu

Date of submission: November 30th, 2023

Contents

1. **Abstract** **3**
2. **Introduction** **3**
3. **Literature Review** **4**
4. **Data Description** **8**
5. **Approach** **11**
6. **Data Preprocessing** **12**
7. **Image Enhancement** **14**
8. **Simple CNN Model** **14**
9. **ResNet50 Model** **16**
10. **VGG16 Model** **19**
11. **Results** **23**
12. **Conclusions and Recommendations** **23**
13. **References** **24**
14. **Terminology** **25**

Abstract

Breast cancer remains a significant challenge in healthcare, where it is reported that one in eight women will be diagnosed with it according to the Canadian Cancer Society. The greatest chance of survival demands innovative approaches for early detection and prognosis and in Canada, provincial and territorial screening programs use mammography to detect abnormalities. Mammography is an x-ray of the breast that uses low dose radiation to capture an image called a mammogram which are then allows radiologists to carefully examine the image for high density regions.

This deep learning capstone project used the breast cancer imagery from the CBIS-DDSM (Curated Breast Imaging Subset of Digital Database for Screening Mammography) dataset to predict breast cancer diagnoses using three popular convolutional neural network (CNN) architectures: simple CNN, ResNet50, and VGG16. The highest accuracy was with the second iteration of VGG16 with an accuracy of 0.666.

Introduction

Breast cancer is the most common cancer among Canadian women and is the second leading cause of death from cancer. The cancer starts in breast cells that can grow and metastasize, spreading to other parts of the body. In 2022, an estimated 28,600 Canadian women would be diagnosed with breast cancer and of those women, 5,500 will die from breast cancer, a mortality rate of 20% (Canadian Cancer Society, 2022). These statistics emphasize the importance of early detection by leveraging artificial intelligence. As mammography is the most

efficient diagnostic method for early detection, utilizing the CBIS-DDSM dataset from screening programs and patient records, this project seeks to develop accurate and reliable predictive models for early detection of breast cancer.

The project is divided into four main components where we will explore the potential of the three deep learning architectures, build the deep learning networks and compile the models, plot the accuracies, and loses of the respective models, and present the final predictions. This project will use three different convolutional neural network (CNN) architectures: a simple CNN architecture created from scratch. ResNet50 model which is a deep residual neural network architecture with 50 layers and does well dealing with vanishing gradient problems. And finally, VGG16 which consists of 16 layers including 13 convolutional layers and 3 fully connected layers.

This capstone project hopes to determine if utilizing the application of deep learning techniques can significantly advance breast cancer prognosis during the screening process with the highest accuracy. Mammography is an important tool, the additional help from AI to aide radiologists in screening expedites prognosis results to patients.

The data set was retrieved from:

https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=22516629

The raw data and processes for this study can be accessed from:

https://github.com/annsam0115/CIND860_CapstoneProject

Literature Review

Advancements in machine learning has transformed the medical industry in aiding in interpreting digital medical data. Machine learning and its subset, deep learning, has been

particularly key in classification and image processing related to breast cancer and other forms of cancer diagnosis. Advanced medical imaging systems such as mammography can provide life saving early detection for cancer patients. The downside is that it can also be an enormous strain on the healthcare system with regards to workload for the respective radiology department to individually analyze and report the results in a timely manner. The use of deep learning techniques aims to assist in these tasks and improve the accuracy of breast cancer detection.

Artificial intelligence as been used in aiding in diagnostic decision making as far back as the mid-90s. Kahn Jr. et al. (1997) were constructing Bayesian networks for breast cancer detection to assist in mammographic diagnosis. They created a Bayesian network model called MammoNet that incorporated five patient-history features, two physical finds, and 15 mammographic findings. The model looked at the presence of calcification or masses within the mammograms and manifestations of malignancy or not. They were able to return results with an accuracy of 85% and where able to train an artificial neural network (ANN) to return results with a receiver operating characteristic (ROC) curve of 0.89. They note that ANNs have advantages as they have the ability to learn patterns directly from observations but have several disadvantages during the training process where adjusting numerical weights need to be done incrementally like a “hill-climbing process”. Its also important to note the dataset used consisted of only 77 cases.

Ten years later, Cruz-Ramierz et al. (2007) would approach computer aided diagnosis using the Bayesian network framework and return 93.04% accuracy on breast cancer lesions by a single observer and 83.31% accuracy on multiple observers. They found that their accuracy decreased with multiple observers due differing observations while looking at samples in the microscope introducing more variance and thus diminishes the performance of classification. They approach their modeling by running seven different Bayesian network classifiers: Naïve Bayes, Bayes-N, MP-Bayes, Greedy, MP-Bayes + Greedy, CBL2, and PC. Its important to note that no neural network techniques were used in this study and so no image classification was

performed. They explicitly note that all the features coded in the dataset were expertly curated by a cytopathologist and no preprocessing was required to run the models. They note that image analysis techniques should be conducted so that a more objective approach could be performed and compare the BN classifiers that way.

In Alanzi et al. (2021) paper on boosting breast cancer detection using convolutional neural network (CNN), returned accuracy results of 87% with the use of a 5-layer CNN and regularization techniques such as dropouts. Using a large dataset of 275,000 image patches, they were able to increase their detection accuracy by 9% by using CNN over machine learning classification algorithms for predicting invasive cancer. Their results were not as accurate as they had hoped and noted limitations within the dataset itself and felt that further data cleaning, trimming, and preparation would have benefited the training and thus yielded a higher accuracy percent.

A specific Bayesian convolutional neural network was proposed as an optimized automated detection technique by Ezzat & Hassanien (2023) employing an approximate Bayesian version of a pre-trained CNN model called ResNet101V2. The ResNet101V2 architecture adds the Monte Carlo drop-out to the predictive distribution and also utilized optimization algorithms to be able to return accuracy results of 93.83%. This study puts weight on the input and parameter noise and handles those uncertainties with the Monte Carlo dropouts. They where able to compare their findings with other Bayesian and non-Bayesian models in recent related works and their results demonstrated a higher efficiency the models across accuracy, balanced-accuracy, F1-score, precision, and sensitivity values.

