6-2021

Development of Novel Deep Learning Models in the Detection of Breast Cancer and Brain Aneurysm

Pradnya Patel

SANTA CLARA UNIVERSITY

DEPARTMENT OF BIOENGINEERING

Date: June 2021

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY

Pradnya R. Patel

ENTITLED

Development of novel deep learning models in the detection of Breast Cancer and Brain Aneurysm

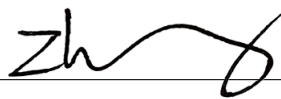BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN BIOENGINEERING (COMPUTATIONAL)

_____   June 10, 2021

Thesis Advisor   Date

_____   June 10, 2021

Department Chair   Date

Development of novel deep learning models in the detection of Breast Cancer and Brain Aneurysm

by

Pradnya Patel

Advisor: Dr. Yuling Yan

Submitted in partial fulfillment of the requirements for the degree
of
Master of Science in Bioengineering (Computational)
School of Engineering
Santa Clara University

Santa Clara, California
June 2021

Development of novel deep learning models for the detection of Breast Cancer and Brain Aneurysm

Pradnya Patel

Department of Bioengineering
Santa Clara University
June 2021

ABSTRACT

Three deep learning models using convolutional neural network (CNN) were developed for the early detection of breast cancer and brain aneurysm. Model 1 was built for the detection of breast mass; it consists of 20 total layers including 5 convolutional layers, 5 maxpool layers with Rectifier Linear Unit as the activation function for feature extraction, one flatten layer, 4 batch normalization with fully connected layers, and one output layer. This CNN model was trained and validated on an open-source breast ultrasound dataset that contains a total of 830 images categorized into three classes: normal (133 images), benign (487 images), and malignant (210 images). The model was tested on our dataset collected at Santa Clara Valley Medical Center (SCVMC) consisting of 300 cases with breast masses of which 194 are benign (64.7%) and 106 were malignant (35.3%). The final accuracy of the model on this test set achieved is 94.89%.

Models 2 & 3 are built for the detection of brain aneurysms as shown in magnetic resonance angiography (MRA) images. The MRA dataset containing contiguous MRA images representing normal and aneurysms were provided by collaborator radiologists at SCVMC. Model 2 comprises of 11 layers (3 convolutional layers, 3 max pool layers, 3 fully connected layers, a softmax layer, and an output layer) and was trained and validated on a larger open-source medical dataset (CBIS-DDSM mammogram) with 90% of which used for the training and 10% for the validation of the model. A transfer learning approach was then used to retrain the model for the detection of aneurysm using the MRA dataset, which contains 29024 images for normal and 25245 images for aneurysm cases collected from 100 healthy subjects and 100 patients with aneurysm. Model 3 represents the retrained model, which achieved a test accuracy of 76.04%, this is a preliminary result for this study.

Acknowledgements

I would like to thank my advisor, Dr. Yuling Yan, for being the greatest advisor in the world. I want to thank her for the countless hours she put in advising and guiding me as my mentor and thesis advisor during my study in the Bioengineering Department at Santa Clara University. Without her guidance, I would not have been able to achieve my dream. Dr. Yuling Yan went the extra mile to rekindle passion in me with her guidance and encouragement, when I had lost hope during this COVID-19 pandemic. I will forever be grateful to Dr. Yuling Yan for her patience, kindness, guidance and mentoring.

I express my deep and sincere gratitude to our collaborators, Dr. Young S. Kang, Dr. Mahesh Patel, Dr. Richard Lee, Dr. Aileen Chang, Dr. Christopher Nguyen, Dr. Ran Pang from SCVMC for giving me the opportunity to do research and for their hard work.

Lastly, I cannot thank my family enough for supporting my dream, sending me abroad for studies and believing in me. You have been with me every step of the way, through good times and bad, it would not have been possible without you all. Thank you for all the unconditional love and support you have given to me, Thank you for everything. I hope I have made you proud.

Table of Contents

List of Figures

Glossary of Terms

1. SCVMC - Santa Clara Valley Medical Center

2. CBIS-DDSM - Curated Breast Imaging Subset of DDSM

3. CNN - Convolutional Neural Network

4. ML - Machine learning

5. DL - Deep Learning

6. AI - Artificial Intelligence

7. ANN - Artificial Neural Network

8. MRA - Magnetic Resonance Angiography

9. MRI- Magnetic Resonance Imaging

10.   MIP- Maximum Intensity Projection

11.   ReLu – Rectifier Linear Unit

Chapter 1

Introduction

1.1   Background and Motivation.

Worldwide, Cancer is one of the leading cause of deaths [1]. There are numerous challenges faced by doctors as well as researchers in the fight of cancer. According to the American cancer society, 96,480 deaths are expected due to skin cancer, 142,670 from lung cancer, 42,260 from breast cancer, 31,620 from prostate cancer, and 17,760 deaths from brain cancer in 2019 (American Cancer Society) [2]. In this study, breast cancer is targeted. Breast cancer is the most frequently diagnosed cancer in women and it poses a serious threat to women's health. Breast cancer is the second most common cancer among women in the United States (some kinds of skin cancer are the most common). In 2017, the latest year for which incidence data are available, in the United States, 1,701,315 new cases of cancer were reported, and 599,099 people died of cancer. For every 100,000 people, 438 new cancer cases were reported and 153 people died of cancer [3]. Thus, early detection and proper treatment can improve patient prognosis and also is a top priority in saving lives. Breast ultrasound is one of the most commonly used modalities for diagnosing and detecting breast cancer in clinical practice. Using different imaging modalities considerable efforts have been made to improve the accuracy of breast cancer diagnosis. Breast ultrasound is one of the imaging technique that is expected to emerge as a complementary screening method for women with mammographically dense breasts, and the screening practice is expected to detect tumors at an early stage and reduce breast cancer mortality in women [4] .

The medical images of breast cancer using ultrasound scan are used in this research to detect cancer. Ultrasound and magnetic resonance imaging have been widely used to detect breast cancers in high risk patients. For this research, an open source Ultrasound dataset has been used [5].

The second disease that this study targets is intracranial aneurysms. An estimated 6.5 million people in the United States or 1 in 50 people, have an unruptured brain aneurysm. The annual rate of rupture is approximately 8 – 10 per 100,000 people. About 30,000 people in the United States suffer a brain aneurysm rupture each year. A brain aneurysm ruptures every 18 minutes [6]. Despite the widespread availability of brain imaging that can detect a ruptured brain aneurysm, misdiagnosis or delays in diagnosis occur in up to one quarter of patients when initially seeking medical attention [7]. In three out of four cases, misdiagnosis results from a failure to do a scan [8]. In the general population intracranial aneurysms are relatively common life-threatening diseases with a prevalence of 3.2% [9] and in the spontaneous subarachnoid hemorrhage patients, intracranial aneurysms account for 85% [10]. Intracranial aneurysms are now detected in more patients at early stages due to the advancement of various imaging techniques. Although aneurysmal subarachnoid hemorrhage accounts for 5–10% of all strokes in the United States [11] , it may cause significantly high mortality [12] , and the survivors may suffer from long-term neuropsychological effects and decreased quality of life [13] . For patients with spontaneous subarachnoid hemorrhage, early diagnosis of underlying disease can both influence clinical management and guide prognosis in intracerebral hemorrhage patients [13]. Henceforth, early and accurate detection of unruptured aneurysms may significantly improve clinical outcomes.

Typically, for cancer diagnosis and intracranial aneurysms diagnosis, visual examination and manual techniques are used. This manual interpretation of medical images demands high time consumption and is highly prone to mistakes. Timely and accurate identification of both, intracranial aneurysms and breast cancer is critical for immediate intervention or surgical management, whereas for patients without Intracranial aneurysms , reliable exclusion of intracranial aneurysms is also important for specialized management [14]. This research aims to aid in early detection of brain aneurysm and breast cancer with the implementation of deep learning models.

