I. Abstract

Cardiovascular diseases are the leading cause of death worldwide. It has been shown that a contributing factor in causing cardiovascular diseases is cardiac adipose tissue, or the layer of visceral fat around the heart. In SIUE's Biomedical Imaging Research Lab, methods are being designed to locate, quantify, and map this fat surrounding the heart using ultrasound imaging, advanced signal processing, and machine learning. A key first step to this process of interpreting ultrasound scans is to identify the image orientation, or view, of each scan. Convolutional neural networks, which are a class of machine learning algorithms, are very effective at recognizing patterns in, and classifying, image data. In this study, a convolutional neural network model will be created to successfully classify ultrasound images corresponding to their specific view. After creating and analyzing the data from the model, more iterations of this model can be created to reduce bias and increase the robustness of the convolutional neural network.

II. Introduction and Significance

The leading causes of death worldwide are cardiovascular diseases. Some of the more well-known of these diseases, such as heart attacks and strokes, can be caused by fat deposits in blood vessels that block blood flow to the heart or brain [1]. However, the presence of an unhealthy amount of cardiac adipose tissue (CAT), which is the layer of visceral fat surrounding the heart, has also been shown to be a major factor in cardiovascular diseases, including coronary artery disease [2].

While MRI and CT scans can both be used to image CAT with comparatively high resolutions, these methods are both expensive and immobile. Ultrasound is another imaging option that is less expensive and more mobile, but has less resolution, which makes identifying and quantifying CAT difficult. However, ultrasound scans contain more data than just images; much like how radar or sonar systems work, there is frequency information contained within the ultrasonic signals. For the past three years, I have worked with [redacted] in SIUE's Biomedical Imaging Research Lab, where we have used specialized software along with machine learning algorithms and signal processing techniques to develop methods with an aim to quantify the amount of CAT in ultrasound images of patients' hearts, in hopes to develop a process to use the ultrasound signal's frequency data to map the epicardial fat accurately using only the ultrasound scans.

As mentioned earlier, machine learning has become a major key in the research done in the Biomedical Imaging Research Lab. In basic terms, machine learning is "a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy" [3]. Both former and current graduate and undergraduate students in the research lab have contributed to the process of identifying and validating the quantity of CAT in ultrasound images using machine learning. Each of the following steps in this process uses its own 'neural network', or a series of machine learning algorithms:

1) Determine the view of the ultrasound image (Figure 1, Part 1).
2) Find the contours of epicardial fat in the ultrasound image (Figure 1, Part 2).
3) Classify regions in this contour as 'fat' or 'non-fat' regions (Figure 2, Part 3).
4) Map the regions of epicardial fat as determined (Figure 2, Part 4).

This proposal will focus on implementing the first step of this process. In previous parts of this ongoing study in the Biomedical Imaging Research Lab, only 5-10 ultrasound images were collected per patient. However, this number is currently being increased to 24 total images (Table 1), each representing a different image cut plane, or 'view' of the heart. Figure 1, Part 1 also shows a 3-D model of the heart with planes intersecting it (labeled 1-5) that represent some of the different views that this project will focus on differentiating. It is our belief that a new neural network can be created to identify these views prior to identifying the epicardial fat contours in the ultrasound image.

The type of machine learning algorithm that will be used to classify the ultrasound views is called a convolutional neural network (CNN). For many reasons, CNNs are widely used in image classification, making them a perfect candidate for classifying ultrasound views [4]. In a CNN, image data are input into a system of convolutional operations in a series of layers. These algorithms will produce an output label, corresponding to a list of the given ultrasound views shown in Table 1. The design of this network is sometimes also called 'deep learning', and mimics the function of the human brain's optical system [4]. The methodology behind creating a deep learning 'model' is to utilize datasets (in this case, a database of cardiac ultrasound images) to train a computer to recognize each of the views. After this model is created, it can then be used to classify future ultrasound images based on their orientation.

III. Literature Review

Echocardiographic view classification via deep learning is a topic that is continuing to grow for several reasons. First, recognizing the ultrasound image orientation is the "essential first step toward comprehensive computer-assisted echocardiographic interpretation" [5]. In conjunction with this, "deep learning, specifically using convolutional neural networks (CNNs), is a cutting-edge machine learning technique that has proven "unreasonably" successful at learning patterns in images and has shown great promise helping experts with image-based diagnosis in radiology" [5]. Based on these remarks, a CNN is the ideal choice for the process of classifying cardiac ultrasound images based on their view.