Looking at a study on using ANN as an automatic classifier, Kaymak et al. (2017) proposed a methodology of using Back propagation Neural Network (BPNN) and radial basis neural networks (RBNN) on 115 images. Their study yielded 59.0% and 70.4% accuracy respectively on the models. In their study they were classifying histopathology images which are biopsy images of cancer tissue under a microscope. Their images appeared to have been preprocessed and filters applied to attempt to mitigate the effects of noise to identify specific wavelets called Haar discrete wavelets which may be more

complex to accurately detect and given the small training set, could be the reason for the lower accuracy results.

Finally, in the last paper studied called Deep Learning to Improve Breast Cancer Detection on Screening Mammography (Shen et al., 2019) using the same dataset in this capstone project, was able to achieve AUC of 0.88 on a single model and AUC of 0.91 on a four-model network using two popular CNN structures: VGG network and ResNet (residual network). To be able to achieve they also compared their trained classifier using ImageNet database to ensure the initialize weights performed the same. By doing so, they were able to adjust the learning rate dependent on the upper and lower layers in a 3-stage training strategy. They also implemented a patch-based classifier classifying the images into 5 classes (normal, benign-mass, benign-calcification, malignant-mass, malignant-calcification) using Resnet50 and VGG16 on the test set. They found that patch classification was critical to the accuracy of the whole image classification technique and believed that the accuracy of while image classification was improved with more or larger patches. An important note that running these networks required a lot of computation power that was not available and required high GPU memory for processing and thus limited their study as a whole.

Clearly, it has been demonstrated, even as far back as the mid-90s that computer aided detection with the use of deep learning can be highly accurate in diagnosing breast cancer and can be highly successful when applied in medical imaging problems and prognoses.

Data Description

The Curated Breast Imaging Subset of Digital Database for Screening Mammography (CBIS-DDSM) dataset is an updated and standardized version of the Digital Database for Screening Mammography (DDSM). The original source files were obtained from the Massachusetts General Hospital, Wake Forest University School of Medicine, Sacred heard Hospital, and Washington University of St. Louis School of Medicine following a grant from the United States Department of Defence Breast Cancer Research program, and the United States Army Research and Material Command.

The CBIS-DDSM collection is a carefully selected subset of the DDSM data where images have been decompressed and converted to DICOM format. The dataset includes updated ROI (Region of Interest) segmentation and bounding boxes, as well as pathologic diagnosis information for training data. Specifically, the CBIS-DDSM dataset addresses the previous challenges of other open-source mammography datasets by providing a well-curated and standardized version of the DDSM for future computer-aided detection research in mammography. 6,775 studies make up the dataset profiling 1,566 patients with a total of 10,239 mammography images. These images are in a Digital Imaging and Communications in Medicine (DICOM) format which is standard for medical images and preserves all the metadata from the original dataset files after image decompression (Sawyer-Lee et al., 2017).

There are three image types within the dataset: full mammograms, cropped images of the full mammograms, and the masked imagery of the region of interest (ROI). The images have been pre-processed and resized to 299x299 images with random flips and rotations. The images are labeled with two labels:

1. Pathology: 0 for negative (Benign) and 1 for positive (Malignant)

2. Multi-class Category: 0 is undetermined, 1 well differentiated, 2 is moderately differentiated, 3 is poorly differentiated, and 4 is undifferentiated

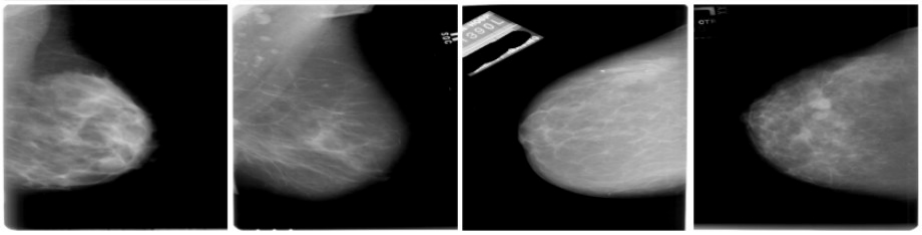


Figure 1: Full Mammograms

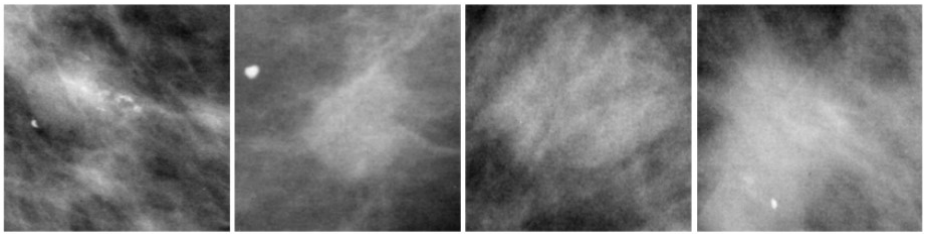


Figure 2: Cropped Images

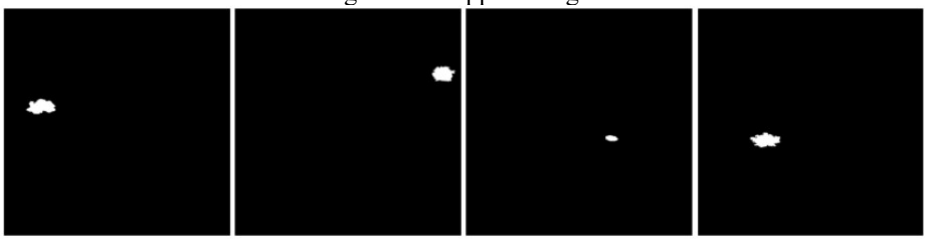


Figure 3: Region of Interest (ROI) Mask

The mass training sets contains the following attribute columns:

```
Data columns (total 14 columns):
 #   Column                  Non-Null Count  Dtype
---  ------                  --------------  -----
 0   patient_id              1546 non-null   object
 1   breast density          1546 non-null   int64
 2   left or right breast    1546 non-null   object
 3   image view              1546 non-null   object
 4   abnormality id          1546 non-null   int64
 5   abnormality type        1546 non-null   object
 6   calc type               1526 non-null   object
 7   calc distribution       1170 non-null   object
 8   assessment              1546 non-null   int64
 9   pathology               1546 non-null   object
 10  subtlety                1546 non-null   int64
 11  image file path         1546 non-null   object
 12  cropped image file path 1546 non-null   object
 13  ROI mask file path      1546 non-null   object
dtypes: int64(4), object(10)
```