In recent years, artificial intelligence (AI) / Deep Learning, has accomplished phenomenal results in image recognition, image detection, natural language processing, self-driving cars, chat bots, speech recognition, computer vision, online recommendation systems, bioinformatics, and videogames etc.

Machine learning (ML) is the study of computer algorithms that improve automatically through experience and by the use of data [15]. It is seen as a subset of AI. ML algorithms build a model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to do so [16]. Deep learning (DL), which is a subset of ML, is a learning technique for computers and is based on artificial neural networks (ANNs) inspired by the structure and function of the biological brain. An ANN is based on a collection of nodes or connected units called artificial neurons, which loosely model the neurons in the brain. The DL model progressively learns eventually executes tasks with accuracy comparable to or surpassing human experts.

DL algorithms are used in a wide variety of applications where it is difficult or unfeasible to develop conventional ML algorithms to perform the needed tasks. DL achieves great power and flexibility by representing the world as a nested hierarchy of concepts, with each concept defined in relation to simpler concepts, and more abstract representations computed in terms of less abstract ones [17].

The learning of a model can be grouped into three categories: supervised learning, unsupervised learning and reinforcement learning. DL facilitates various techniques, and in this study main focus is on image detection. The most widely encountered applications are image classification, object detection, image segmentation, and synthetic imaging. Each application involves wide array of techniques based upon which the models are built. The end goal of these models is image detection. DL image classification and image detection in recent years has proven very promising in medical field as well. With more and more medical data being digitized as well as more medical data-set being published as open source, this culture has made it possible to utilize this situation to our advantage. For this study specifically as mentioned before, breast cancer and brain aneurysm detection are targeted. This goal of detecting breast cancer and brain aneurysm is achieved using supervised learning for image classification with DL model, in particular, convolutional neural network (CNN).

1.2 Previous Research

Research being conducted on the detection of disease using AI has been increasing substantially in past few years with the advancement of AI and increase in the digitization of checkup reports. A number of attempts have been made to detect breast cancer and brain aneurysm using DL models. Majority of researchers use pre-trained models such as AlexNet [18], which won the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) with overwhelming results for image classification. Other popular models include VGGNet [19], GoogLeNet [20], ResNet [21], and DenseNet [22].

With regard to the cancer detection DL models based on ultrasound images, some research has been conducted by Han et al. [23] used GoogLeNet [20] to distinguish the malignancy of breast masses on ultrasound with a large number of B-mode images collected by them to train a deep neural network with 3154 malignant and 4254 benign samples. They reported that the DL model achieved an accuracy of 91%, a sensitivity of 86%, a specificity of 93%, and an area under the curve (AUC) of greater than 0.9 [23].

Mango et al. [24] evaluated the utility of Koios DS Breast, a ML based diagnostic support system (Koios, https://koiosmedical.com/), by performing reading tests using the ultrasound images of 900 breast lesions. Among 15 physicians, the mean reader AUC for cases reviewed using ultrasound only was 0.83 versus 0.87 for ultrasound plus Koios DS Breast. Thus, Koios DS Breast improved the accuracy of sonographic breast lesion assessment while reducing inter- and intra-observer variability [24]. This diagnostic support system has been approved by the US Food and Drug Administration (FDA) and commercially launched.

Zhang et al. [25] built a DL architecture and evaluated its performance in the differentiation of benign and malignant breast tumors on a set of shear wave elastography (SWE) images with 135 benign and 92 malignant tumor cases. The DL architecture displayed good classification performance, with an accuracy of 93.4%, a sensitivity of 88.6%, a sensitivity of 97.1%, and an AUC of 0.947.

For intracranial aneurysm detection, the DNN model proposed by M. M. S. Bhurwani et al. [26] achieved an accuracy and ROC AUC of 72.4% and 0.80 respectively on un-normalized coiled data, 87.9% and 0.95 respectively on normalized coiled data, 73.9% and 0.79 respectively on un-normalized flow-diverted data, 85.3% and 0.80 respectively on normalized flow-diverted data, 62.9% and 0.64 respectively on un-normalized combined data, 64.8% and 0.73 respectively on normalized combined data.

The approach in this thesis differs from the research reported in the literature in terms of the data set used, the classification scheme and also the features used. Though both studies for detecting breast cancer using ultrasound showed high accuracy similar to results in this study, the work presented in this paper uses different approach than the given papers. Rather than using already trained models, this research introduces customized model, which is smaller in size but efficient and achieved better accuracy than the above mentioned models.

The research presented in this paper for detecting brain aneurysm is one of its kind according to the literature search as of June 2021. The novelty lies in that the model is pre trained on a mammogram dataset for breast cancer classification, and then it is transferred to learning the detection of brain aneurysm on MRA dataset. To my knowledge no similar approach has ever been used .

1.3 Aim and Contributions

As an aiding technology to possibly curb cancer and brain aneurysm and significantly improve patient outcomes, this thesis work focuses on the detection of breast cancer and brain aneurysm. To achieve this goal, I built and trained custom CNN models to differentiate between benign and malignant sonographic breast masses and to detect intracranial aneurysms using data input from MRA source images and maximum intensity projection (MIP) images.

Chapter 2

Overview and Concepts background

This section covers basic concepts used in the research.

2.1 CNN Architecture

There are several types of DL models including CNNs, Long Short Term Memory Networks (LSTMs), Recurrent Neural Networks (RNNs), and Generative Adversarial Networks (GANs). Of particular significance is the CNN model that has contributed greatly to the field of computer vision and medical image analysis.

CNNs are a class of Deep Neural Networks (DNNs) that can recognize and classify particular features from images and are widely used for analyzing visual images. Their applications range from image and video recognition, image classification, medical image analysis, computer vision and natural language processing.

The term 'Convolution" in CNN denotes the mathematical function of convolution which is a special kind of linear operation wherein two functions are multiplied to produce a third function which expresses how the shape of one function is modified by the other. In simple terms, in a given layer, the convolutional kernel slides along input matrix to generate a feature map which is fed as an input to the next layer.

There are three types of layers that make up the CNN, which are the convolutional layers, pooling layers, and fully-connected (FC)or dense layers. When these layers are stacked, a CNN architecture is formed as illustrated in figure 2.1. There are also two more important additions: dropout layer and activation function, which are described below.
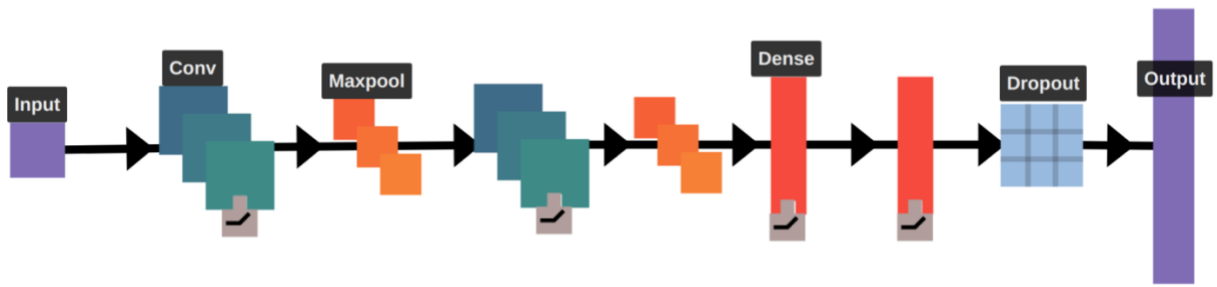
Figure 2.1: CNN architecture

Convolutional Layer: This layer is the first layer that is used to extract the various features from the input images. In this layer, the mathematical operation of convolution is performed between the input image and a filter of a particular size MxM. By sliding the filter over the input image, the dot product is taken between the filter and the parts of the input image with respect to the size of the filter (MxM).The output is termed as the Feature map which gives the information about the image such as the corners and edges. Later, this feature map is fed to other layers to learn several other features of the input image.

