



Ain Shams University

Faculty of Computer Science and Information Sciences

Scientific Computing Department

Echo-Based Cardiac Function Assessment

This documentation submitted as required for the degree of bachelor's in computer and Information Sciences.

By

Abanoub Gamal Fekry	[Scientific Computing Department]
Kerolos Nabil Mounir	[Scientific Computing Department]
Yassa Kamille AbdEllatef	[Scientific Computing Department]
Kerolos Helal Zaki Soltan	[Scientific Computing Department]
Ganna Muhamed Ahmed	[Scientific Computing Department]
Kerolos Waheed Kamel	[Scientific Computing Department]

Under Supervision of

Dr. Manal Mohsen Tantawi

Associate Professor at Scientific Computing Department,
Faculty of Computer and Information Sciences,
Ain Shams University.

TA. Radwa Reda Hossieny

Teaching Assistant of Scientific Computing Department,
Faculty of Computer and Information Sciences,
Ain Shams University.

**Cairo
July 2024**



Ain Shams University

Faculty of Computer Science and Information Sciences

Scientific Computing Department

Echo-Based Cardiac Function Assessment

This documentation submitted as required for the degree of bachelor's in computer and Information Sciences.

By

Abanoub Gamal Fekry	[Scientific Computing Department]
Kerolos Nabil Mounir	[Scientific Computing Department]
Yassa Kamille AbdEllatef	[Scientific Computing Department]
Kerolos Helal Zaki Soltan	[Scientific Computing Department]
Ganna Muhamed Ahmed	[Scientific Computing Department]
Kerolos Waheed Kamel	[Scientific Computing Department]

Under Supervision of

Dr. Manal Mohsen Tantawi

Associate Professor at Scientific Computing Department,
Faculty of Computer and Information Sciences,
Ain Shams University.

TA. Radwa Reda Hossieny

Teaching Assistant of Scientific Computing Department,
Faculty of Computer and Information Sciences,
Ain Shams University.

**Cairo
July 2024**

Acknowledgements

First and foremost, we thank the Merciful Allah, who has given everything to prepare and present this work. He gave me strength in my times of weakness, hope in my times of despair, and patience in our troubles. He gave me inspirational signs to overcome all the difficulties I faced

We're profoundly grateful to our parents and families for their unwavering support and encouragement throughout our years of study. Their love and sacrifices have been the foundation of our success, and we hope to one day repay their kindness and dedication.

Our deepest appreciation goes to our esteemed supervisor, Dr. Manal Mohsen Tantawi, whose unwavering support and guidance have been invaluable. Their mentorship, insightful advice, and boundless dedication were instrumental in overcoming the challenges we encountered. Without their expertise and encouragement, navigating through this project would have been exceedingly daunting.

Special thanks are extended to T.A. Radwa Reda for their invaluable assistance and unwavering support, which greatly enriched our project's development. We also extend our sincere appreciation to all individuals who played a part, no matter how big or small, in the success of our endeavor.

Lastly, we acknowledge the indispensable efforts of each member of our team, whose collective dedication and hard work were indispensable to the project's success. Together, we have achieved a milestone worth celebrating, and for that, we are profoundly grateful.

Abstract

Recently, Left Ventricle Ejection Fraction (LVEF) become a crucial medical measurement that quantifies the percentage of blood pumped out of the left ventricle of the heart during each contraction or heartbeat. Precise assessment of Left Ventricular Ejection Fraction (LVEF) holds significant importance in the care of individuals dealing with cardiovascular issues. Additionally, LVEF serves as a valuable indicator for forestalling unfavorable results in patients with conditions such as congestive heart failure, following a heart attack, or subsequent to revascularization procedures. Left ventricular ejection fraction indicates whether a person is normal or has a heart disease. Determining the value of the left ventricular ejection fraction quickly can guide doctors to identify the disease in its early stages and then treat it.

In this work, the proposed system provides an effective approach to determine whether a person is suffering from a heart problem by identifying the ejection fraction of a heartbeat from the echocardiography video and that was achieved using Deep learning models (DL). First, the system performs the video analysis phase of the echocardiogram video and divides it into a number of frames and identifies whether this frame contains the state of the heart contracting or contracting. Then determine the position of the left atrium using the binary mask. In addition to calculating the size of the heart using the parameters of the left atrium of the heart. And last but not least, the system is able to calculate the value of the ejection fraction of the heartbeat based on the size of the heart in systole and diastole.

Finally, it is classified whether a person is normal or sick using a threshold method for the ejection fraction value of the heartbeat. The proposed system achieves effective experimental results as it achieved a mean square error of 1.69 and achieved a dice coefficient similarity of 91.8. The proposed system is represented in usable and accessible application that meets the needs of doctors everywhere and helping them to provide an accurate diagnosis of heart failure to the Patients. The application has the potential to make a positive impact on the diagnosis process.

Table of Contents

Acknowledgements.....	I
Abstract.....	II
Table of Contents.....	III
List of Figures.....	V
List of Tables.....	VI
List of Abbreviations.....	VII
Chapter 1: Introduction.....	2
1.1 Problem Definition.....	2
1.2 Motivation.....	6
1.3 Objective.....	7
1.4 Time Plan.....	8
1.5 Documentation organization.....	8
Chapter 2: Literature Review.....	11
2.1 General Overview.....	11
2.2 Related Studies.....	13
2.3 Competitive analysis.....	16
Chapter 3: System Architecture.....	18
3.1 Overview.....	18
3.2 Preprocessing.....	18
3.2.1 Image Resizing.....	19
3.2.2 Image Orientation.....	20
3.2.3 Grayscale.....	21
3.2.4 Random Flips.....	21
3.2.5 Random Rotations.....	22
3.2.6 Random Exposure.....	22
3.2.7 Random Noise.....	23
3.3 Deep learning Models.....	24
3.4 Ejection Fraction (LVEF) Calculation.....	25
3.5 Classification.....	26
Chapter 4: Implementation Techniques.....	28
4.1 Video Analysis.....	28
4.2 End-Systolic and End-Diastolic Frame Detection.....	29
4.3 Find the Binary Mask of Left Ventricle (LV)	31
4.4 Left Ventricle Landmark Detection.....	33
4.5 Left Ventricle Volume Calculation.....	34
4.6 Left Ventricle Ejection Fraction (LVEF) Calculation.....	35
4.7 Thresholding Classification.....	35

Chapter 5: Experimental Results.....	37
5.1 Datasets.....	37
5.2 Experimental Results.....	38
5.3 User Interface.....	40
Chapter 6: Conclusions and Future work.....	45
6.1 Conclusions.....	45
6.2 Future Work.....	46
References.....	47