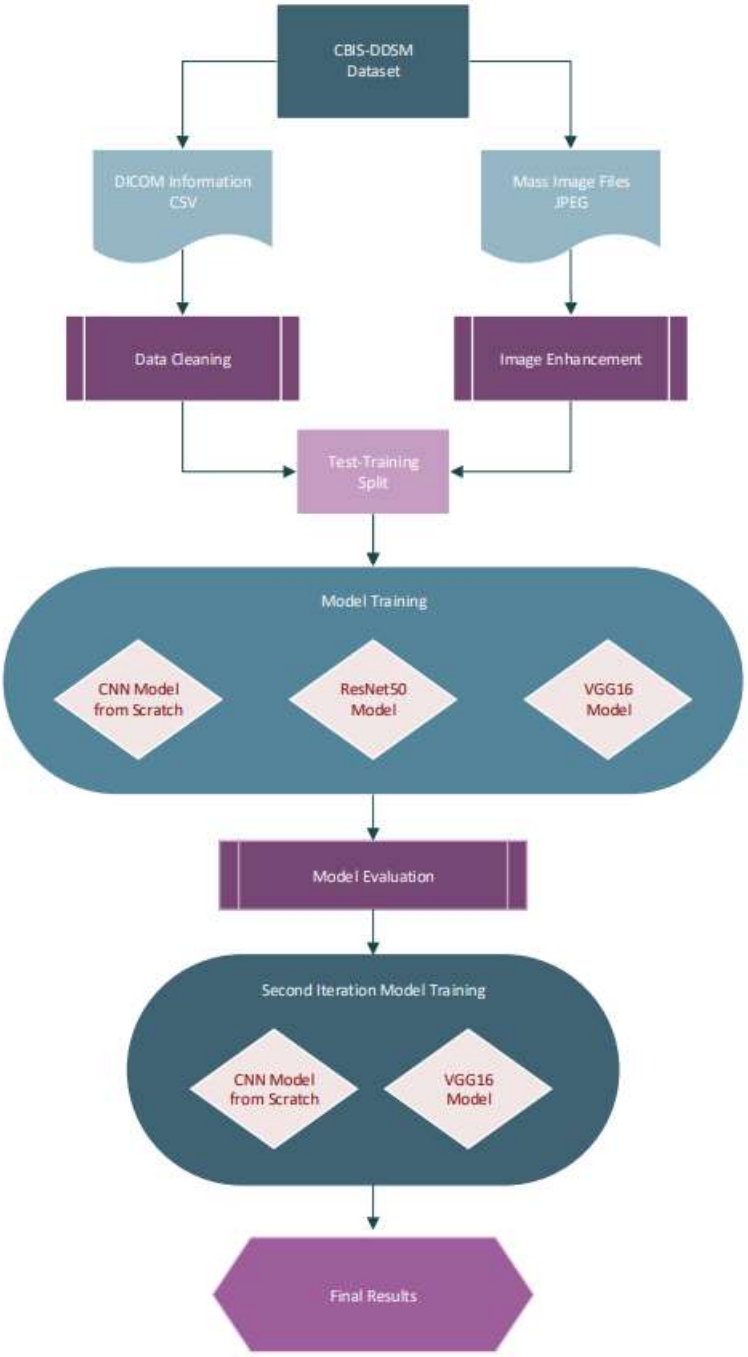
Figure 4: Dataframe information on the training and test sets within the CBIS-DDSM dataset

The DICOM data contain the following attribute columns :

```
Data columns (total 38 columns):
 #   Column                                        Non-Null Count  Dtype
---  ------                                        --------------  -----
 0   file_path                                     10237 non-null  object
 1   image_path                                    10237 non-null  object
 2   AccessionNumber                               0 non-null      float64
 3   BitsAllocated                                 10237 non-null  int64
 4   BitsStored                                    10237 non-null  int64
 5   BodyPartExamined                              10237 non-null  object
 6   Columns                                       10237 non-null  int64
 7   ContentDate                                   10237 non-null  int64
 8   ContentTime                                   10237 non-null  float64
 9   ConversionType                                10237 non-null  object
 10  HighBit                                       10237 non-null  int64
 11  InstanceNumber                                10237 non-null  int64
 12  LargestImagePixelValue                        10237 non-null  int64
 13  Laterality                                    9671 non-null   object
 14  Modality                                      10237 non-null  object
 15  PatientBirthDate                              0 non-null      float64
 16  PatientID                                     10237 non-null  object
 17  PatientName                                   10237 non-null  object
 18  PatientOrientation                            10237 non-null  object
 19  PatientSex                                    0 non-null      float64
 20  PhotometricInterpretation                     10237 non-null  object
 21  PixelRepresentation                           10237 non-null  int64
 22  ReferringPhysicianName                        0 non-null      float64
 23  Rows                                          10237 non-null  int64
 24  SOPClassUID                                   10237 non-null  object
 25  SOPInstanceUID                                10237 non-null  object
 26  SamplesPerPixel                               10237 non-null  int64
 27  SecondaryCaptureDeviceManufacturer            10237 non-null  object
 28  SecondaryCaptureDeviceManufacturerModelName   10237 non-null  object
 29  SeriesDescription                             9671 non-null   object
 30  SeriesInstanceUID                             10237 non-null  object
 31  SeriesNumber                                  10237 non-null  int64
 32  SmallestImagePixelValue                       10237 non-null  int64
 33  SpecificCharacterSet                          10237 non-null  object
 34  StudyDate                                     9671 non-null   float64
 35  StudyID                                       10237 non-null  object
 36  StudyInstanceUID                              10237 non-null  object
 37  StudyTime                                     9671 non-null   float64
dtypes: float64(7), int64(12), object(19)
```

Figure 5: Dataframe information on the DICOM_info.csv file within the CBIS-DDSM dataset