With deep learning classification, there are several issues to look out for, which mainly stem from a lack of data. Machine learning is similar to human learning in the way that there is a positive correlation between learning successfully and the number of 'examples' worked. In the case of echocardiography and view classification, accuracy "greatly depends on large training sample datasets" [6]. Smaller datasets not only affect accuracy, but can cause the machine learning model to adopt biases in its measurements and classification [7]. Another issue that a successful research project using machine learning will look out for is the "black box" effect; machine learning models and how they work can be "obscure" at a first glance. Simply feeding data into a CNN and getting results from it is not sufficient in research. It is important to explore the 'why' and 'how' in terms of how a machine learning model reaches its conclusions [5].

There are a few ways that these issues can be counteracted. The most straightforward way is to acquire more images. The research done in the Biomedical Imaging Research Lab uses images solely from subjects that are imaged in the lab by professional sonographers on a specialized ultrasound system that allows for the capturing of raw radiofrequency data along with the ultrasound images. In order to supplement this data, a possibility that other research methods have utilized is to use online ultrasound image databases [6], [8]. However, since the images from these online datasets may not be acquired using the same machine/processes as in our study, there are other options that could prove to be more viable, such as altering the original architecture of the CNN to make the model more robust, or assigning weights based on analyzing the data from the model [8]. These options are explored further in the Data Analysis section of this proposal.

IV. Objectives, and Operating Hypothesis

Objectives

1) Divide a dataset of ultrasound images into data to train the CNN model (called training data), and data to feed into the model after it is trained in order to test the model's accuracy (called testing data).
2) Create a working CNN to classify these cardiac ultrasound images.
3) Alter the dataset and/or CNN to improve the model's classification results.

Operating Hypothesis

A working convolutional neural network (CNN) can be created to classify cardiac ultrasound images with an accuracy of 85%, and will improve the detection of epicardial fat contours in the future.

V. Materials, Procedure, and Timeline

Materials Required and Uses

- Mindray Zonare Z3 Ultrasound Machine – used for capturing ultrasound images and corresponding frequency data.
- Cardiac P4-1c Ultrasound Probe – probe used for cardiac ultrasound imaging.
- 2 EVGA RTX 2070 Super Graphics Processing Units (GPUs) – used for computing power to greatly accelerate machine learning processes in computers.
- EVGA NVLink SLI Bridge – links the two GPUs above to combine their processing power for increased efficiency.

Target Population and Sampling Methods

Due to an increased amount of CAT being found in females as opposed to males, and an increase in CAT also being correlated to higher body weights, this study has focused on a target population of women between the ages of 18 and 40, with a body-mass index between 30 and 39.99 kg/m^2. Having a homogeneous sample of imaging subjects helps have a controlled variable in this research project, and limits the number of independent variables. Imaging of these subjects is done by professional sonographers with a Mindray Zonare Z3 Ultrasound Machine. 24 total ultrasound images are captured per subject, as listed in Table 1.

Project Design and Methods for Data Collection

As mentioned, a CNN will be used to classify the ultrasound images as one of the different views listed in Table 1, meaning that the creation of the CNN model will be the main task for this research project. Prior research done in the Biomedical Imaging Research Lab has used the software library 'TensorFlow', which is an open-source framework for machine learning algorithms. This software library will be used to create the initial framework of the CNN model.

However, prior to designing the model, data must be prepared for its creation. Ultrasound scans from our Mindray Zonare Z3 Ultrasound Machine come in the form of "IQ" data, which are files that contain the sample values of the ultrasound's radiofrequency signals that our system allows us to obtain. The first step in the data preparation process is to convert this IQ data from each ultrasound scan into an image file, since the CNN's input must be in the form of an image. The images can be represented in two main formats. The first format, shown in Figure 3, represents what a sonographer would see on the ultrasound machine, and shows actual size and scale of landmarks in the ultrasound scan. Because of this, this format gives a better understanding of quantity and location of CAT and other landmarks in the heart (such as the ventricles, myocardium, and atria). However, the second format, shown in Figure 4, represents the 'signal domain' or 'IQ domain'. This format shows how the signal is received by the ultrasound transducer. At this point, it is uncertain which image format is 'better' regarding how the CNN will classify views, so each image format will be tested to see which, if either, obtains better results.