List of Figures

Fig.1.1. Cardiac Cycle.....	3
Fig.1.2. Cardiomyopathy causes the heart to lose its ability to pump blood well.....	3
Fig.1.3. Cardiovascular Diseases.....	4
Fig.1.4. Ejection Fraction indicators.....	4
Fig.1.5. Cardiologist performs Echocardiograph on a patient.....	5
Fig.1.6. Relation between crude mortality rate and fractional regression.....	6
Fig.1.7. Mobile Application for assessing cardiac function.....	7
Fig.1.8. Project Time Plan.....	8
Fig 2 .1 Classification Techniques.....	12
Fig.3.1. Traditional System Architecture.....	18
Fig 3.2. Image Resizing Effect.....	20
Fig 3.3. Image Orientation.....	20
Fig 3.4. Image Gray Scaling.....	21
Fig 3.5. Random Flips.....	21
Fig 3.6. Random Rotations.....	22
Fig 3.7. Random Exposure.....	23
Fig 3.8. Removing Random Noise.....	23
Fig.4.1. Proposed System Architecture.....	28
Fig.4.2. converting 3D video to frameset.....	28
Fig.4.3. ResNetAE + BERT Encoders Architecture.....	30
Fig.4.4. U-Net Network Architecture.....	32
Fig.4.5. Left Ventricle Landmarks.....	34
Fig. 5.1 Samples of Echo-Net dataset.....	37
Fig. 5.2 Frames Detection Results.....	39
Fig. 5.3 Starting window.....	40
Fig. 5.4 Main window.....	41
Fig. 5.5 Sign-Up window.....	41
Fig. 5.6 Upload window.....	42
Fig. 5.7 Loading window.....	42
Fig. 5.8 Result window.....	43

List of Tables

Table 2.1 Comparison between the Related Studies.....	16
Table 5.1 Echo-Net dataset label variables.....	38
Table 5.2 Left Ventricle Detection Results.....	39

List of Abbreviations

<u>Abbreviation</u>	<u>Stands for</u>
AI	Artificial Intelligence
ARVC	Arrhythmogenic Right Ventricular Cardiomyopathy
CNN	Convolutional Neural Network
DCM	Dilated Cardiomyopathy
DL	Deep Learning
ED	End Diastolic
ES	End-Systolic
GLS	Global Longitudinal Strain
HCM	Hypertrophic Cardiomyopathy
LVEF	Left Ventricle Ejection Fraction
ML	Machine Learning
RCM	Restrictive Cardiomyopathy
SIFT	Scale Invariant Feature Transform
SVM	Support Vector Machine
WHO	World Health Organization

Chapter 1

Introduction

Chapter 1: Introduction

In this chapter, providing a comprehensive overview will be laid out of the key aspects of the study. Initially, the Issues section will outline the specific issues that the project aims to address, highlighting the challenges faced by clinicians to calculate accurate and reliable assessments of cardiac function. and current limitations of current assistive technologies. Following this, the objectives section will outline the primary goals of the project, and the scope will define the boundaries and focus areas of the research. Additionally, the significance of the study will be discussed, emphasizing the potential impact and benefits of the developed system. Finally, the structure of the document will be presented, offering a roadmap of the subsequent chapters and their content.

1.1 Problem Definition

The heart is the central organ of the circulatory system, responsible for maintaining the flow of blood throughout the body. Its function is crucial for sustaining life, as it ensures that oxygen and nutrients are delivered to tissues and organs while removing carbon dioxide and metabolic waste products. The heart operates through a coordinated cycle of contraction and relaxation. During the contraction phase, known as systole, the heart pumps blood out of its chambers. Conversely, during the relaxation phase, called diastole, the heart chambers fill with blood, preparing for the next contraction [1].

The left ventricle, one of the four chambers of the heart, plays a particularly vital role in this process. It is responsible for pumping oxygenated blood into the systemic circulation, which supplies nearly all the organ systems of the body. The efficient function of the left ventricle is essential for overall health, as it ensures that oxygen-rich blood reaches tissues and organs. If the left ventricle fails to function properly, a condition known as left ventricular failure, it can lead to widespread impairment of other organ systems. This is because the diminished cardiac output results in insufficient blood flow as shown in the Fig1.1, compromising the delivery of oxygen and nutrients, and ultimately affecting the functionality of vital organs.

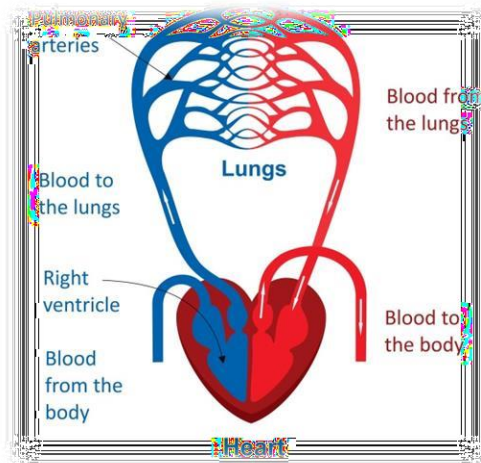


Fig.1.1 Cardiac Cycle

Cardiomyopathy encompasses a spectrum of diseases affecting the heart muscle, resulting in impaired cardiac function, as shown in the Fig1.2. These conditions can lead to severe complications, such as heart failure, arrhythmias (irregular heart rhythms), and sudden cardiac arrest, underscoring the need for timely and accurate diagnosis.

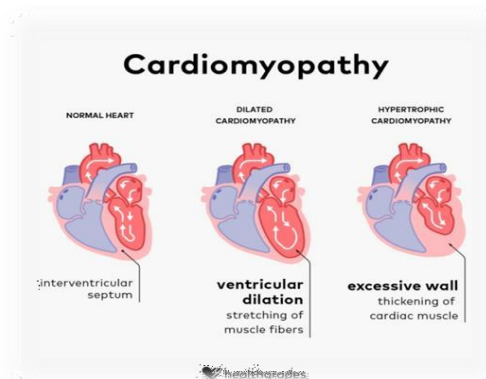


Fig.1.2 Cardiomyopathy causes the heart to lose its ability to pump blood well.

Cardiomyopathy has several types as shown in the Fig1.3. each with unique characteristics. Dilated cardiomyopathy (DCM) involves an enlarged and weakened left ventricle, reducing its ability to pump blood and potentially leading to heart failure and arrhythmias. Hypertrophic cardiomyopathy (HCM) is marked by abnormal thickening of the heart muscle, often the septum, which can obstruct blood flow and increase the risk of sudden cardiac arrest, especially in young athletes. Restrictive cardiomyopathy (RCM) results in stiffened heart muscle, impairing its ability to fill properly with blood and causing symptoms like fatigue and fluid retention. Arrhythmogenic right ventricular cardiomyopathy (ARVC) is a

rare condition where fatty or fibrous tissue replaces the muscle of the right ventricle leading to arrhythmias and a heightened risk of sudden cardiac arrest.