Approach


CBIS-DDSM Dataset
DICOM Information CSV
Mass Image Files JPEG
Data Cleaning
Image Enhancement
Test-Training Split
Model Training
CNN Model from Scratch
ResNet50 Model
VGG16 Model
Model Evaluation
Second Iteration Model Training
CNN Model from Scratch
VGG16 Model
Final Results

Data Preprocessing

With regards to the DICOM information, little to no data manipulation was required. As the project focused on image classification, most of the DICOM information was only edited for data visualization purposes. There was two options for classification, a binary classification for pathology: malignant or benign, or the option to classify into 4 categories for Breast Cancer Assessment: undetermined, well-differentiated, moderately-differentiated, poorly-differentiated, and undifferentiated. For the purposes of this project, it was decided to go with just the 2-class classification models due the imbalance of pathology within the categories. For the pathology types, BENIGN_WITHOUT_CALLBACK indicates an abnormal image but was later deemed to be benign and therefore was grouped together with the BENIGN class during modelling.

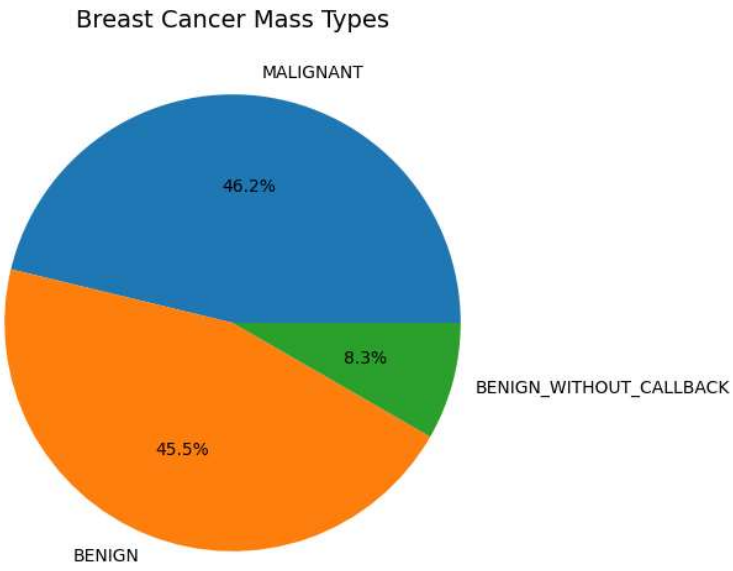


Figure 6: Breast Cancer Mass Types by Pathology.

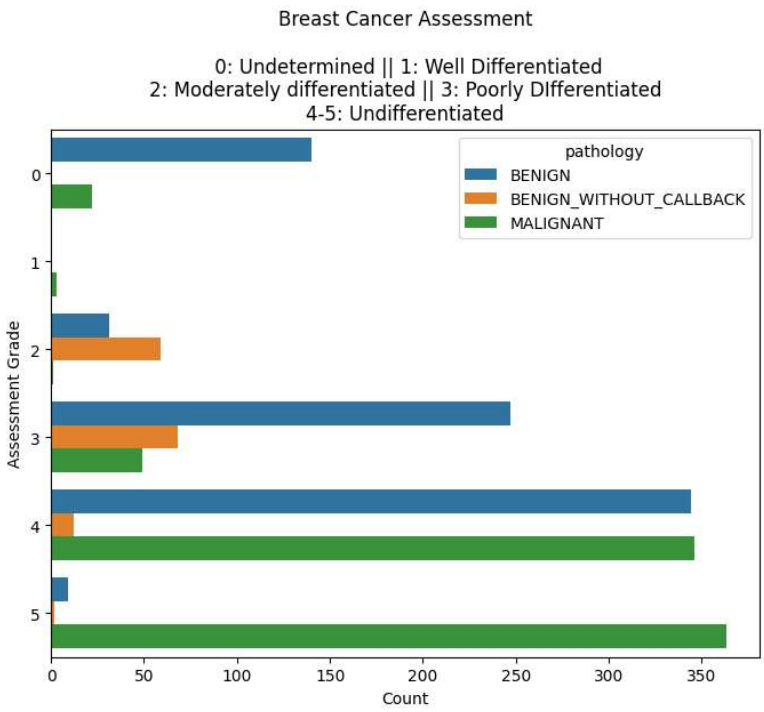


Figure 7: Breast Cancer Assessment by Pathology

Within the image files, there where 3 types of mammography images: cropped images, full mammogram images, and the ROI (region of interest) mask images. When looking at the distribution of image types, they were all fairly even and therefore there were no additional dataset imbalance processes used.