In terms of data being given to the model, the total number of images needs to be split into training data and testing data. Training data is used to train the model; in essence, images are shown to the model, along with the image's 'label' (the label, in this case, is the image's view). From this information, the model attempts to recognize patterns corresponding to each image to correctly correlate the images with their respective label.

After the training process is completed, the testing data is used; images are now presented as input to the model without their labels, and the CNN attempts to classify the image with a certain label based on what it learned from the training data. As mentioned in the literature review, more training data will tend to increase the efficacy of the CNN. Because of this, roughly 80% of the total images will be assigned to be training data, and the remaining 20% will be testing data.

After preparing the input data for the model, the CNN can now be created by performing three processes. Each image, as shown in Figures 1 and 2, are greyscale in nature. This means that the images can be modeled as arrays, where each value in the array corresponds to the brightness of each pixel in the image. The first process of a CNN is convolution (the namesake of the CNN). This process uses convolution, a mathematical operation involving matrix multiplication and addition, in order to assign weights to important regions of the input image. The next step is called pooling, where the CNN decreases the size of the convolved image to only include its most important features. Finally, the CNN compares these features to those that it extracted from the training data, and attempts to classify the input images based on this. [9]

Data Analysis

Machine learning models traditionally use several different metrics for data analysis: accuracy, precision, sensitivity, and specificity. In the case of this research project, accuracy is calculated by comparing the number of correctly identified views to the total number of images that were given to the model to classify.

$$Accuracy = \frac{Correctly\ Classified\ Images}{Total\ Number\ of\ Images}$$

Precision can be measured for each different view (labeled in the equation by 'View #N', representing an arbitrary view), by comparing the number of correct classifications of that view to the number of total classifications of that view.

$$Precision = \frac{Corectly\ Classified\ Images\ for\ View\ \#N}{Total\ Images\ Classified\ as\ View\ \#N}$$

Sensitivity measures the number of correctly classified images for a certain view, compared to the number of images of that view that were fed into the model.

$$Sensitivity = \frac{Correctly\ Classified\ Images\ for\ View\ \#N}{Total\ Images\ of\ View\ \#N\ Given\ to\ Model}$$

Specificity, on the other hand, compares the number of images that were classified as not being a certain view, compared to the number of images given to the model that were not that view. [10]

$$Specificity = \frac{Images\ Correctly\ Classified\ as\ NOT\ View\ \#N}{Total\ Images\ that\ are\ NOT\ View\ \#N}$$

As mentioned in the literature review, looking at these metrics can help decipher the '"black box" nature of deep learning models [5]. Determining whether the model has a tendency to classify false-positives or false-negatives for certain views, and looking at these specific images, can give helpful insight into which aspects of the model need to be altered.

Iterations and Design Improvements

After gaining insights from analyzing data, useful information can be gathered which can assist in improving the model. One of the most common ways to improve a machine learning model is to assign weights to the different data labels [5], [6]. For instance, if a certain view is consistently being classified as other views, a higher 'weight' can be assigned to that view, which, ideally, will increase the number of correctly classified images for that specific view. In contrast, if a view is consistently being overclassified, it can be assigned a lower weight, meaning that, ideally, less images will be incorrectly given the label corresponding to that view.

Research Timeline

May 2023 – August 2023: During this time period, ultrasound imaging of the target population listed above will be conducted in order to gather as much data as possible. In addition to this, the conversion of collected ultrasound data from the "IQ" format to image format will be completed.

August 2023 – October 2023: A first-pass CNN will be created using the TensorFlow software library. From this CNN, data including accuracy, precision, sensitivity, and specificity can be collected and analyzed to determine ways in which to improve the CNN's performance.

October 2023 – December 2023: An improved CNN will be created using information from the data metrics mentioned above.

December 2023 – February 2023: Perform final data analysis and alterations to the CNN model. Also, begin writing a research paper/summary in preparation for the URCA Research Symposium.

February 2023 – March 2023: Compose the project design poster for the URCA Research Symposium.