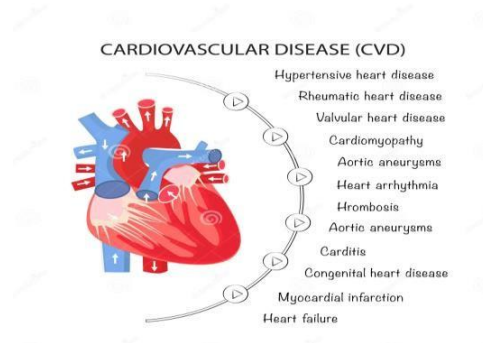


Fig.1.3 Cardiovascular Diseases

Echocardiography is a pivotal diagnostic tool for cardiomyopathy. This non-invasive imaging test utilizes high-frequency sound waves to produce detailed images of the heart, enabling healthcare providers to assess its structure and function. An echocardiogram can reveal crucial information about the size and shape of the heart chambers, the thickness and movement of the heart walls, and the condition of the heart valves [2].

One of the essential measurements derived from an echocardiogram is the Left Ventricle Ejection Fraction (LVEF), which quantifies the percentage of blood the left ventricle ejects with each contraction, as shown in the Fig1.4. A normal LVEF ranges from 55% to 70%. Lower values may indicate heart failure or cardiomyopathy, while higher values can suggest hyperdynamic states or hypertrophic cardiomyopathy.

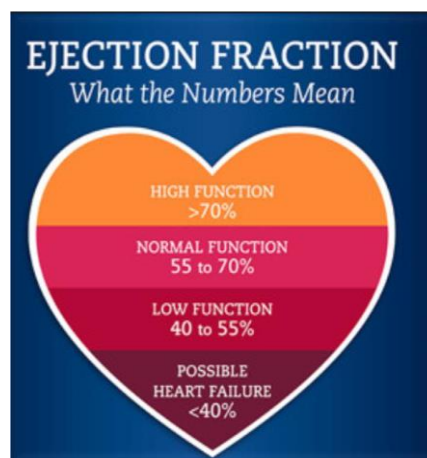


Fig.1.4. Ejection Fraction indicators

Newly graduated and less experienced cardiologists often face challenges in diagnosing cardiomyopathies accurately and swiftly. Interpreting complex echocardiographic data and identifying specific patterns associated with different types of cardiomyopathies can be daunting for those with limited clinical experience. Misinterpretation can lead to incorrect diagnoses, potentially delaying appropriate treatment. In addition to diagnostic accuracy, the speed of decision-making is critical in managing cardiomyopathy. Timely intervention is essential to prevent severe outcomes, but inexperienced cardiologists may take longer to reach a diagnosis due to less familiarity with the clinical presentation and echocardiographic findings.

To address these challenges, several strategies can be employed. Advanced diagnostic tools, such as artificial intelligence (AI) and machine learning algorithms, can assist in analyzing echocardiographic data, identifying patterns, and providing diagnostic suggestions, thereby enhancing both accuracy and speed. Telemedicine allows less experienced cardiologists to consult with seasoned specialists remotely, ensuring more accurate diagnoses and appropriate treatment plans [3].

Continuous education and training are vital for newly graduated cardiologists to improve their diagnostic skills and stay updated with the latest advancements in cardiology. Regular training programs, workshops, and collaborative learning opportunities can help them gain confidence and competence in diagnosing and managing cardiomyopathies. Moreover, involving multidisciplinary teams in patient care, including experienced cardiologists, radiologists, and other specialists, can ensure comprehensive evaluation and management. This team-based approach leverages the expertise of various healthcare providers, enhancing diagnostic accuracy and patient outcomes.

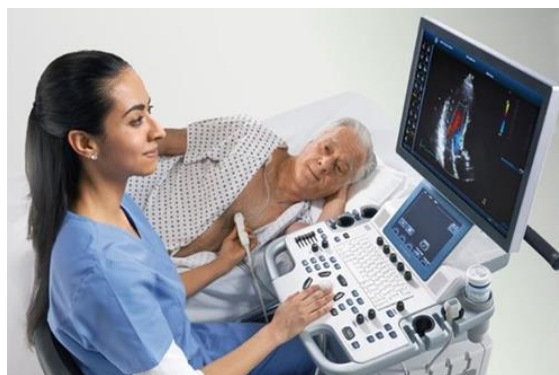


Fig.1.5 Cardiologist performs Echocardiograph on a patient

1.2 Motivation

According to the World Health Organization (WHO) reports, over 500 million individuals worldwide are still impacted by cardiovascular diseases, resulting in approximately 20.5 million fatalities in 2021. This accounted for nearly one-third of all global deaths, showcasing an overall rise from the previously estimated 121 million deaths caused by cardiovascular diseases, as shown in Fig 1.6. Reductions in LVEF portend worse cardiovascular outcomes. Although there is an inverse relationship between LVEF and all-cause mortality rate, this plateaus at an EF of 40% to 45%, above which EF is unrelated to mortality. Ejection fraction assessment in humans varies due to heart rate irregularities and the labor-intensive manual tracing needed for each heartbeat's ventricle size measurement. Guidelines suggest averaging measurements from several heartbeats, but often only one is used, leading to high variability and imprecise results, with inter-observer differences ranging from 7.6% to 13.9%. [4]

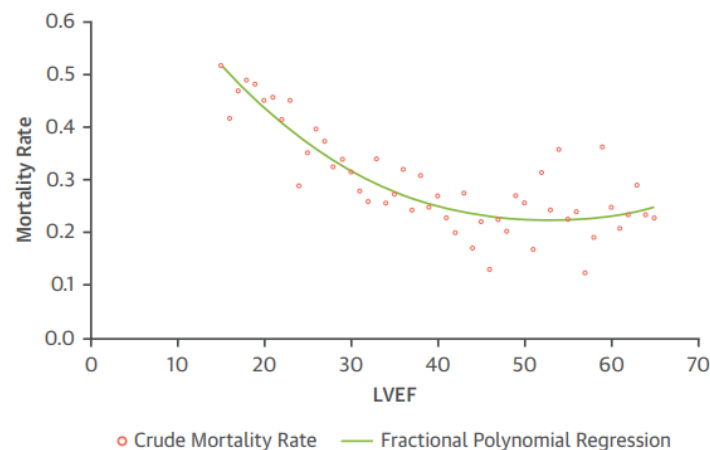


Fig.1.6 Relation between crude mortality rate and fractional Regression

Given the significant impact of cardiovascular diseases, proactive health management is crucial. Regular heart assessments and monitoring can play a vital role in reducing risks and improving health outcomes.

- **Early Detection:** Heart assessments help identify risk factors and early signs of cardiovascular diseases, allowing for timely interventions.
- **Personalized Care:** Heart assessment apps offer personalized health insights and advice on lifestyle modifications, medication adherence, and preventive measures.
- **Accessibility:** Digital health technologies make it easier for individuals to

monitor their heart health from home, especially in areas with limited healthcare access.