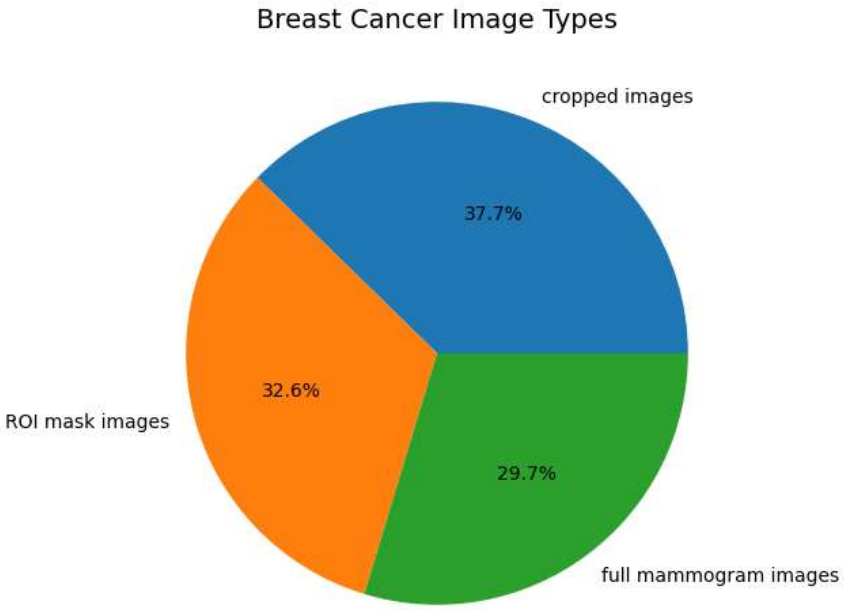


Figure 8: Breast Cancer Image Types

Image Enhancement

To prepare the images for modelling, image enhancement process was performed by creating a function to process all the images by first obtaining the absolute file paths of the image file, then reads the image file and returns a NumPy array to represent the image file. Then it will convert the colour space of the image to enhance the image. It will also resize the image to the set image target size. It will also normalize the pixel values of the image. And finally, it will return a processed image as a NumPy array ready for modelling. This step was definitely the most time consuming portion of the modelling process. Due to processing capacity, it was often take 15-20 minutes for all the images to complete.

```
def image_processor(image_path, target_size):
    absolute_image_path = os.path.abspath(image_path)
    image = cv2.imread(absolute_image_path)
    image = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)
    image = cv2.resize(image, (target_size[1], target_size[0]))
    image_array = image / 255.0
    return image_array
```

Figure 9: Image Enhancement function for processing the image files.

Simple CNN Model

The CNN modelling from scratch consisted of developing an overly simple CNN model and then slowly building it up. The first model consists of a 3 convolutional layers with 32, 64, and 128 filters respectively with interleaved max-pooling to reduce the spatial dimensions between the previous convolutional layer. A flattened layer is added and then a fully connected dense layer with 48 units. A drop out later is also added at a rate of 0.5 to help prevent overfitting. Finally, a final dense layer with a single neuron with sigmoid activation generates the

output.

```
model0 = models.Sequential()

model0.add(layers.Conv2D(32, (3, 3), activation='relu', input_shape=(224, 224, 3)))
model0.add(layers.MaxPooling2D((2, 2)))
model0.add(layers.Conv2D(64, (3, 3), activation='relu'))
model0.add(layers.MaxPooling2D((2, 2)))
model0.add(layers.Conv2D(128, (3, 3), activation='relu'))
model0.add(layers.MaxPooling2D((2, 2)))

model0.add(layers.Flatten())
model0.add(layers.Dense(48, activation='relu'))
model0.add(layers.Dropout(0.5))
model0.add(layers.Dense(1, activation='sigmoid'))
```

Figure 10: CNN model0 architecture

Model0 was trained on only 50 epochs (due to multiple crashes) and the batch size is set to 32 since larger batch sizes require more memory. The test accuracy returned a value of 0.633 but when looking at the Training and Validation Accuracy and Validation Loss graphs, we can see that model showing signs of overfitting in the accuracy graph but training well in the loss graph.

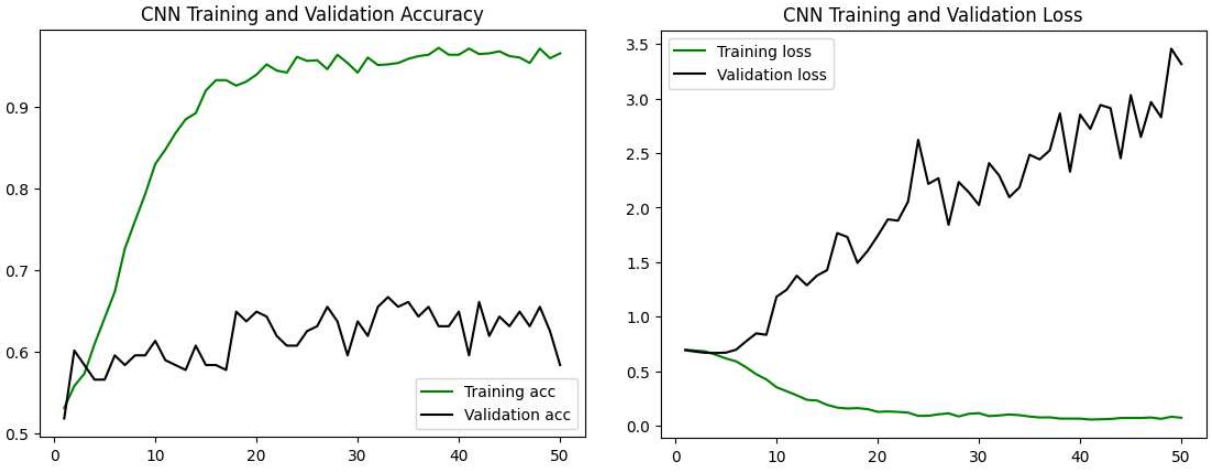


Figure 11: CNN Model0 Training and Validation Accuracy and Loss graphs.

When looking at the confusion matrix of the model0, we can definitely see overfitting occurring:

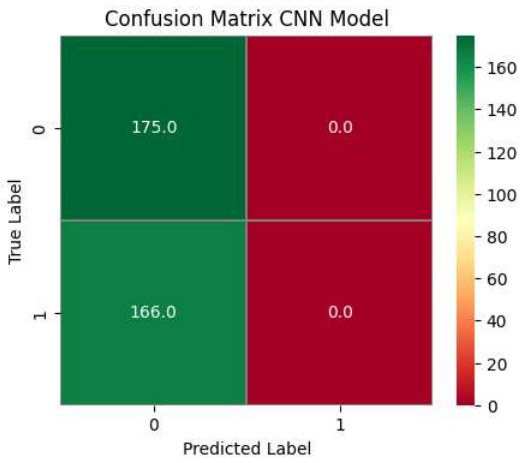


Figure 12: CNN Model0 confusion matrix

In a second iteration of the CNN modeling, model1 was built on with additional connected layers in an attempt to improve the model but the test accuracy returned 0.628, lower than our first iteration.

ResNet50 Model

ResNet50 (Residual Network – 50 layers) is a deep residual neural network architecture is runs deeper than CNN models with 50 layers. ResNet50 is known for its excellent performance in larger and complex image classification tasks and typically is more powerful.