April 2023: Finalize all aspects of the project and present the project and poster at the URCA Research Symposium.

VI. References

[1] "Cardiovascular diseases (CVDs)." https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) (accessed Mar. 07, 2023).

[2] E. E. Ha and R. C. Bauer, "Emerging Roles for Adipose Tissue in Cardiovascular Disease," *Arterioscler. Thromb. Vasc. Biol.*, vol. 38, no. 8, pp. e137–e144, Aug. 2018, doi: 10.1161/ATVBAHA.118.311421.

[3] "What is Machine Learning? | IBM." https://www.ibm.com/topics/machine-learning (accessed Mar. 11, 2023).

[4] T. Wood, "Convolutional Neural Network," *DeepAI*, May 17, 2019. https://deepai.org/machine-learning-glossary-and-terms/convolutional-neural-network (accessed Mar. 09, 2023).

[5] A. Madani, R. Arnaout, M. Mofrad, and R. Arnaout, "Fast and accurate view classification of echocardiograms using deep learning," *Npj Digit. Med.*, vol. 1, no. 1, Art. no. 1, Mar. 2018, doi: 10.1038/s41746-017-0013-1.

[6] S. Liu *et al.*, "Deep Learning in Medical Ultrasound Analysis: A Review," *Engineering*, vol. 5, no. 2, pp. 261–275, Apr. 2019, doi: 10.1016/j.eng.2018.11.020.

[7] J. Zhang *et al.*, "Fully Automated Echocardiogram Interpretation in Clinical Practice," *Circulation*, vol. 138, no. 16, pp. 1623–1635, Oct. 2018, doi: 10.1161/CIRCULATIONAHA.118.034338.

[8] J. P. Howard *et al.*, "Improving ultrasound video classification: an evaluation of novel deep learning methods in echocardiography," *J. Med. Artif. Intell.*, vol. 3, p. 4, Mar. 2020, doi: 10.21037/jmai.2019.10.03.

[9] J. C. Martinez, "Introduction to Convolutional Neural Networks CNNs." https://aigents.co/data-science-blog/publication/introduction-to-convolutional-neural-networks-cnns (accessed Mar. 12, 2023).

[10] M. Teh, "How to Interpret the Confusion Matrix: Accuracy, Sensitivity, Specificity, Precision & F1 Score," *UpLevel*, Jul. 30, 2020. https://projects.uplevel.work/insights/confusion-matrix-accuracy-sensitivity-specificity-precision-f1-score-how-to-interpret (accessed Mar. 11, 2023).

VII. Budget Justification

Item	Cost
Crucial Ballistix 3200 MHz DDR4 DRAM (16 GB)	$105
Samsung 990 Pro 2 TB SSD	$250
URCA Symposium Poster Creation & Lamination	$25
TOTAL	**$380**

Many of the materials mentioned earlier in this proposal, such as the ultrasound machine, GPUs, and SLI bridge, have already been purchased through grants from prior research in the Biomedical Imaging Research Lab. However, computing power is of vital importance to the efficiency of the deep learning process. Since GPUs account for the bulk of this speed (and the bulk of the cost), it is fortunate that we already have access to 2 EVGA RTX 2070 Super Graphics Processing Units. However, there are other computer components that also provide computing power, such as RAM (random access memory), and internal hard drives. For this reason, new RAM in the form of 16 GB of Crucial Ballistix 3200 MHz DDR4 DRAM is needed to allocate more memory to the deep learning process. In addition, a 2 TB solid state drive, specifically a Samsung 990 Pro 2 TB SSD, will assist in speeding up this process as well.

Appendix

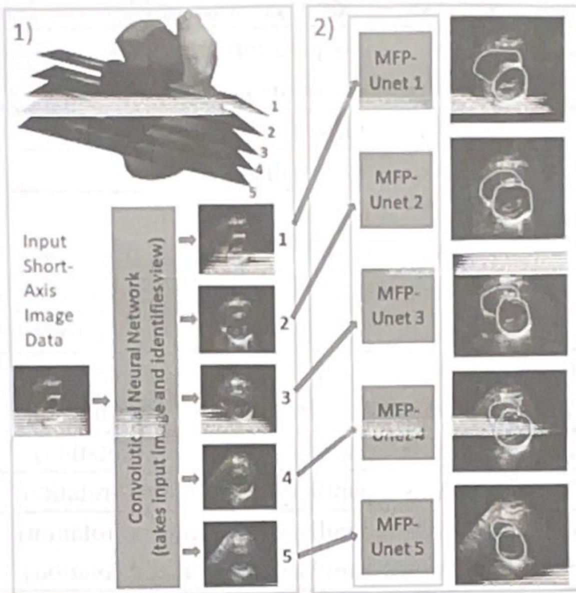

*Figure 1 - **Part 1** - Identifying ultrasound view; **Part 2** - Identify epicardial fat contours*