- **Patient Empowerment:** Heart assessment apps empower individuals to take charge of their health, encouraging proactive management and adherence to treatment plans.

Understanding the seriousness of cardiovascular diseases and the benefits of early detection and continuous monitoring, it's essential to use modern technology for heart health assessment. By incorporating heart assessment apps into daily routines, individuals can significantly lower their risk of severe cardiovascular conditions.

1.3 Objective

The development of a sophisticated system capable of accurately assessing cardiac function and rapidly estimating the Ejection Fraction of the left ventricle (LVEF) from echocardiogram videos is a significant advancement in cardiovascular diagnostics, as shown in Fig 1.7. This system aims to determine whether a patient's heart is functioning normally or abnormally, thus aiding in early diagnosis and effective management of heart conditions.

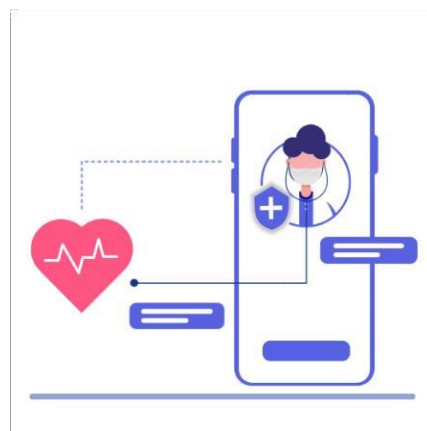


Fig.1.7. Mobile Application for assessing cardiac function

1.4 Time Plan

The deadlines were met in time and the schedule was very flexible and organized giving us the space to complete each task without any delay.

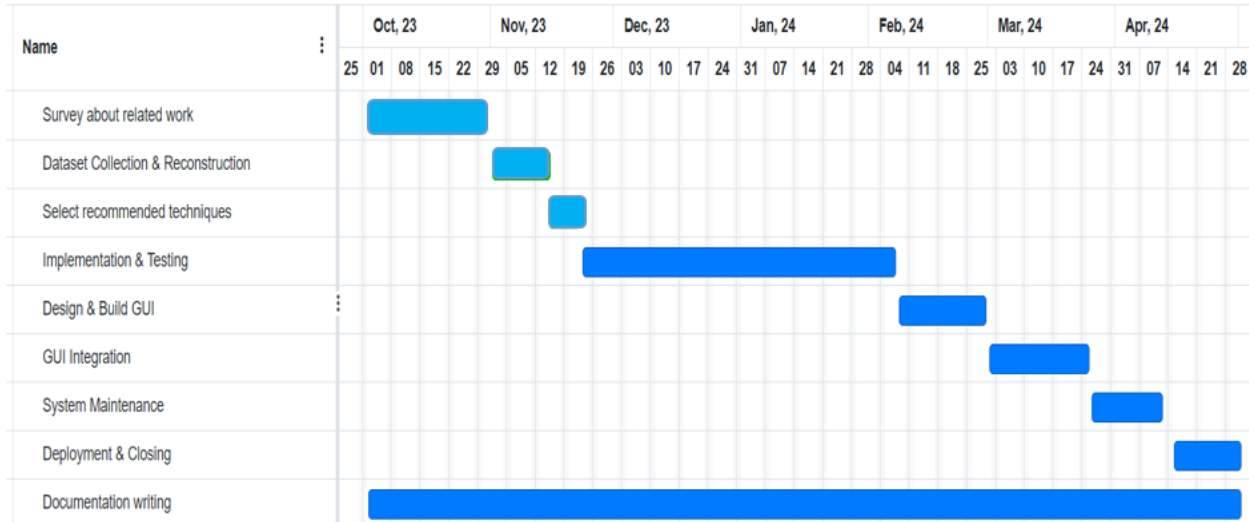


Fig.1.8 Project Time Plan

1.5 Documentation organization

The Project includes six chapters, that's the first. the chapters' description is presented briefly as follows:

Chapter 2 (Related work): This chapter presents a review of previous papers and projects that address similar topics as proposed project. It discusses the methodologies and results of these prior attempts, providing context and background for our work.

Chapter 3 (System Architecture): This chapter explains prior techniques used in common and common systems and their impact on model performance. It details how each technique contributes to improving the overall results of the system.

Chapter 4 (System Implementation): In this chapter, we will discuss the deep learning (DL) models used in all phases of the proposed system. It includes an overview of their architectures and a detailed description of their experimental results.

Chapter 5 (Experimental Result & User Interface): This chapter presents the experimental results for all phases of the system. It also showcases the design of the system's user

interface, describing how users interact with the smart glasses and the functionalities provided.

Chapter 6 (Conclusion and Future Work): The final chapter summarizes the key findings and contributions of the project. It also outlines potential future work, suggesting improvements and additional features that could further enhance the system's effectiveness

Chapter 2

Literature Review

Chapter 2: Literature Review

Before starting the project, you should have seen what the researchers had done before. We then scan the papers that address our interests to see what techniques and methods the researchers used in dealing with our project, so we read many papers that dealt with how to determine the ejection fraction of the heart.

In this chapter, we discuss some papers that follow the general system architecture and present the parts of the prior systems, classification methods, and results mentioned in each paper. For each paper, we show what data set was used, what classification algorithm was applied, and what are the criteria for calculating the system performance that the researchers obtained in the paper.

2.1 General Overview

The accurate assessment of left ventricular systolic function is crucial in the diagnosis and treatment of cardiovascular diseases. Left ventricular ejection fraction (LVEF) and global longitudinal strain (GLS) are the most critical indexes of cardiac systolic function. Echocardiography has become the mainstay of cardiac imaging for measuring LVEF and GLS because it is non-invasive, radiation free, and allows for bedside operation and real-time processing. However, the human assessment of cardiac function depends on the sonographer's experience, and despite their years of training, inter-observer variability exists. In addition, GLS requires post-processing, which is time consuming and shows variability across different devices.

Researchers have turned to artificial intelligence (AI) to address these challenges. The powerful learning capabilities of AI enable feature extraction, which helps to achieve accurate identification of cardiac structures and reliable estimation of the ventricular volume and myocardial motion. Hence, the automatic output of systolic function indexes can be achieved based on echocardiographic images. This review attempts to thoroughly explain the latest progress of AI in assessing left ventricular systolic function and differential diagnosis of heart diseases by echocardiography and discusses the challenges and promises of this new field [5].