For the ResNet50 architecture consisted on the first layer being the ResNet50 model, adding the global average pooling that reduces spatial dimensions, adding 3 dense layers of 256, 128, and 64 units to learn specific features and patterns from the previous layers outputs. There was also a dropout layer added to introduce regularization and a final dense layer of 2 neurons.

```
resnet50_model = ResNet50(weights='imagenet', include_top=False, input_shape=(224, 224, 3))

for layer in resnet50_model.layers:
    layer.trainable = False

model2 = models.Sequential()
model2.add(resnet50_model)
model2.add(layers.GlobalAveragePooling2D())
model2.add(layers.Dense(256, activation='relu'))
model2.add(layers.Dense(128, activation='relu'))
model2.add(layers.Dense(64, activation='relu'))
model2.add(layers.Dropout(0.5))
model2.add(layers.Dense(2, activation='sigmoid'))
```

Figure 13: ResNet50 Model2 architecture

Model2 was trained on only 50 epochs (due to multiple crashes) and the batch size is set to 32 since larger batch sizes require more memory. The test accuracy returned a value of 0.513 and I noticed that during the training, all the validation accuracy values were constant at 0.5179 after the second epoch. When looking at the Training and Validation Accuracy and Validation Loss graphs, we can see that model2 training accuracy becomes stagnant at 0.5746 after epoch 38. Stagnant accuracy is a sign that the model may not have enough capacity or complexity to learn or alternatively be too limited there is something preventing the model from learning. I attempted a second iteration, model3 to modify the ResNet50 multiple times to resolve this sudden plateauing but only to discover on other forums that this issue is related to batch normalization and that my initial preprocessing step to enhance the image files. I only later learned that when using ResNet50, a specific ResNet50 preprocess_input function needed to be performed on the input image files, which is clearly the issue regarding the stagnant accuracy results.

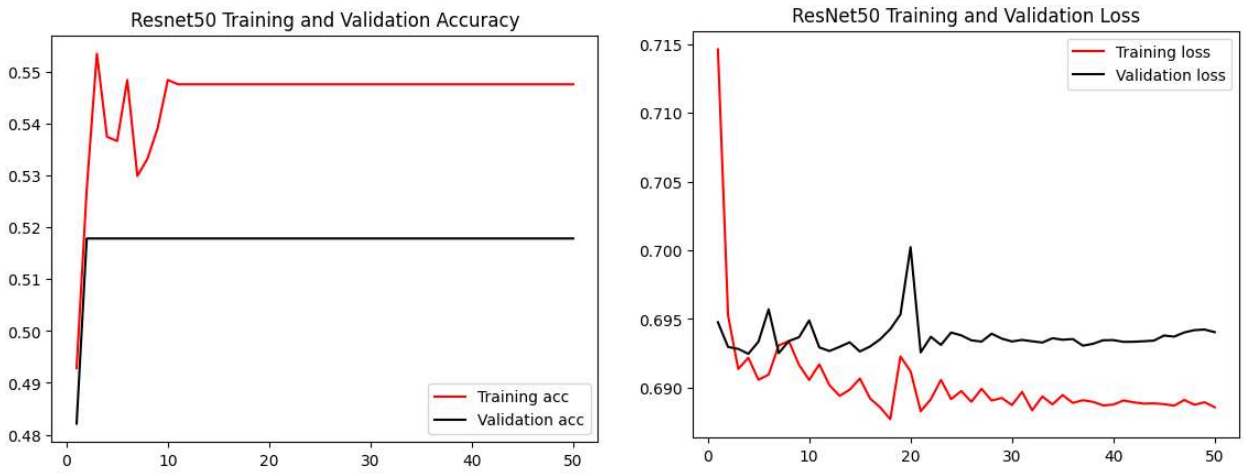


Figure 14: ResNet50 Model2 Training and Validation Accuracy and Loss graphs.

The confusion matrix for the ResNet50 model, curiously return the same output as the CNN Model0 showing an overfit model. I did perform a second iteration of the ResNet50 model but due to the fact that I was not able to resolve the initial image preprocessing issue, my second iteration return the exact same results even after adding and removing layers.

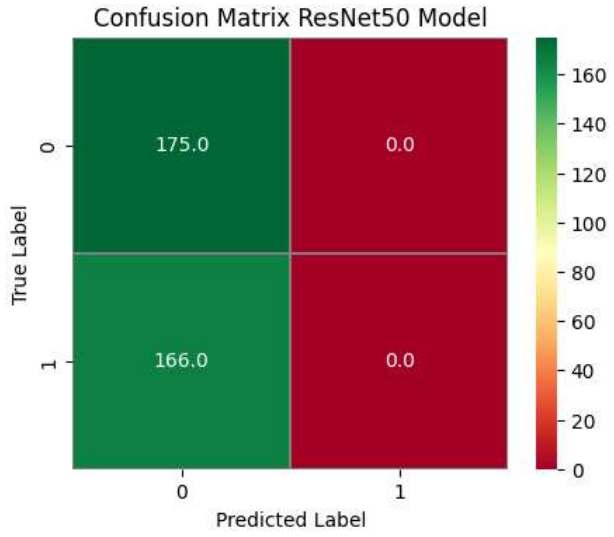


Figure 15: ResNet50 Model2 Confusion Matrix

VGG16 Model

The VGG16 (Visual Geometry Group – 16 layers) is a convolutional model architecture that performs well on a wide range of image classification tasks. The VGG16 model is the first layer and serves as a feature extractor for the subsequent layers. A global average pooling layer is added to reduce spatial dimensions. A dropout layer is added to prevent overfitting by introducing regularization and reducing interdependencies. 3 additional layers were added with 256, 128, and 64 units respectively, with a final dense layer with 2 neuron output. In the previous CNN and ResNet50 model, I used sigmoid as the activation but for VGG16, I decided to use softmax instead.

```
vgg16_model = tf.keras.applications.vgg16.VGG16(
    weights = 'imagenet',
    include_top = False,
    input_shape = (224, 224, 3)
)

for layer in vgg16_model.layers:
    layer.trainable = False

model4 = tf.keras.Sequential()
model4.add(vgg16_model)
model4.add(tf.keras.layers.GlobalAveragePooling2D())
model4.add(tf.keras.layers.Dropout(0.5))
model4.add(tf.keras.layers.Dense(256, activation = 'relu'))
model4.add(tf.keras.layers.Dense(128, activation = 'relu'))
model4.add(tf.keras.layers.Dense(64, activation = 'relu'))
model4.add(tf.keras.layers.Dense(2, activation = 'softmax'))
```

Figure 16: VGG16 Model4 architecture

Model4 was trained on only 50 epochs (due to multiple crashes) and the batch size is set to 32 since larger batch sizes require more memory. The test accuracy returned a value of 0.660. When looking at the Training and Validation Accuracy and Validation Loss graphs, we can see that model4 showed that both the training and validation accuracy values were increasing as the

number of epochs were increasing which means that model4 is learning well and the same can be said for the training and validation loss graph showing ideal conditions where the loss value is decreasing with increasing number of epochs, telling us that the model is not overfitting.