Pooling Layer: In most cases, a Convolutional Layer is followed by a Pooling Layer. The primary aim of this layer is to decrease the size of the convolved feature map to reduce the computational costs. This is performed by decreasing the connections between layers and independently operates on each feature map. Depending upon method used, there are several types of Pooling operations. In Max Pooling, the largest element is taken from feature map. Average Pooling calculates the average of the elements in a predefined sized Image section. The total sum of the elements in the predefined section is computed in Sum Pooling. The Pooling Layer usually serves as a bridge between the Convolutional Layer and the FC Layer.

Fully Connected Layer: The Fully Connected (FC) layer consists of the weights and biases along with the neurons and is used to connect the neurons between two different layers. These layers are usually placed before the output layer and form the last few layers of a CNN Architecture. In this, the input image from the previous layers are flattened and fed to the FC layer. The flattened vector then undergoes few more FC layers where the mathematical functions operations usually take place. In this stage, the classification process begins to take place.

Dropout: Usually, when all the features are connected to the FC layer, it can cause overfitting in the training dataset. Overfitting occurs when a particular model works so well on the training data causing a negative impact in the model's performance when used on a new data. To overcome this problem, a dropout layer is utilized wherein a few neurons are dropped from the neural network during training process resulting in reduced size of the model. On passing a dropout of 0.3, 30% of the nodes are dropped out randomly from the neural network.

Activation Functions: Finally, one of the most important hyperparameters of the CNN model is the activation function. They are used to learn and approximate any kind of continuous and complex relationship between variables of the network. In simple words, it decides which information of the model should fire in the forward direction and which ones should not at the end of the network. It adds non-linearity to the network. There are several commonly used activation functions such as the ReLU, Softmax, Tanh and the Sigmoid functions. Each of these functions has a specific usage. For example, for a binary classification CNN model, sigmoid and softmax functions are preferred and for a multi-class classification, generally softmax is used.

2.2   Transfer Learning

A pre-trained model is a saved network that was previously trained on a dataset, typically on a large-scale image-classification task. Either the pretrained model is used as is or transfer learning is used to customize this model to a given task.

The intuition behind transfer learning for image classification is that if a model is trained on a large and general enough dataset, this model will effectively serve as a generic model of the visual world. Advantage can be taken of these learned feature maps without having to start from scratch by training a large model on a large dataset.

17

There are two ways to customize a pretrained model which are as follows:

1. By Extracting Features: Feature extraction can be used to customize pretrained network. Use the representations learned by a previous network to extract meaningful features from new samples. Simply add a new classifier, which will be trained from scratch, on top of the pretrained model so that you can repurpose the feature maps learned previously for the dataset.
There is no need to (re)train the entire model. The base convolutional network already contains features that are generically useful for classifying pictures. However, the final, classification part of the pretrained model is specific to the original classification task, and subsequently specific to the set of classes on which the model was trained.

2. By Fine Tuning: Unfreeze a few of the top layers of a frozen model base and jointly train both the newly-added classifier layers and the last layers of the base model. This allows us to "fine-tune" the higher-order feature representations in the base model in order to make them more relevant for the specific task.

2.3   Breast Cancer

Breast cancer is a type of cancer that starts in the breast. Cancer starts when cells begin to grow out of control. Breast cancer cells usually form a tumour that can often be seen on an x-ray or felt as a lump. Breast cancer occurs almost entirely in women, but men can get breast cancer, too. Most breast lumps are benign and not cancerous (malignant). Non-cancerous breast tumours are abnormal growths, but they do not spread outside of the breast and thus are not life threatening. However, some types of benign breast lumps can increase a woman's risk of getting breast cancer. Any breast lump or change needs to be checked by a health care professional to determine if it is benign or malignant and if it might contribute to future cancer risk.

2.4.   Brain Aneurysm

A brain aneurysm, also known as a subarachnoid hemorrhage (SAH), is a weak spot in the wall of a blood vessel inside the brain. Think of a weak spot in a balloon and how it feels stretched out and thin. A brain aneurysm is like that.

That area of the blood vessel gets worn out from constant flow of blood and bulges out, almost like a bubble. It can grow to the size of a small berry. There are different types:

a. Saccular aneurysms are the most common type of brain aneurysm. They bulge out in a dome shape from the main artery. They're connected to that artery by a narrow "neck."

b. Fusiform aneurysms aren't as common as saccular aneurysms. They don't pouch out in a dome shape. Instead, they make a widened spot in the blood vessel.

2.5.   Medical Imaging Techniques

Biomedical imaging is a useful tool for measuring the biodistribution, targeting, and elimination of nanostructures in real time. This is especially needed at the whole organism level. In order to provide sufficient imaging contrast, biomedical nanodevices can be designed with reporting functions or moieties that provide signal in conventional medical imaging modalities. These include gamma scintigraphy, magnetic resonance imaging (MRI), computed tomography (CT), positron emission tomography (PET), and ultrasound imaging.

2.5.1 MRA

Magnetic resonance angiography–also called a magnetic resonance angiogram or MRA–is a type of MRI that looks specifically at the body's blood vessels. Unlike a traditional angiogram, which requires inserting a catheter into the body, magnetic resonance angiography is a far less invasive and less painful test.

During magnetic resonance angiography, patient lie flat inside the magnetic resonance imaging scanner. This is a large, tunnel-like tube. In some cases, a special dye, known as contrast, may be added to patient's bloodstream to make your blood vessels easier to see. When needed, the contrast is given with an intravenous (IV) needle.

2.5.2 MRI

Magnetic resonance imaging (MRI) is a medical imaging technique that uses a magnetic field and computer-generated radio waves to create detailed images of the organs and tissues in your body. Most MRI machines are large, tube-shaped magnets. When you lie inside an MRI machine, the magnetic field temporarily realigns water molecules in your body. Radio waves cause these aligned atoms to produce faint signals, which are used to create cross-sectional MRI images — like slices in a loaf of bread. The MRI machine can also produce 3D images that can be viewed from different angles.

2.5.3 Mammogram

A mammogram is an X-ray picture of the breast. Doctors use a mammogram to look for early signs of breast cancer. Regular mammograms are the best tests doctors have to find breast cancer early, sometimes up to three years before it can be felt.

2.5.4 Ultrasound

Ultrasound imaging uses sound waves to produce pictures of the inside of the body. It is used to help diagnose the causes of pain, swelling and infection in the body's internal organs and to examine a baby in pregnant women and the brain and hips in infants. It's also used to help guide biopsies, diagnose heart conditions, and assess damage after a heart attack. Ultrasound is safe, noninvasive, and does not use ionizing radiation.

2.6 Data Sources used in this study

In Deep learning, data plays a crucial role as the data defines the models accuracy. Following datasets have been used in this study to train, validate and test the model-

2.6.1. Ultrasound dataset (open source)

An open-source dataset (by Al. Dhabyani et al [5]) is used in this thesis work. This is the first breast ultrasound dataset publicly available and hence provides researchers with the scope to

detect breast cancer in Ultrasound images and improve patient outcomes. The total number of images in this dataset is 830.

2.6.2. CBIS-DDSM dataset(open source)

The CBIS-DDSM (Curated Breast Imaging Subset of DDSM), includes decompressed images, data selection and curation by trained mammographers, updated mass segmentation and bounding boxes, and pathologic diagnosis for training data, formatted similarly to modern computer vision data sets. The data set contains total 1644 mass cases, providing a data-set size capable of analyzing decision support systems in mammography. [27]

2.6.3. SCVMC ultrasound dataset (our dataset)

This dataset was collected in 2018 at SCVMC and consists of a total of 600 images. The data collected at baseline include breast ultrasound images among women aged 25 to 75.

2.6.4. SCVMC MRA and MIP dataset (our dataset)

This dataset includes MRA source and MIP images with intracranial aneurysms that were collected from between January 2017 through December 2019. The dataset contains 25245 MRA images out of 100 cases with aneurysms and 29024 MRA images out of 100 cases without aneurysms.

