

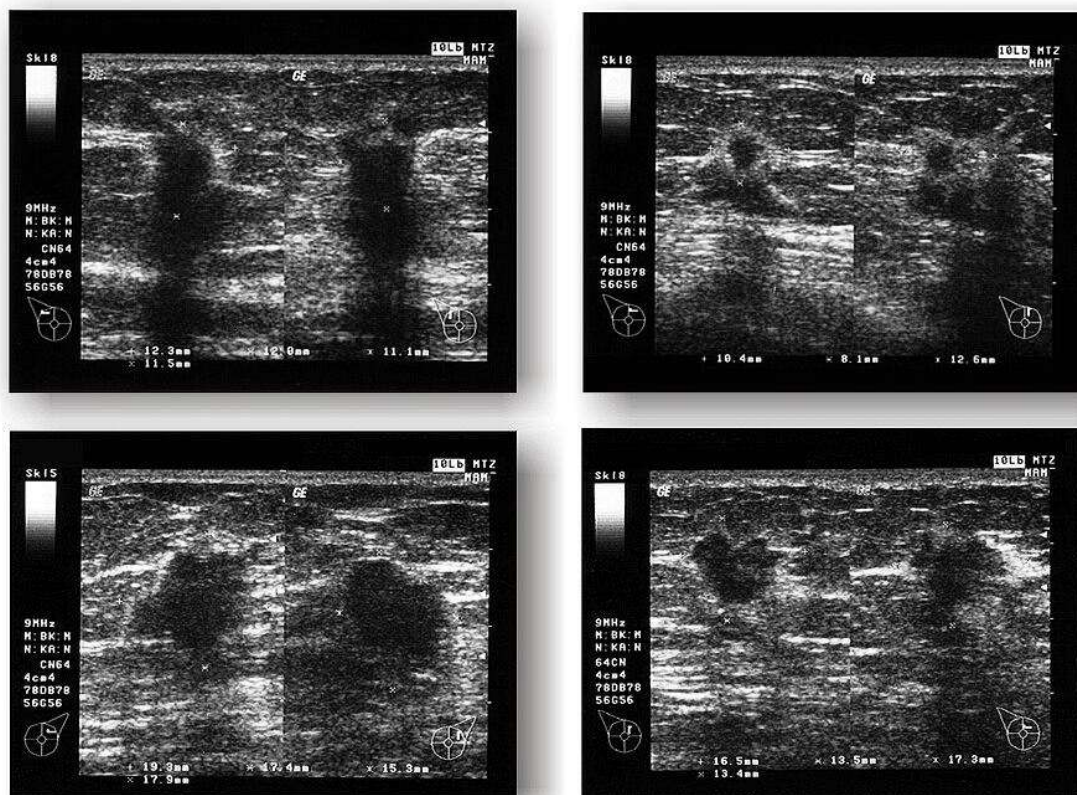
# RESEARCH PROPOSAL

## BREAST CANCER ULTRASOUND IMAGE ANALYSIS

---

Jack McMorro and Alejandra Mejia

---



## Introduction

In this project, we will utilize a breast ultrasound images dataset which contains 780 images in order to develop a classification model that has the capability to predict if the ultrasound shows a normal, benign or malignant formation on the breasts.

---

---

## Why breast cancer ultrasounds image analysis?

Breast cancer is a significant health concern worldwide, and early detection is crucial for improving patient outcomes. Deep learning models have emerged as a promising tool for breast cancer diagnosis and segmentation. These models for breast cancer ultrasound image analysis, with a focus on segmentation, hold significant potential in improving the accuracy and efficiency of breast cancer diagnosis and treatment enhancing the capabilities of healthcare professionals, aiding in early detection and precise localization of breast cancer lesions.

## Dataset

The dataset we will use has been retrieved from [Kaggle](#) and contains of 780 images with an average image size of 500\*500 pixels. The data contains ultrasound images among women in ages 25 and 74 years old. The images are categorized in three classes: normal, benign and malignant.

## Deep Learning Techniques and Metrics

The deep learning techniques that we will be implementing will be essential to the image classification task. We will primarily focus on convolutional neural networks, as they tend to perform well on image classification techniques. This will require us to perform different image augmentation techniques as well as experiment with different architectures of the neural network.

We will also like to experiment with image segmentation. Ideally, this would require a segment model that would segment the image to highlight the classified malignant tumor. This will require research on our part to ensure the segmentation is done properly and accurately.

## Reference Materials

Data Source:

<https://www.kaggle.com/datasets/aryashah2k/breast-ultrasound-images-dataset>

---

[https://pytorch.org/tutorials/beginner/blitz/cifar10\\_tutorial.html](https://pytorch.org/tutorials/beginner/blitz/cifar10_tutorial.html)

<https://glassboxmedicine.com/2020/05/29/grad-cam-visual-explanations-from-deep-networks/>

## Preliminary Work Schedule

Activities/ Tasks		Timeline ( Nov 2023- Dec 2023) Starting Nov 2			
		W1 11/2- 11/9	W2 11/10- 11/17	W3 11/18 11/25	W4 11/26- 12/03
<b>Act 0. Inception</b>					
Task 0.1	Github Set up				
Task 0.2	Dataset selection				
<b>Act1. Proposal and data cleaning</b>					
Task 1.1	Short proposal document		D1		
Task 1.2	Data preprocessing (if needed)				
<b>Act.2 Training and Model Implementation</b>					
Task 2.1	Split and train dataset				
Task 2.2	Implement MLP and CNN				
Task 2.3	Image segmentation				
Task 2.4	Webpage layout				
<b>Act 3 Data viz and final reports</b>					
Task 3.1	Implementation website				
Task 3.2	Start individual reports				

Task 3.3	Layout Presentation Outline				
Task 3.4	Continue Work on Individual Reports				
<b>Act 4. Project wrapup</b>					
Task 4.1	Fully Make and Rehearse Presentation				
Task 4.2	Finish Individual Reports				
Task 4.3	Fully Prepare Github for submission				
Task 4.4	Give Presentation on 12/5				

Deliverable Schedule		Timeline ( Nov 2023- Dec 2023) Starting Nov 2			
		W1 11/2- 11/9	W2 11/10- 11/17	W3 11/18 11/25	W4 11/26- 12/03
<b>Act 0. Inception</b>					
D1	Short Proposal document		D1		
D2	Individual report submission				D2
D3	Group report and resources submissions				D3
D4	Group presentation				D4