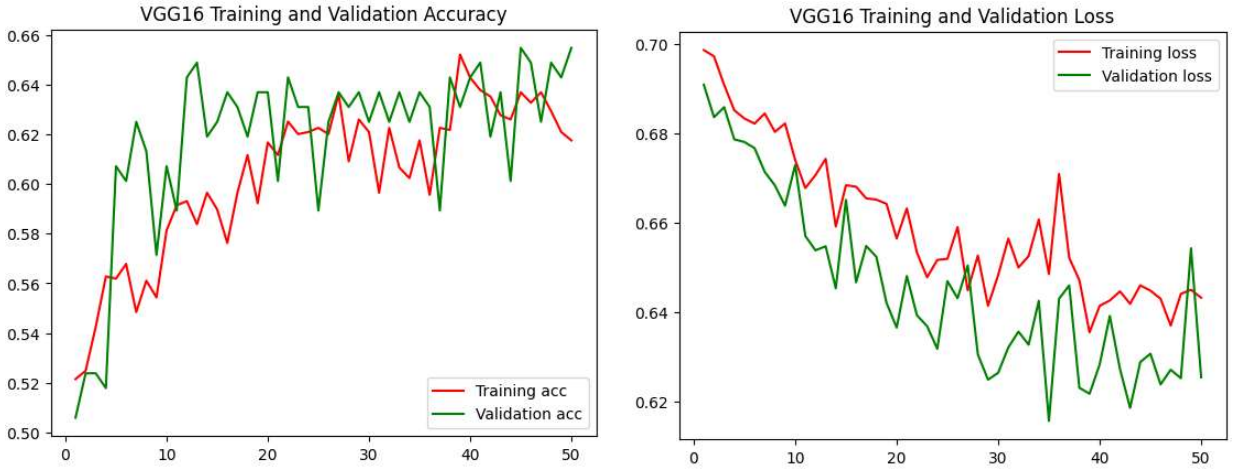


Figure 17: VGG16 Model4 Training and Validation Accuracy and Loss Graphs

The classification report for model4 also showed more impressive results compared to the CNN and ResNet50 models:

Classification Report:

	precision	recall	f1-score	support
0	0.73	0.53	0.62	175
1	0.62	0.80	0.69	166
accuracy			0.66	341
macro avg	0.67	0.66	0.66	341
weighted avg	0.68	0.66	0.65	341

Figure 18: VGG16 Model4 Classification Report

The confusion matrix reflects the same results showing a more realistic image classification results:

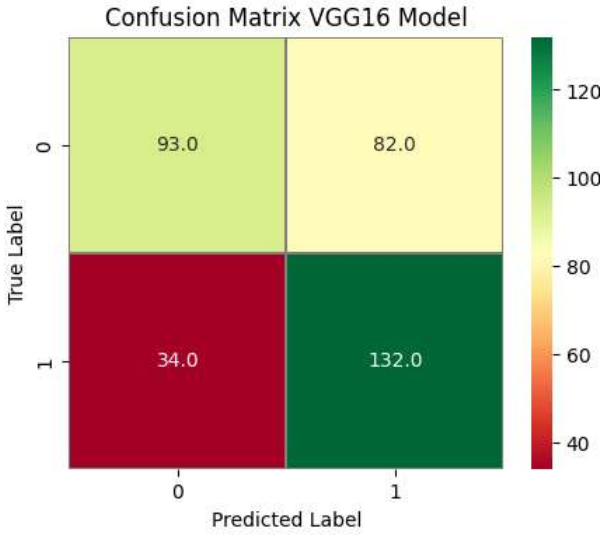


Figure 19: VGG16 Model4 Confusion Matrix

During the second iteration for the VGG16 model architecture, model5 included 2 additional layers with 32 and 16 units respectively and changed the activation back to sigmoid. The number of epochs and batch size remained the same and model5 was able to return a test accuracy of 0.667, a slight improvement from the first iteration model4. The training and validation accuracy and loss graphs had the same trend as the first iteration model 4 and returned a more precise confusion matrix.