Chapter 3

Methods and Evaluation

3.1. Detecting Breast Cancer using Ultrasound Images

3.1.1 Overview

     The first model (model-1) presented here is specifically designed for detecting breast cancer with Ultrasound images. The model is a CNN built with TensorFlow 2.0 and Keras open source framework along with matplotlib, OpenCV2, NumPy, pandas, python etc.
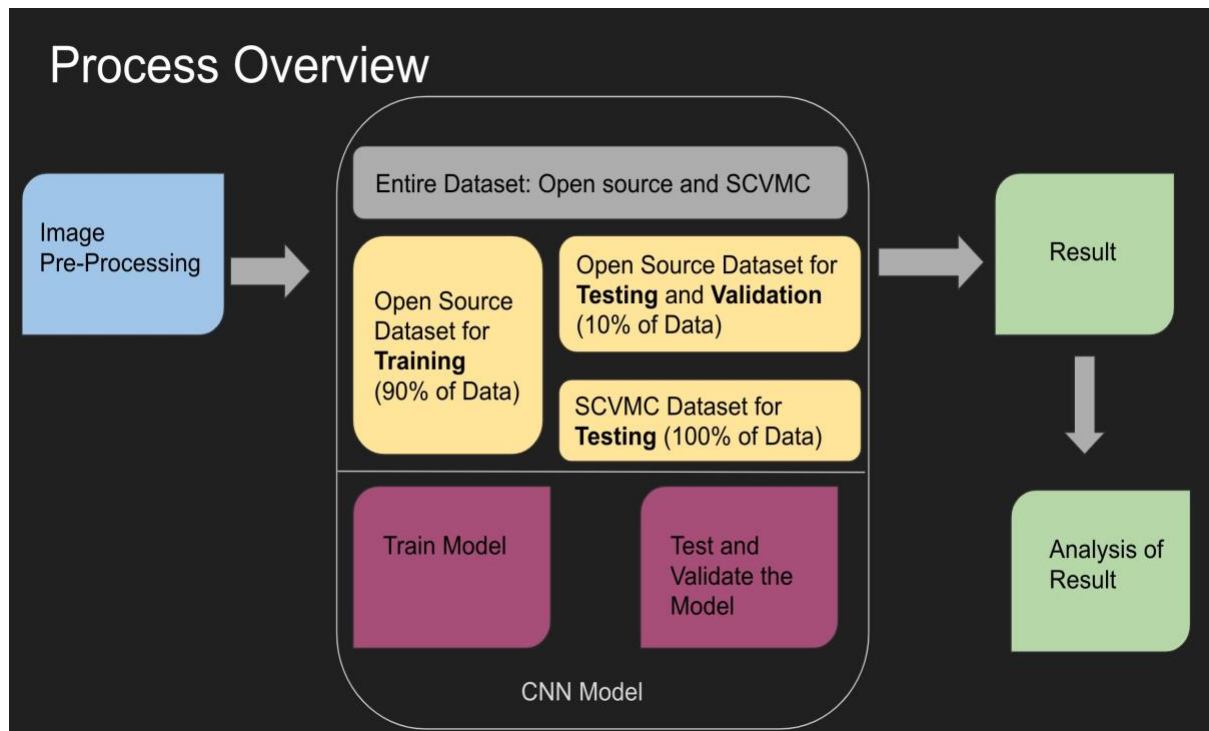

Figure 3.1: Process Overview

This model uses open source BUSI Ultrasound Dataset [5] for training and validation, and is subsequently tested on our dataset which was collected by collaborators at SCVMC.

3.1.2 Dataset

In this study, two sets of data are used as explained below in detail:

a. The Open Source BUSI Ultrasound Dataset: This open-source dataset, comprised of ultrasound images and the corresponding ground truth depicting the region of interest (ROI), is used for training the model. It was collected in 2018 and includes breast images collected from 600 female patients aged 25 to 75. The dataset consists of a total of 830 images stored in PNG format. These images are categorized into three classes: normal (133 images), benign (487 images), and malignant (210 images) [5].
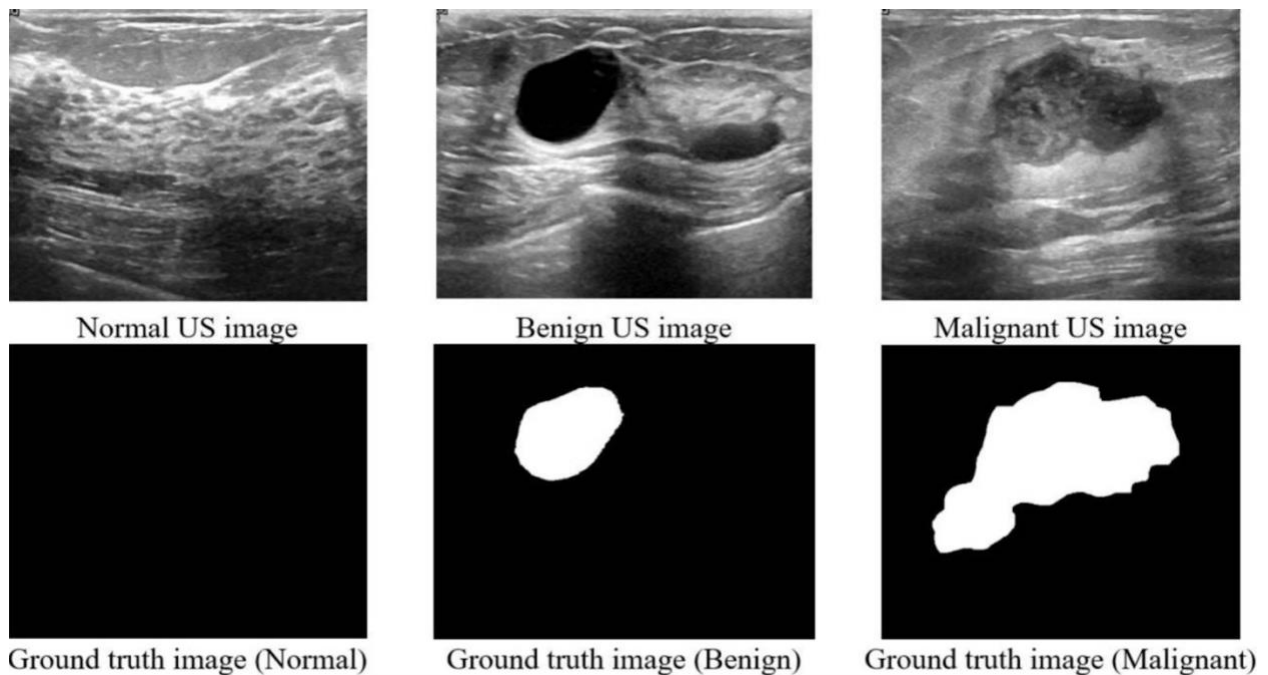

Figure 3.2: Sample images from BUSI open source dataset [5]

b. Our Ultrasound Dataset from SCVMC: This dataset is used for testing the model. This dataset was collected during 2017 - 2019 at our collaborator's institution (SCVMC). The dataset consists of 600 images in jpg format. The images are categorized into two classes: benign (388 images) and malignant (212 images). This data does not have ROI mask images.
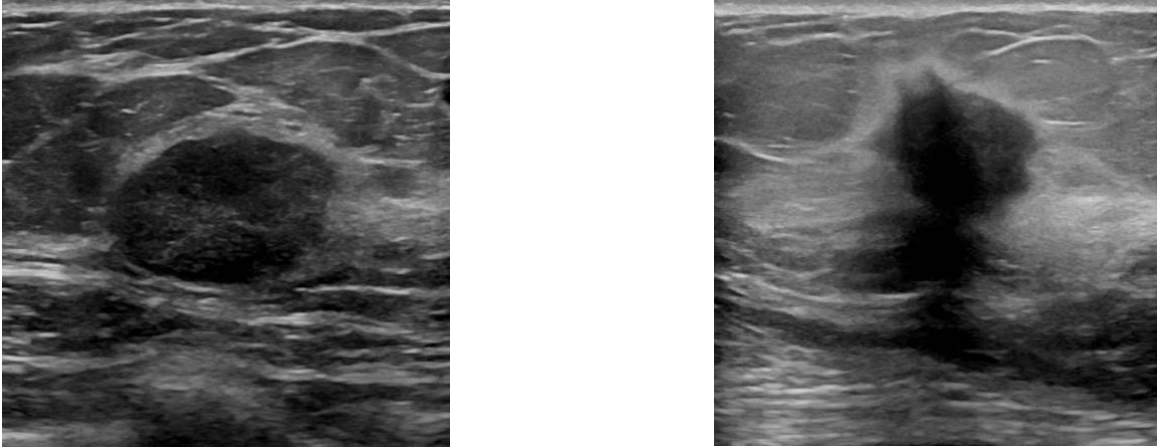
Figure 3.3: Sample ultrasound images from SCVMC dataset for each class, Benign (Left) and Malignant (Right)