Artificial Intelligence (AI) is an evolving field that focuses on creating intelligent machines capable of performing tasks that typically require human intelligence, as shown in Fig 2.1. It involves the development of algorithms and models that enable machines to learn from data, make decisions, and adapt to new information. AI encompasses various subfields, such as machine learning and deep learning and all related work is recent

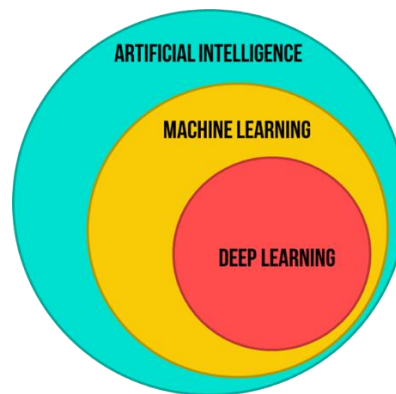


Fig 2 .1 Classification Techniques

Machine learning is a subset of artificial intelligence (AI) that focuses on using data and algorithms to mimic human learning and improve accuracy over time. In this approach, a machine learning model is trained on a dataset to recognize patterns by utilizing a reasoning algorithm. The traditional machine learning model follows a three-step process. Firstly, preprocessing cleans the dataset can use and improves the accuracy of ML model, using filters like low pass filter, high pass filter, and band pass filter. Secondly, feature extraction selects the most effective features from the dataset using loss function like zero crossing, mean frequency, root mean square, and standard deviation to minimize error between learning and datasets. Finally, classification employs suitable algorithms like SVM, RF, and LDA to accurately classify the dataset [6].

Deep learning (DL), a subset of machine learning (ML), enables computers to learn hierarchies of concepts from data, reducing the need for explicit human programming. Introduced in the 1950s, DL has revolutionized fields like voice recognition and medical imaging, categorizing complex concepts such as colorectal abnormalities with high accuracy. DL architectures include supervised networks like convolutional neural networks (CNNs), which use labeled data for training, and unsupervised networks like autoencoders, which learn from unlabeled data to

identify patterns. Hybrid deep networks combine these architectures to enhance performance in diverse applications, including medical image analysis, so Therefore, all work related to the system for determining the ejection fraction of the heart uses deep learning techniques and models.

2.2 Related Studies

Based on the image segmentation, researchers have devoted themselves to developing fully automatic models to assess LVEF.

Ouyang et al. [7] presented a video-based AI model called EchoNet Dynamic to evaluate cardiac function. The study included apical four chamber views of 10,030 patients with annotations of only end-systole and end-diastole frames. They generalized these labels with a weak supervision approach and developed an atrous convolution model to generate frame-level semantic segmentation of the whole cardiac cycle. The weak supervision approach could cut the cost of labeling considerably. The atrous convolution model can integrate sufficient information into the temporal domain, provide frame level segmentation of the left ventricle, and draw a left ventricular volume curve with multiple cardiac cycles to evaluate LVEF. The results showed that the EchoNet-Dynamic segmented the left ventricle accurately with a Dice similarity coefficient greater than 0.9, both at the end-systole and end-diastole levels, as well as across the cardiac cycle. The researchers further analyzed the accuracy of LVEF assessment by external validation of 2895 patients. It was revealed that the automatic measurements were consistent with the expert assessments and had good repeatability. In addition, the study discussed a series of problems that clinicians might experience in clinical practice, including the accuracy of model measurements of cardiac function in patients with arrhythmias and the measurements of different video qualities, using different instruments, and under different imaging conditions. The results showed that EchoNet-Dynamic was robust to variation in heart rhythm and video acquisition conditions.

Grant Duffy et al. [8] The method proposed here improves upon the segmentation approach by predicting the z-depth of the LV at each pixel. Any pixels outside the LV are zero. Figure 3 visualizes an example LV depth map. Depth maps are predicted for every frame of an echocardiogram video. The volume for each frame is calculated by summing the pixel depths. From the frame volumes, end systole and end diastole are selected and used to calculate EF. These depth maps can be easily visualized as a 3D surface over the LV. This approach is analogous to doing

a regression to depth for each pixel instead of classification for each pixel as seen in a segmentation model. CNNs have previously been used to predict depth information and is an approach that has become commonplace in computer vision and autonomous AI¹⁵. To calculate the volume for a given prediction, the pixel depths are simply summed together. This allows more of the work to be done by the model but is still closely aligned with the human method and therefore is easily interpreted.

Xin Liu et al. [9] A new strategy of feature extraction and fusion could enhance the accuracy of automatic LVEF assessment based on Multiview 2-D echocardiographic sequences. High diagnostic performance for the determination of heart failure was obtained by using DPS-Net in cases with different phenotypes of heart diseases. High performance for left ventricle segmentation was obtained by using DPS-Net, suggesting the potential for a wider range of application in the interpretation of 2DE images.

Ashley P. Akerman et al. [10] presented a novel AI HFpEF model which, based on only a single routinely acquired TTE video clip, accurately detected HFpEF, provided fewer nondiagnostic outputs than current clinical scores, and identified patients with worse survival. The application of this classifier in the screening for HFpEF, particularly when their diagnosis is uncertain, has the potential to automate an accurate detection process for a complex clinical syndrome, resulting in more patients getting a correct and expeditious diagnosis.

Anesth Analg. et al. [11] reported newly diagnosed HFrEF within 730 days postoperative to occur in 0.41% of adult patients surviving major non-cardiac surgery. Through machine learning, we model patterns within preoperative data and intraoperative response profiles to surgical and anesthetic interventions to detect HFrEF potentially in early stages. As developed by this study, the technique of intraoperative record segmentation by overlapping anesthetic and surgical interventions may prove to be a useful paradigm for future studies seeking analytic approaches to granular intraoperative data. Our findings may guide development of perioperative systems for early diagnosis and management of HFrEF; however, future studies must first be performed to (i) externally validate the detection algorithms, (ii) assess the feasibility of embedding algorithms into the EHR at the point of care, and (iii) understand the clinical effectiveness of such clinical decision support algorithms.

Francesc Formiga et al. [12] Proved that HFpEF is a very common condition that is associated with high morbidity and mortality. However, confirming a diagnosis of HFpEF is challenging, as affected patients have many comorbidities that can mimic the condition. Additionally, HFpEF is not defined based on a single criterion but on a cluster of parameters, mainly increased natriuretic peptide levels and specific echocardiographic alterations. We present a comprehensible algorithm that can easily be applied to real-world patients and prove useful when confirming or ruling out a diagnosis of HFpEF.

Sunita Pokhrel Bhattarai et al. [13] presented a comprehensive analysis of ECG features to predict reduced LVEF, highlighting the potential of automated 12-lead ECG analysis in enhancing diagnostic accuracy for heart failure. By employing the LASSO model, we identified key ECG features that significantly contribute to the estimation of $<30\%$ LVEF, with the model demonstrating a high sensitivity and moderate specificity. The study also acknowledges the limitations posed by the retrospective design and the need of validation in broader, prospective, and clinical trial populations. Nonetheless, our results offer promising insights into the utility of ECG-based diagnostic tools and artificial intelligence in cardiology, potentially paving the way for more accurate, real-time assessment of cardiac function and risk stratification in clinical settings, which could positively impact patient outcomes.