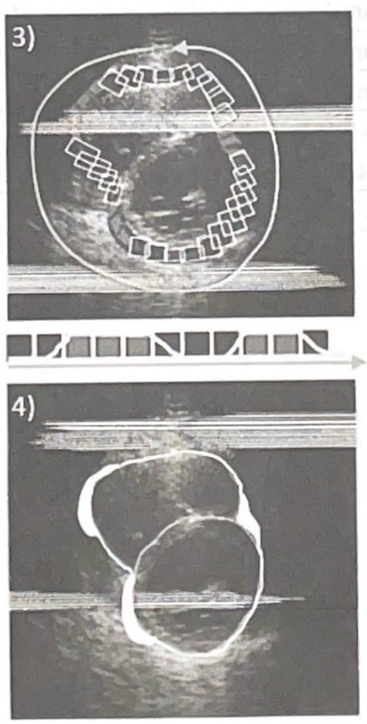

*Figure 2 - **Part 3** - Identify 'fat' & 'non-fat' regions; **Part 4** - Map epicardial fat*

Ultrasound Image Views	
1.	Long Axis – traditional
1.	Long Axis – slide angle inferior (down a rib from 0)
2.	Long Axis – low parasternal, LV/apex
3.	Short Axis – ascending aorta
4.	Short Axis – aortic valve
5.	Short Axis – RCA ostia
6.	Short Axis – left main
7.	Short Axis – mitral valve level (normal rotation)
8.	Short Axis – mitral valve level (medial rotation)
9.	Short Axis – mitral valve level (lateral rotation)
10.	Short Axis – chordae level (normal rotation)
11.	Short Axis – chordae level (medial rotation)
12.	Short Axis – chordae level (lateral rotation)
13.	Short Axis – papillary level (normal rotation)
14.	Short Axis – papillary level (medial rotation)
15.	Short Axis – papillary level (lateral rotation)
16.	Short Axis – apex level superior (normal rotation)
17.	Short Axis – apex level superior (medial rotation)
18.	Short Axis – apex level superior (lateral rotation)
19.	Short Axis – apex level (normal rotation)
20.	Short Axis – apex level (medial rotation)
21.	Short Axis – apex level (lateral rotation)
22.	RV-focused apical view
23.	Apical 4-chamber with medial rotation

Table 1 - List of all 24 different ultrasound image views

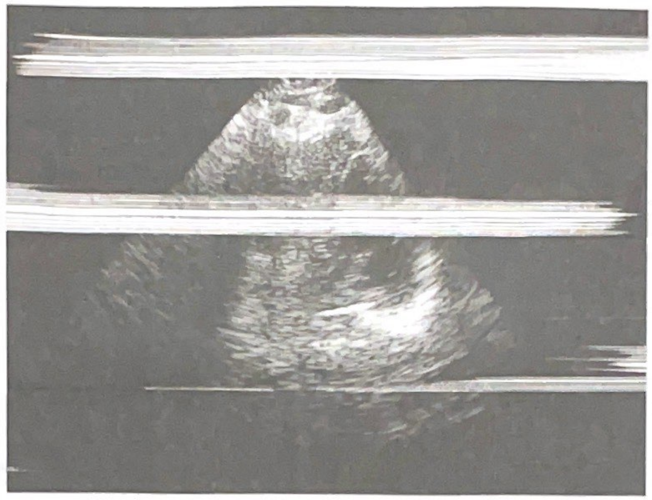

Figure 3 *- First image format, showing size and scale of different cardiac landmarks*

Figure 4 *- Second image format, showing linearly how the ultrasound transducer receives the radio-frequency signals*