3.1.3 Image Pre-processing and Data Augmentation

Image preprocessing in this project is performed that involved removing the background pixels and cropping the image using ground truth mask images provided in the open source dataset. The original image is multiplied with its mask image and then a boundary is marked around ROI and the rest of the image is cropped. This image pre-processing was not applied to images in the category of normal class. The pre-processing step is expected to improve the computational efficiency.

Data augmentation methods are applied to images to artificially expand the dataset with more data for training and validating the model. The data augmentation techniques applied include rescaling the image by 1/255, zoom with 0.2, rotating images horizontally and vertically, height and width shifting, shearing and horizontal flipping. Some sample images after a random combination of aforementioned operations are shown below:
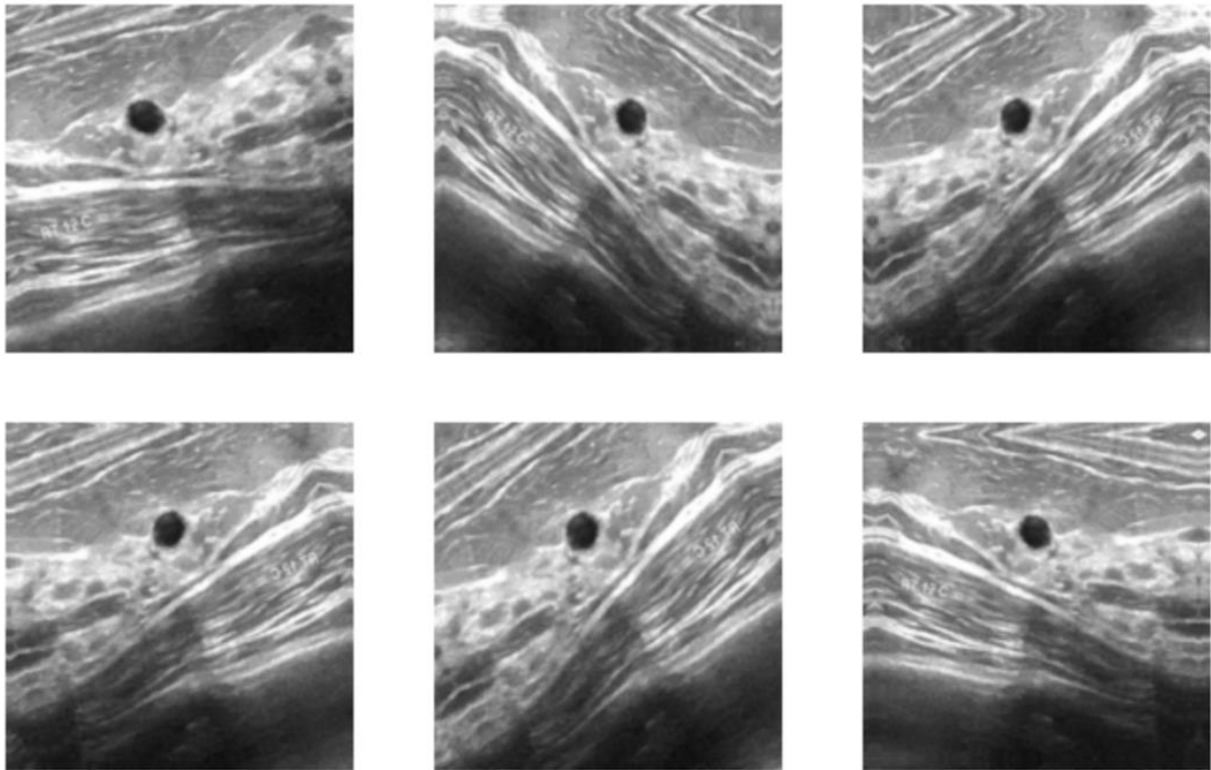
Figure 3.4: Sample images with random aforementioned data augmentation techniques applied

3.1.4 CNN Model-1 Architecture

Our CNN model-1 was trained and validated on an open source breast ultrasound dataset [5] with 100% of data used for training and the entire SCVMC dataset which is about 90% of the BUSI dataset was used for testing. The pre-processed images as discussed in previous section, the augmented images and the original images were used to train the model. The total number of parameters were 14,508,769.

The proposed model-1 consists of 20 total layers (including 15 layers of convolution and maxpool layers with ReLu as the activation function for feature extraction,1 flatten layer, dropout layer, global average pooling layer, and an output layer with Sigmoid function). The architecture of the model-1 with its parameters are represented below:
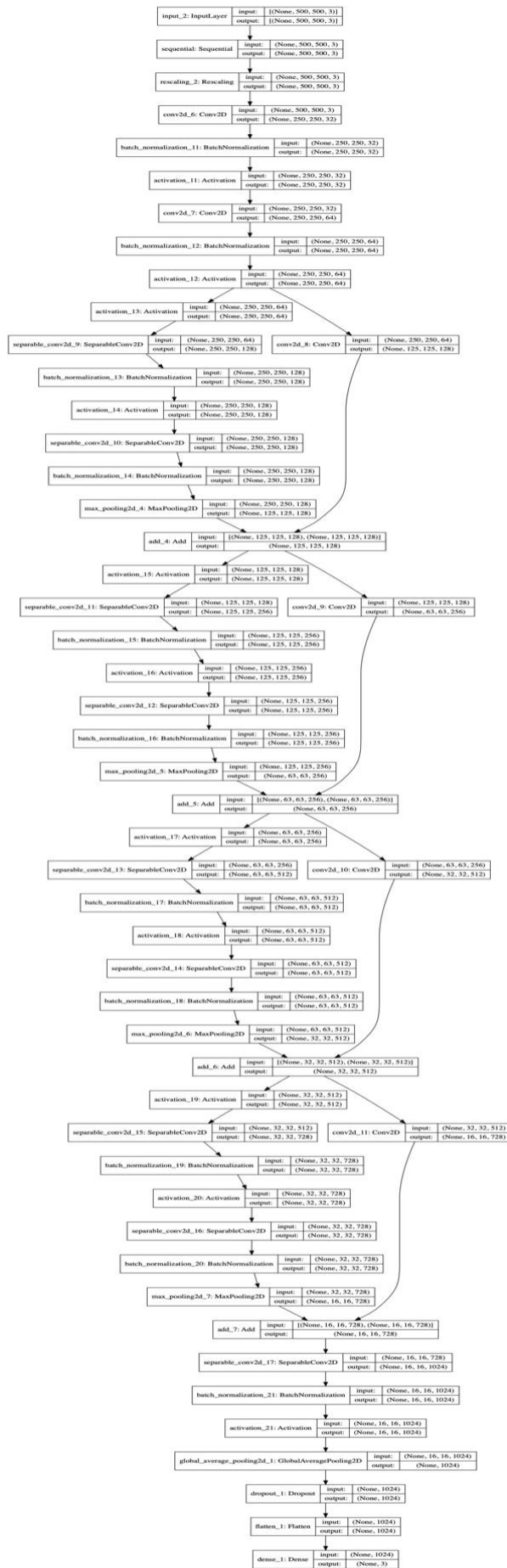
Figure 3.5: CNN model-1 architecture.