Bing Feng et al. [14] presented a method to calculate the LVEF without edge detection. Compared with QGS LVEF, our method gave better results for small LVs in computer simulations. In patient studies, our method gave results similar to those of the QGS method when imaging patients with large hearts. On the basis of our computer simulations, one can infer that our method gave more accurate measurements of the LVEF when imaging patients with small hearts, whereas the QGS method overestimated the LVEF. A method for calculating the LVEF without calculating the edges of the endocardial wall of the LV in images was developed. The method improves the accuracy of calculating functional parameters from in patients with small hearts. In patient studies, QGS overestimated the LVEFs in small hearts by $<9\%$ compared with those obtained with our method. This study concentrated on the calculation of the LVEF but the extension of this method to calculating other parameters such as heart wall thickness would not be difficult.

2.3 Competitive analysis

Tables 2.1 shows a brief comparison between the main recent studies. The table is compared in terms of three main criteria: dataset, classification model, and achieved results.

Table 2.1 Comparison between the Related Studies

Authors	Year	Model	Dataset	Results
Reynaud H. et al. [15]	2021	Transformer	EchoNet-Dynamic size=10,030	MAE = 5.95%, $R^2 = 0.52$
Ouyang A. et al. [7]	2020	EchoNet-Dynamic	EchoNet-Dynamic size=10,030	LV segmentation (DSC = 0.92), LVEF assessment (MAE = 4.1%), HFpEF classification (AUC 0.97).
Grant Duffy et al. [8]	2022	DeepLabV3	EchoNet-Dynamic size=10,030	MAE = 6.55%, $R^2 = 0.61$
Xin Liu et al. [9]	2021	DpsNet, DpsNet-V2	EchoNet-Dynamic size=10,030, CAMUS size=500	DC = 0.92

According to the current studies in the literature presented in the previous section, the following observations can be drawn:

- **Dataset Size:** The size of the dataset plays a crucial role in classification performance. EchoNet-Dynamic dataset, is more complex but it contains large number of 3D videos, yielded higher performance.
- **Classification model:** Deep learning models are more general, making them suitable for domain-specific problems. This is what makes it the most widely used.

Chapter 3

System Architecture

Chapter 3: System Architecture

In this chapter, the system architecture for a cardiac function assessment using deep learning techniques will be explained. First, the system contains many stitches, each one had many techniques and algorithms that can be applied. So, this chapter will describe each stage and its benefits and the most used and famous algorithms that can have been applied.

3.1 Overview

The traditional Machine Learning approaches includes different phases. As shown in Fig.3.1, The detailed description of the approach's phases will be presented in the following sub-sections.

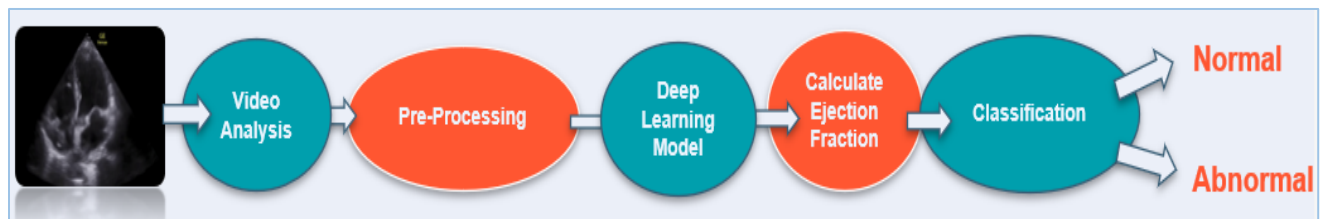


Fig.3.1 Traditional System Architecture

That the system architecture includes five stages be processing, feature extraction, classification and finally software application. The system architecture of cardiac function assessment. System using deep learning is critical for achieving accurate and efficient diagnosis of Such a serious organ which is heart. This chapter presents an overview of the architectural design, key components, and data flow of the system. The architecture leverages the power of deep learning algorithms to enhance. The detection and diagnosis capabilities of the system as shown in Fig 3.1, the architecture of the project.

3.2 Preprocessing

This is not to say data quality is not an a priori concern. Striving to collect high quality data for the task at hand is always important. But there are instances where deep learning engineers may blindly apply preprocessing and augmentation steps that reduce model performance on the same data. And, even when having high quality data, preprocessing allows the best possible results to be obtained. Understanding what preprocessing and augmentation are at their core enables data

scientists to get the most out of their input data. Image preprocessing are the steps taken to format images before they are used by model training and inference. This includes, but is not limited to, resizing, orienting, and color corrections.

Image augmentation are manipulations applied to images to create different versions of similar content in order to expose the model to a wider array of training examples. For example, randomly altering rotation, brightness, or scale of an input image requires that a model consider what an image subject looks like in a variety of situations. Image augmentation manipulations are forms of image preprocessing, but there is a critical difference: while image preprocessing steps are applied to training and test sets, image augmentation is only applied to the training data. Thus, a transformation that could be an augmentation in some situations may best be a preprocessing step in others [16].

Consider altering image contrast. A given dataset could contain images that are generally low contrast. If the model will be used in production on only low contrast in all situations, requiring that every image undergo a constant amount of contrast adjustment may improve model performance. This preprocessing step would be applied to images in training and in testing. However, if the collected training data is not representative of the levels of contrast the model may see in production, there is less certainty that a constant contrast adjustment is appropriate. Instead, randomly altering image contrast during training may generalize better. This would be augmentation. Knowing the context for data collecting and model inference is required to make informed preprocessing and augmentation decisions.

Identifying the correct preprocessing and augmentation steps most useful for increasing model performance requires a firm understanding of the problem, data collected, and production environment. What may work well in one situation is not appropriate in all others. Thus, considering techniques and why each may be valuable enables informed decisions. In this post, we'll surface considerations and provide recommendations that are generally best. Again, there is no free lunch, so even "generally best" tips can be disproven.

3.2.1 Image Resizing

Changing the size of an image sounds trivial, but there are considerations to take into account. Many model architectures call for square input images, but few devices capture perfectly square images. As shown in Fig 3.2, Altering an image to be a square call for either stretching its dimensions to fit to be a square or keeping its aspect ratio constant and filling in newly created "dead space" with new pixels.

Moreover, input images may be various sizes, and some may be smaller than the desired input size. Best tips: preserving scale is not always required, filling in dead pixels with reflected image content is often best, and down sampling large images to smaller images is often safest [17].