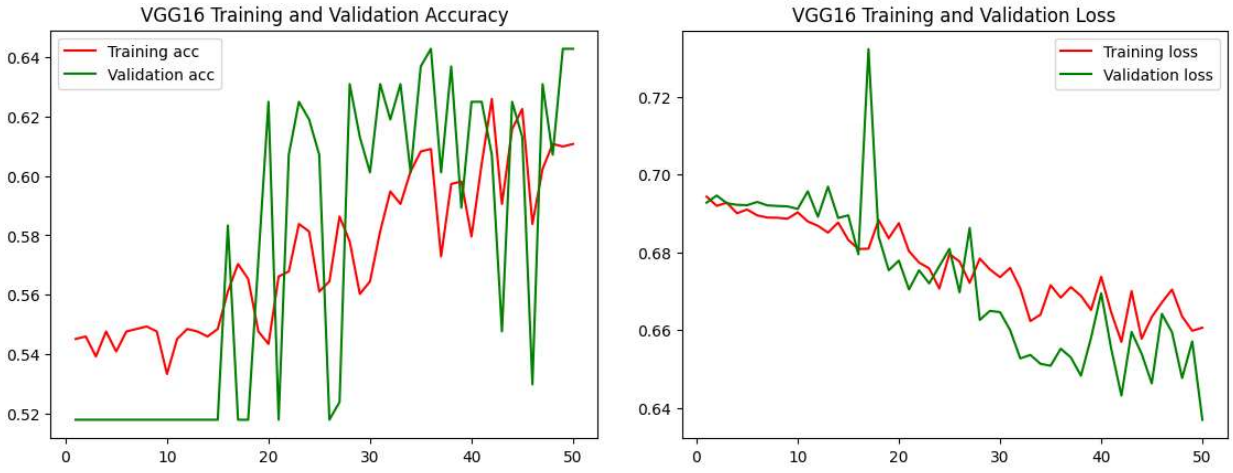


Figure 20: VGG16 Model5 Training and Validation Accuracy and Loss Graphs

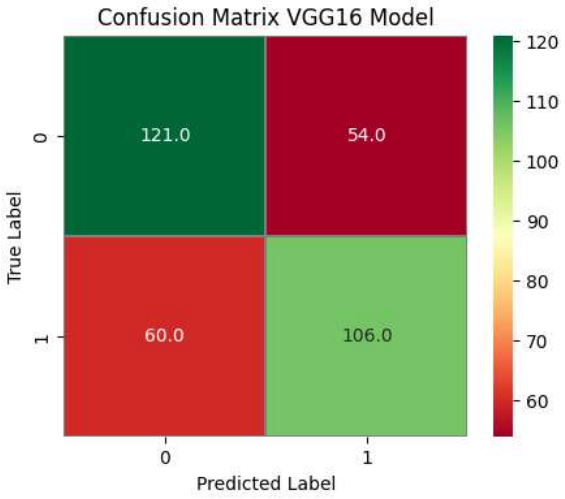


Figure 21: VGG16 Model5 Confusion Matrix

Results

The results of image classification modelling are as follows:

Model	Model#	Test Accuracy	Iteration
CNN	Model0	0.633	1
CNN	Model1	0.628	2
ResNet50	Model2	0.513	1
ResNet50	Model3	0.513	2
VGG16	Model4	0.660	1
VGG16	Model5	0.667	2

Conclusions and Recommendations

The best performing model from this project was with the use of VGG16 model architecture with the additional 5 dense layers and dropout. While the test accuracy is higher in the CNN model, the model was clearly over fitting. Unfortunately, the ResNet50 models were unable to correctly perform due to image preprocessing inadequacies. The biggest issue faced during this process was definitely computing power to handle the image processing. It is recommended that if this project were to be replicated, a computing with a powerful GPU is necessary to be able to run all the models without the environment crashing. I believe that with additional image enhancements such as image augmentation would be beneficial during the training process and running larger batch sizes.

References

1. Canadian Cancer Society. *Breast Cancer Statistics*. Retrieved from https://cancer.ca/en/cancer-information/cancer-types/breast/statistics
2. Sawyer-Lee, R., Gimenez, F., Hoogi, A., & Rubin, D. (2016). Curated Breast Imaging Subset of Digital Database for Screening Mammography (CBIS-DDSM) [Data set]. *The Cancer Imaging Archive*. Retrieved from https://doi.org/10.7937/K9/TCIA.2016.7O02S9CY
3. Kahn Jr, C. E., Roberts, L. M., Shaffer, K. A., & Haddawy, P. (1997). Construction of a Bayesian network for mammographic diagnosis of breast cancer. *Computers in biology and medicine, 27*(1), 19-29.
4. Cruz-Ramierz, N., Acosta-Mesa, H.G., Carillo-Calvet, H., Nava-Fernandez, L. A. & Barrientos-Martinez, R.E. (2007). *Diagnosis of breast cancer using Bayesian networks: A case study*. Alanazi, S. A., Kamruzzaman, M. M., Sarker, N. I., Alruwaili, M., Alhwaiti, Y., Alshammari, N., & Siddiqi, M. H. (2021). Boosting Breast Cancer Detection Using Convolutional Neural Network. *Journal of Healthcare Engineering, 2021*, 5528622.
5. Ezzat, D., & Hassanien, A. E. (2023). Optimized Bayesian convolutional neural networks for invasive breast cancer diagnosis system. *Applied Soft Computing*, 110810.
6. Kaymak, S., Helwan, A., & Uzun, D. (2017). Breast cancer image classification using artificial neural networks. *Procedia computer science, 120*, 126-131.
7. Shen, L., Margolies, L. R., Rothstein, J. H., Fluder, E., McBride, R., & Sieh, W. (2019). Deep learning to improve breast cancer detection on screening mammography. *Sci Rep, 9*. Retrieved from https://doi.org/10.1038/s41598-019-48995-4

Terminology

Benign	Non-cancerous conditions such as atypical hyperplasia and cysts. These can also lead to non-cancerous tumors such as intraductal papillomas.
Breast Cancer	A disease in which the cells within the breast change or no longer grow or behave abnormally without control. A cancerous tumour is a group of cancer cells that can destroy nearby tissue.
Breast Density	The proportion of fat to fibrous tissue. Mammography is more effective when screening breasts of less density.
Calcification	Deposits of calcium in the breast tissue. They are common and most are not associated with cancer but are often found in mammograms.
Carcinoma	Cancer that begins from cells that line glands and in the lining of internal organs.
Cyst	A fluid-filled sac, usually benign.
Diagnosis	Identification of the current condition
Lesion	Any focal abnormal area in the body. Can be used to describe a benign or malignant growth.
Malignant	Cancerous conditions such as ductal or lobular carcinoma. Less common breast cancer types include inflammatory breast cancer, Paget disease of the breast, and triple negative breast cancer. Rare types include non-Hodgkin lymphoma and soft tissue sarcoma.
Mammography	Low-dose x-ray imaging method used to examine the breast for the early detection of cancer and other breast related diseases
Mass	Lump, growth, or mass in the breast formed from a growth of tissue. Not all masses are cancerous.
Metastasize	The spread of a cancerous tumour beyond the primary site, often to other parts of the body
Oncology	The study and treatment of cancer. Doctors who specialize in oncology are called oncologists
Pathologist	Physician who identifies diseases by studying tissues or cells under a microscope
Prognosis	Description of the condition’s implications for future health
Radiologist	Physician who uses x-rays, ultrasound, MRI, etc. to aid in a diagnosis
Screening	Checking for the disease without the presence of any symptoms of the disease. Screening tests such as mammography or self screening by conducting self breast examinations