3.1.5   Results and Evaluation

The performance of model-1 was measured in 100 epochs, and the following results were obtained after final fine tuning of the model.
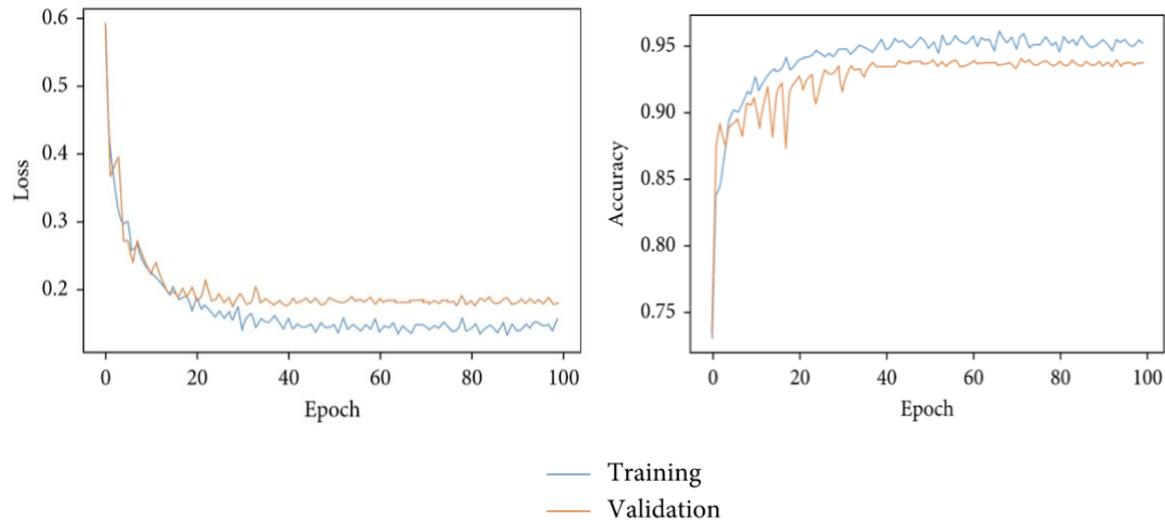

Figure 3.6: Performance of model-1: loss (left) and accuracy (right)

Of the 300 masses within the SCVMC dataset, 194 were benign (64.7%) and 106 malignant (35.3%). The proposed CNN model-1 demonstrated a high accuracy (94.89%) tested on the SCVMC dataset. The high performance of the CNN model-1 on a small dataset suggests its ability to utilize a smaller database for training. Preliminary results suggest that CNN may be utilized to better differentiate between benign and malignant sonographic breast masses. It also provides a way for radiologists to double check their work.

3.2 Detecting Brain Aneurysm using MRA

3.2.1 Overview

Models 2 and 3 are built for the detection of brain aneurysm using mammogram and MRA datasets. Model-2 uses the open source CBIS-DDSM Dataset for pre-training and validation, and Model-3 is retrained and tested on our SCVMC dataset.

3.2.2 Dataset

In this study two different datasets are used, which are explained in detail below:

a. CBIS-DDSM (Curated Breast Imaging Subset of DDSM): This dataset is used for training the model that includes decompressed images, data selection and curation by trained mammographers, updated mass segmentation and bounding boxes, and pathologic diagnosis for training data, formatted similarly to modern computer vision data sets. The data set contains 753 calcification cases and 891 mass cases [27].
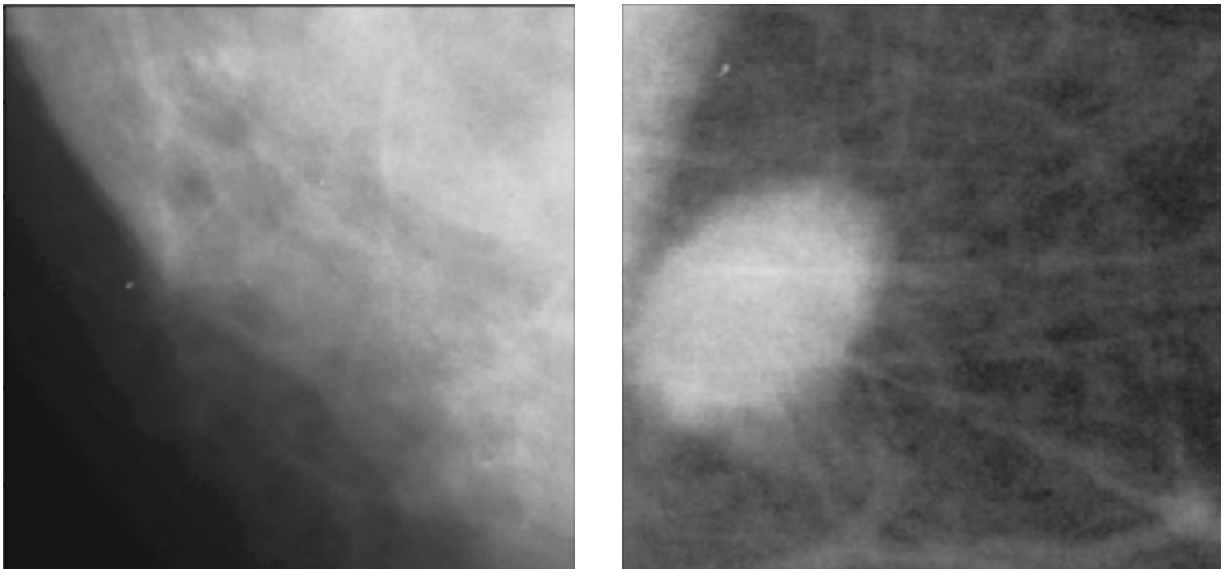

Figure 3.7: Sample images from the CBIS-DDSM dataset: Normal (Left) and Malignant (Right) [27]

b. MRA dataset from SCVMC: This dataset includes MRA source and MIP images with intracranial aneurysms. The images were divided into two datasets for training and testing respectively . This dataset contains 25245 MRA images out of 100 cases with aneurysms and 29024 MRA images out of 100 cases without aneurysms.
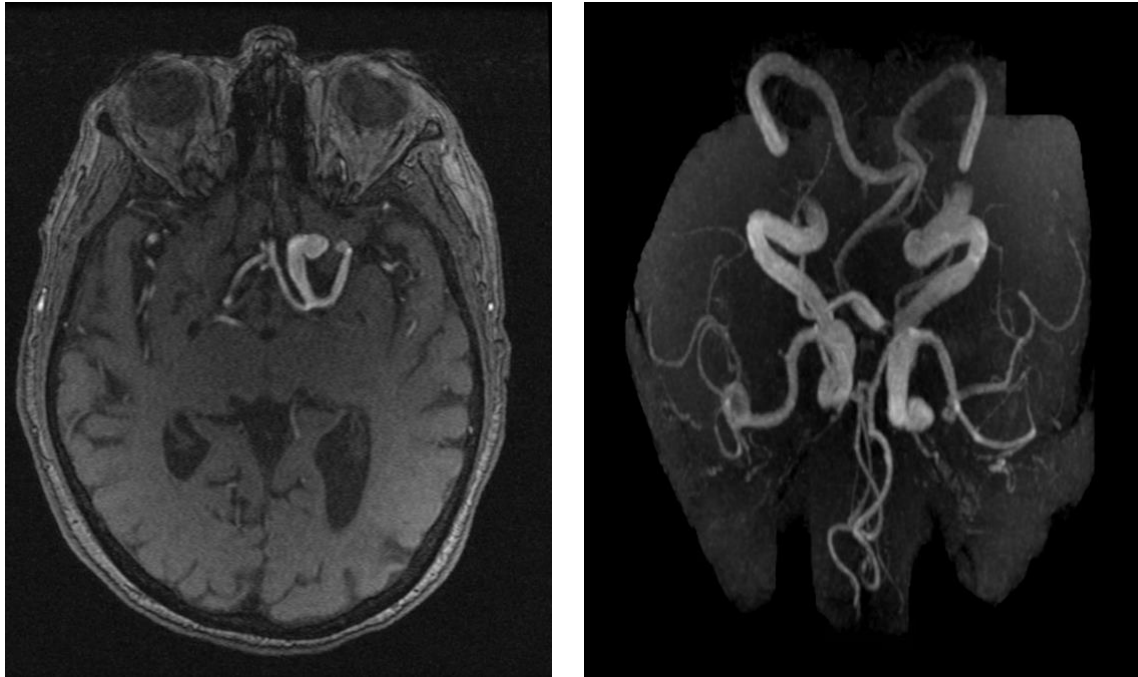
Figure 3.8: Sample images from SCVMC dataset: MRA (Left) and MIP (Right)