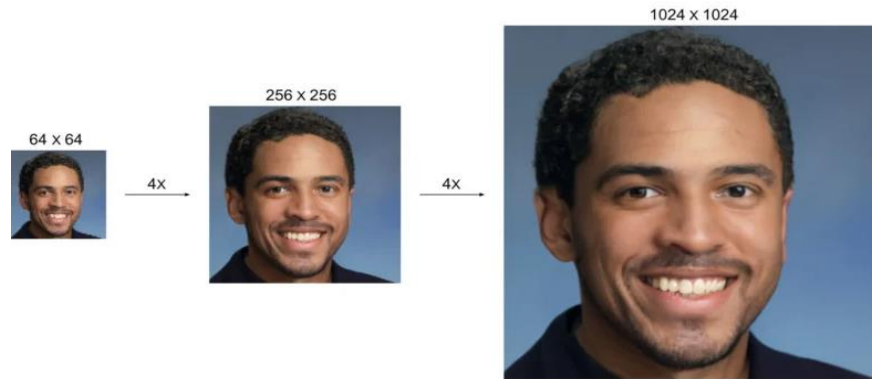


Fig 3.2 Image Resizing Effect

3.2.2 Image Orientation

When an image is captured, it contains metadata that tells our machines the orientation by which to display that input image relative to how it is stored on disk. That metadata is called its EXIF orientation, and inconsistent handling of EXIF data has long been a bane of developers everywhere. As shown in Fig 3.3, This applies to models, too: if we've created annotated bound boxes on how we perceived an image to be oriented, but our model is "seeing" that image in a different orientation, we're training the model completely wrong! Best tips: strip EXIF orientation from images and ensure pixels are all ordered the same way [18].



Fig 3.3 Image Orientation

3.2.3 Grayscale

Color changes are an example of image transformations that may be applied to all images (train and test) or randomly altered in training only as augmentations. Generally, gray scaling is a color change applied to all images. While we may think “more signal is always better; we should show the model color,” we may see more timely model performance when images are grayscale. As shown in Fig 3.4, Color images are stored as red, green, and blue values, whereas grayscale images are only stored as a range of black to white. This means for CNNs, our model only needs to work with one matrix per image, not three. Best tips: grayscale is fairly intuitive. If the problem at hand explicitly requires color (like delineating a white line from yellow line on roads), it’s not appropriate [19].



Fig 3.4 Image Grayscale

3.2.4 Random Flips

Randomly mirroring an image about its x- or y-axis forces our model to recognize that an object need not always be read from left to right or up to down. Flipping may be illogical for order-dependent contexts, like interpreting text. As shown in Fig 3.5, for most real-world objects, flipping is a strong way to improve performance [20].



Fig 3.5 Random Flips

3.2.5 Random Rotations

Rotating an image is particularly important when a model may be used in non-fixed position, like a mobile app. Rotating can be tricky as it, too, generates “dead pixels” on the edges of our images and, for bounding boxes, requires trigonometry to update any bounding boxes. As shown in Fig 3.6. Best tips: if an object may be a variety of orientations relative to the captured images, rotation is a good option. This would not be true for, say, screenshots, where the image content is always in a fixed position [21].

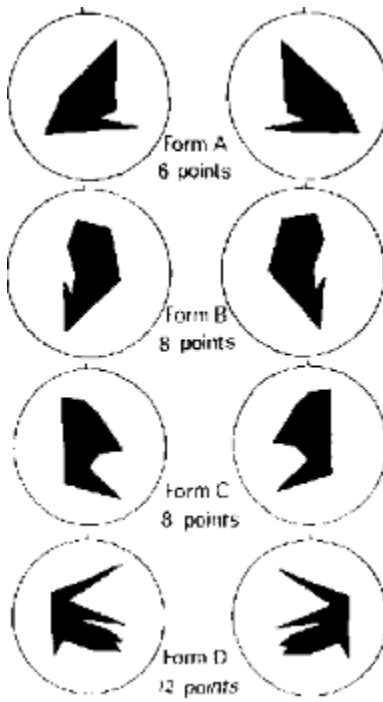


Fig 3.6 Random Rotations

3.2.6 Random Exposure

Adjusting image brightness to be randomly brighter and darker is most applicable if a model may be required to perform in a variety of lighting settings. It's important to consider the maximum and minimum of brightness in the room, as shown in Fig 3.7. Best tips: fortunately, brightness is fairly intuitive as well. Adjust brightness to match conditions the model will see in production relative to the images available for training [22].



Fig 3.7 Random Exposure

3.2.7 Random Noise

As shown in Fig 3.8, Adding noise to images can take a variety of forms. A common technique is “salt and pepper noise,” wherein image pixels are randomly converted to be completely black or completely white. While deliberately adding noise to an image may reduce training performance, this can be the goal if a model is overfitting on the wrong elements. Best tips: if a model is severely overfitting on image artifacts, salt and pepper noise can effectively reduce [23].



Fig 3.8 Removing Random Noise

3.3 Deep learning Models

Classification is a process of categorizing a given set of data into classes, it can be performed on both structured or unstructured data. The process starts with predicting the class of given data points. The classes are often referred to as target, label or categories.

The classification predictive modeling is the task of approximating the mapping function from input variables to discrete output variables. The main goal is to identify which class/category the new data will fall into. For classification stage, there are many models and techniques that can be used.

Deep learning models are used in many stages, from video analysis to calculating a value, which are the following stages:

- End-Systolic and End-Diastolic Frame Detection
- Find the Binary Mask of Left Ventricle (LV)
- Left Ventricle Landmark Detection
- Left Ventricle Volume Calculation

End-Systolic and End-Diastolic Frame Detection: Preprocessing involves normalizing numerical values, handling missing data, encoding categorical variables, and augmenting data to improve model accuracy and generalization. Techniques like scaling and normalization ensure consistent, representative data from real-world scenarios. The training process requires fixed-length videos for accurate identification of end-systolic and end-diastolic frames. The dataset contains single frames, so sequences of 128 are created, down-sampling longer ones. To improve transformer performance, two strategies were employed: Guided Random Sampling and Mirroring. Random sampling extended the distance between labeled frames, padded sequences with black frames, and masked them to avoid attention head interference. Mirroring created longer sequences. These techniques align with standard practices for transformer models, ensuring all seen ES and ED frames are labeled correctly while maintaining spatiotemporal coherence without introducing empty frames.

Initially, in find the binary mask of left ventricle talked about how segmentation consists of classification and localization. Let's understand how a u net architecture performs these two tasks and why it is so apt for segmentation. U-Net is the most used network for this stage. The encoder network is also called the contracting network. This encoder network consists of 4 encoder blocks. Each block contains two convolutional layers with a kernel size of 3×3 and valid padding, followed by a

Relu activation function. This is inputted to a max pooling layer with a kernel size of 2×2 . With the max pooling layer, we have halved the spatial dimensions learned, thereby reducing the computation cost of training the model. The decoder network is also called the expansive network. Our idea is to up-sample our feature maps to the size of our input image. This network takes the feature map from the bottleneck layer and generates a segmentation mask with the help of skip connections.

left ventricle landmark detection stage, it contains three steps. Firstly, identify the key vertical line, find the upper point by determining the maximum y-coordinate in the binary mask. Find the bottom points by determining the minimum coordinates, and the midpoint from these points. Ensure all identified points are within the binary mask. Then Divide the vertical line into 20 equal segments and add perpendicular lines at each division point, ensuring they extend to the binary mask edges. Finally, identify binary mask edges at perpendicular lines and mark landmark points as 21 lines or 42 points, including vertical line endpoints and perpendicular line intersections. By following these steps, you can accurately determine the Volume of the left ventricle from a binary mask using Simpson's method. This method involves precise identification of key points and careful division of the mask into segments for accurate volume calculation to get the EF from Simpson's Method by get ED and ES volume.