3.2.3. Image Pre-processing and Data Augmentation

The pre-processing of these images included rescaling to the pixel size for each image to 500 × 500 pixels. Similar pre-processing steps were performed (see previous section). In brief, the background pixels were removed by cropping the image, and each image was multiplied with its mask image, which is followed by making a boundary around the ROI and cropping the rest of the image to generate new input images each with reduced dimensions.

The preprocessing of MRA and MIP images also included removal of skull, darker background area and cropping the image to focus on ROI . The images were then resized to 500 x 500 pixels. Data augmentation techniques were applied to artificially expand the dataset with more data for training and validating the model. The data augmentation techniques applied include rescaling the image by 1/255 and rotating images horizontally and vertically.

3.2.4 CNN Model-2 & Model-3Architecture

Model-2 was trained and validated on CBIS-DDSM. The source task of this model was to detect breast cancer; it made predictions with 93.4% accuracy on the validation dataset (10% of CBIS-DDSM dataset). A new model for the target task of detecting brain aneurysm, model-3, was derived from model-2 by retraining and fine tuning the model using the transfer learning techniques. In particular, the MRA source images were used to train the new classifier, which was trained from scratch on top of model-2 to repurpose its previously learned feature maps. The MIP images were then used to fine tune the model. Few of the top layers were unfrozen and these layers were trained using above classifier and model-2's last layers. This allowed the model to be more relevant to detecting the brain aneurysm.

The 200 MRA cases from our SCVMC dataset were used, out of which 100 cases were of aneurysm and 100 cases were normal.

From the SCVM dataset: Training set included aneurysm and normal (70 cases each-17672 images) Test set aneurysm and normal (30 cases each- 7573 images). The images were again divided from their categories on the basis of MRA and MIP. Each of these image categories were used for above mentioned purpose.

The transfer learning approach proved advantageous as the learning was faster, more accurate and enabled us to generate great results with comparatively small MRA dataset. With a limited dataset, one can hardly train a complex CNN from scratch, and thus using a pretrained model is recommended for a classification task like the one in this study. Usually pretrained models are trained on images that are not related to medical field, thus they do not perform as well as those trained on a medical dataset of similar image features.

For this study, model-2 was trained on the open-source CBIS-DDSM dataset as mentioned above. The architecture of the model is shown in fig. 3.7. The model comprises of 6 layers of convolution and max pool layers, 4 dense layers, and an output layer.
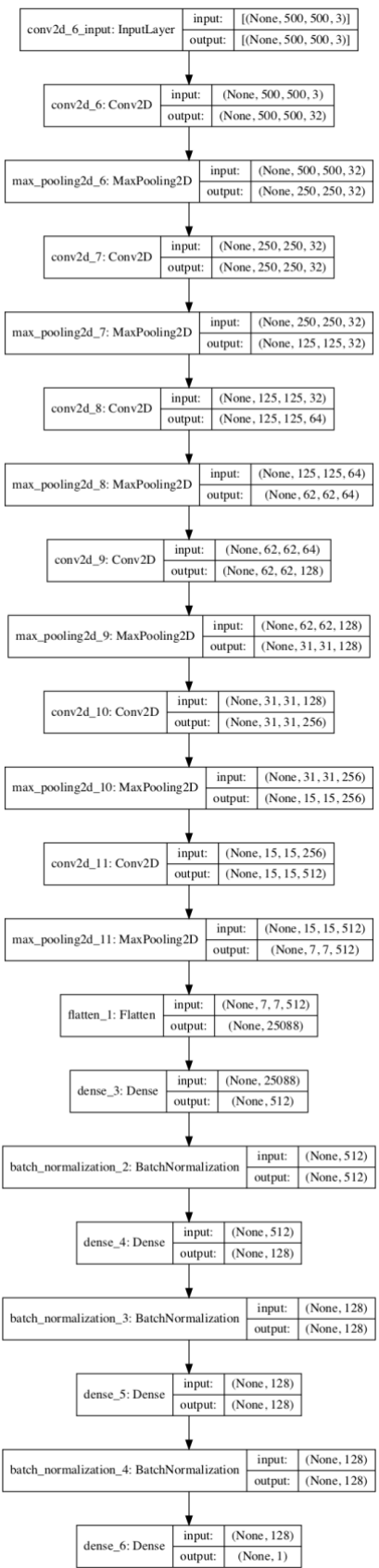
Figure 3.9: CNN model-2 & 3 architecture.

3.2.5 Results and Evaluation

The repurposed model for the detection of intracranial aneurysms in MRA source and MIP images achieved a test accuracy of 76.04%, this is only a preliminary result for this study. The model can be further enhanced to address the problem of overfitting. In addition, many other techniques such as adding custom layers, fine tuning the model with different learning rate etc. may be applied to improve the accuracy.
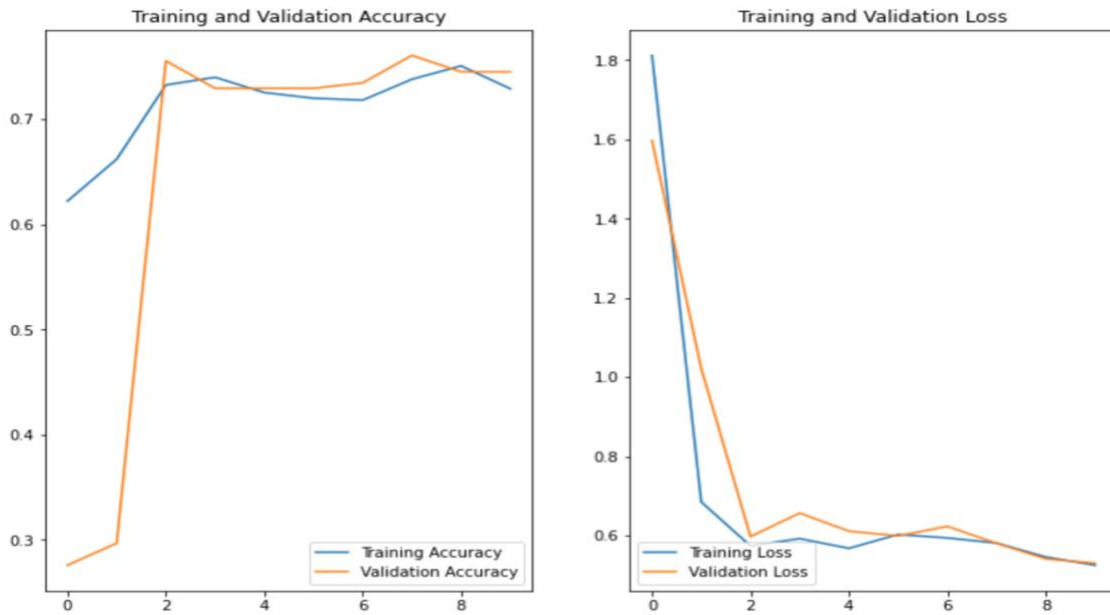

Figure 3.10: Performance of Model 3: Accuracy (Left)l and Loss (Right)

Since transfer learning was used to retrain the model that was pre-trained on radiologic images, we were able to use a smaller training set than typical with this novel approach.

3.3 Radiologist's evaluation aided by the Ultrasound Model

Our model was used to aid the radiologists at SCVMC in making diagnostic decisions on the SCVMC dataset. When the model results were incorporated into the sonographic images, the radiologist's performance showed meaningful improvements. In particular, the accuracy for breast cancer detection achieved by the radiologist was initially 71.7%, and with the aid of the model prediction results, the recall diagnosis was done accordingly and the radiologist's performance improved and resulted in a detection accuracy of 89.0%.

Chapter 4

Conclusion

4.1 Discussion

The three models presented in this study are all built from scratch and differentiates this study from its peers as majority of the approaches use transfer learning approach based on existing pretrained models.

For the breast cancer detection with ultrasound images dataset, the 300 masses collected at SCVMC, 194 were benign (64.7%) and 106 were malignant (35.3%). On this test set, the CNN model-1 achieved an accuracy of 94.89%. The model demonstrated high accuracy on our test dataset. On a small dataset, high diagnostic performance of the CNN model is accomplished in this study.

Our model for the detection of intracranial aneurysms within MRA images achieved an accuracy of 76.04 on our SCVMC test dataset. This preliminary study aimed to demonstrate how to use one medical dataset to pretrain a model and then transfer it for another medical application by retraining and fine-tuning the model on the target dataset. The preliminary results have shown the potential of this new approach being applied to broader medical applications.

It is critical to make accurate diagnosis of breast cancer and brain aneurysm at early stage so that appropriate treatment can be instituted. In this thesis work, three CNN models were built, trained and validated for this purpose. In particular, model-1 for the detection of breast cancer achieved a high accuracy of 94.89% and has shown great potential in assisting radiologists in their decision making.

4.2 Future Work

This thesis work focused on image classification for breast cancer and brain aneurysm detections. The next steps are working to further improve both models' accuracy and then to target the localization along with the prediction. This addition can better assist radiologists in identifying particular lesions or abnormal regions. As part of the future work, various prognosis tasks may also be targeted.

Bibliography

[1] L. A. Torre, F. Bray, R. L. Siegel, J. Ferlay, J. Lortet-Tieulent, and A. Jemal, "Global cancer statistics, 2012," *CA: a cancer journal for clinicians*, vol. 65, no. 2, pp. 87–108, 2015.

[2] K. Munir, H. Elahi, A. Ayub, F. Frezza, and A. Rizzi, "Cancer diagnosis using deep learning: a bibliographic review," *Cancers*, vol. 11, no. 9, p. 1235, 2019.

[3] cdc.gov, "Breast cancer statistics," Jun 2020.

[4] J. H. Youk and E.-K. Kim, "Supplementary screening sonography in mammographically dense breast: pros and cons," *Korean journal of radiology*, vol. 11, no. 6, p. 589, 2010.

[5] W. Al-Dhabyani, M. Gomaa, H. Khaled, and A. Fahmy, "Dataset of breast ultrasound images," *Data in brief*, vol. 28, p. 104863, 2020.

[6] B. A. Foundation, "Statistics and facts," Nov 2020.

[7] E. S. Nussbaum, L. Sebring, and D. Y. Wen, "Intracranial aneurysm rupture presenting as delayed stroke secondary to cerebral vasospasm," *Stroke*, vol. 28, no. 10, pp. 2078–2080, 1997.

[8] M. H. Vlak, A. Algra, R. Brandenburg, and G. J. Rinkel, "Prevalence of unruptured intracranial aneurysms, with emphasis on sex, age, comorbidity, country, and time period: a systematic review and meta-analysis," *The Lancet Neurology*, vol. 10, no. 7, pp. 626–636, 2011.

[9] J. Van Gijn, R. S. Kerr, and G. J. Rinkel, "Subarachnoid haemorrhage," *The Lancet*, vol. 369, no. 9558, pp. 306– 318, 2007.

[10] F. Rincon, R. H. Rossenwasser, and A. Dumont, "The epidemiology of admissions of nontraumatic subarachnoid hemorrhage in the united states," *Neurosurgery*, vol. 73, no. 2, pp. 217–223, 2013.

[11] E. S. Connolly Jr, A. A. Rabinstein, J. R. Carhuapoma, C. P. Derdeyn, J. Dion, R. T. Higashida, B. L. Hoh, C. J. Kirkness, A. M. Naidech, C. S. Ogilvy, *et al.*, "Guidelines for the management of aneurysmal subarachnoid hemorrhage: a guideline for healthcare

professionals from the american heart association/american stroke association," *Stroke*, vol. 43, no. 6, pp. 1711–1737, 2012.

[12] Z. Taufique, T. May, E. Meyers, C. Falo, S. A. Mayer, S. Agarwal, S. Park, E. S. Connolly, J. Claassen, and J. M. Schmidt, "Predictors of poor quality of life 1 year after subarachnoid hemorrhage," *Neurosurgery*, vol. 78, no. 2, pp. 256–264, 2016.

[13] A. Ois, E. Vivas, G. Figueras-Aguirre, L. Guimaraens, E. Cuadrado-Godia, C. Avellaneda, B. Bertran-Recasens, A. Rodr´ıguez-Campello, M.-P. Gracia, G. Villalba, *et al.*, "Misdiagnosis worsens prognosis in subarachnoid hemorrhage with good hunt and hess score," *Stroke*, vol. 50, no. 11, pp. 3072–3076, 2019.

[14] R. Agid, T. Andersson, H. Almqvist, R. Willinsky, S.-K. Lee, R. Farb, M. Soderman,¨ *et al.*, "Negative ct angiography findings in patients with spontaneous subarachnoid hemorrhage: when is digital subtraction angiography still needed?," *American journal of neuroradiology*, vol. 31, no. 4, pp. 696–705, 2010.

[15] T. M. Mitchell *et al.*, "Machine learning," 1997.

[16] J. S. Gero and F. Sudweeks, *Artificial Intelligence in Design '96*. Springer Science & Business Media, 2012.

[17] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep learning*, vol. 1. MIT press Cambridge, 2016.

[18] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, pp. 1097–1105, 2012.

[19] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[20] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1–9, 2015.

[21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[22] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700–4708, 2017.

[23] S. Han, H.-K. Kang, J.-Y. Jeong, M.-H. Park, W. Kim, W.-C. Bang, and Y.-K. Seong, "A deep learning framework for supporting the classification of breast lesions in ultrasound images," *Physics in Medicine & Biology*, vol. 62, no. 19, p. 7714, 2017.

[24] V. L. Mango, M. Sun, R. T. Wynn, and R. Ha, "Should we ignore, follow, or biopsy? impact of artificial intelligence decision support on breast ultrasound lesion assessment," *American Journal of Roentgenology*, vol. 214, no. 6, pp. 1445–1452, 2020.

[25] Q. Zhang, Y. Xiao, W. Dai, J. Suo, C. Wang, J. Shi, and H. Zheng, "Deep learning based classification of breast tumors with shear-wave elastography," *Ultrasonics*, vol. 72, pp. 150–157, 2016.

[26] M. M. S. Bhurwani, A. R. Podgorsak, A. R. Chandra, R. A. Rava, K. V. Snyder, E. I. Levy, J. M. Davies, A. H. Siddiqui, and C. N. Ionita, "Feasibility study of deep neural networks to classify intracranial aneurysms using angiographic parametric imaging," in *Medical Imaging 2019: Computer-Aided Diagnosis*, vol. 10950, p. 109502A, International Society for Optics and Photonics, 2019.

[27] R. S. Lee, F. Gimenez, A. Hoogi, K. K. Miyake, M. Gorovoy, and D. L. Rubin, "A curated mammography data set for use in computer-aided detection and diagnosis research," *Scientific data*, vol. 4, no. 1, pp. 1–9, 2017.