In left ventricle volume calculation stage, the major axis is the length between these two points. Subsequently, the disks are positioned orthogonal to the major axis, spaced at the positions obtained by dividing the major axis into 20 equal parts. At each disk position, a region of interest (ROI) is defined as a disk shape centered at that position, and the pixels within this ROI are extracted. To ascertain the diameter of each disk, the maximum distance between any two points within the ROI is computed. Fig 4.5 show the segmented images and their corresponding feature extraction at the end-systole and end-diastole, respectively.

3.4 Ejection Fraction (LVEF) Calculation

One crucial parameter for assessing heart function is the left ventricular (LV) ejection fraction (EF). EF is a measure of the heart's efficiency that is based on the percent volume change between the end diastolic volume (EDV) and the end systolic volume (ESV). It is the percentage of the left ventricle volume that is pumped out with each heartbeat. It has been demonstrated that low EF, even slightly diminished EF, is significantly suggestive of cardiac disease.

3.5 Classification

The classification process based on the value of ejection fraction is done in several ways. The simplest one is thresholding Normal ranges for two-dimensional echocardiography obtained LVEF as per the American Society of Echocardiography and the European Association of Cardiovascular Imaging. On the other hand, it can be applied more complex classification models as some deep learning models like: CNN, it is distinguished by the detection of the most accurate features from inputs without human intervention, which increases the effectiveness and accuracy of prediction.

It can be used Inception also specially inception v3, because of it solves the large problem of computations. This problem increases by adding a pooling layer. A new inception version is developed to make the calculation process less expensive. These models can be used also for classification stage: Residual Network, VGG network, and Yolo.

Chapter 4

Implementation Techniques

Chapter 4: Implementation Techniques

The system implementation chapter delves into the practical aspects of the proposed system. The proposed system meets the needs of doctors everywhere and helping them to provide an accurate diagnosis of heart failure to the Patients. The proposed system consists of several phases as shown in Fig 4.1

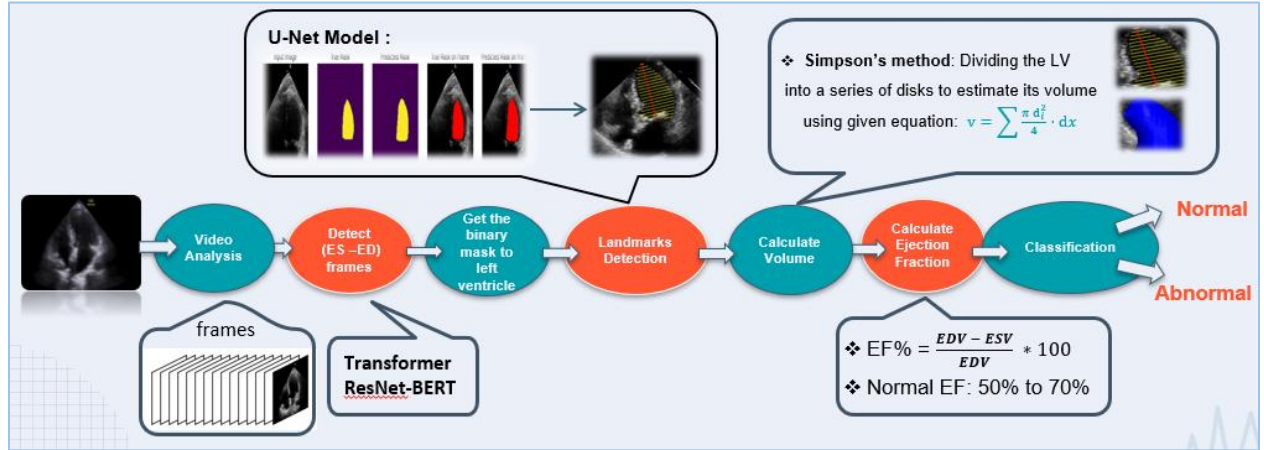


Fig.4.1 Proposed System Architecture

4.1 Video Analysis

In the video analysis phase, the process begins by converting the video into individual frames using Python code, typically with libraries such as OpenCV. As shown in Fig 4.2, This allows for detailed examination and manipulation of each moment captured in the footage. This frame-by-frame analysis is crucial for achieving precision and clarity in next phase which is Detecting the ES and ED Frames.

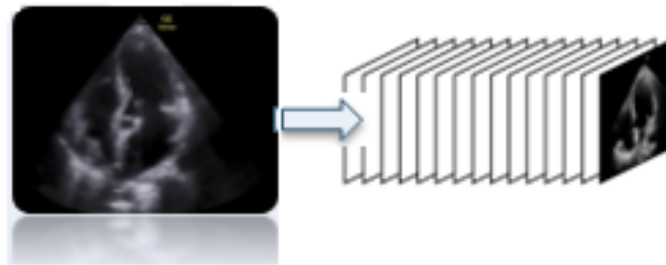


Fig.4.2 converting 3D video to frameset

4.2 End-Systolic and End-Diastolic Frame Detection

This phase is preceded by a preprocessor process. Preprocessing in deep learning models is a crucial step that involves transforming raw data into a suitable format for training. This process includes normalizing numerical values, handling missing data, encoding categorical variables, and augmenting data to increase the diversity and volume of training samples. Effective preprocessing enhances the quality of the data, reduces noise, and ensures that the model can learn the underlying patterns more efficiently. Techniques like scaling, normalization, and data augmentation help in improving model accuracy and generalization by ensuring that the data fed into the model is consistent and representative of real-world scenarios.

For training, we require videos of a fixed length to enable batch processing and accurately identify the ES (end-systolic) and ED (end-diastolic) frames. Our dataset contains only single ES and ED frame labels per video, leaving many true frames unlabeled. We opted to create sequences of 128 frames, based on the typical distances between ES and ED frames in the training data. Sequences longer than 128 frames are down sampled by a factor of 2 [30].

To address the issue of unlabeled ES and ED frames, which would negatively impact transformer performance, we used two strategies:

- Guided Random Sampling: We sampled the labeled frames and the frames in between them, extending by 10% to 70% of the distance between the labeled frames before and after the samples. We padded sequences with black frames to achieve the required 128-frame length and masked these frames so the transformer’s attention heads ignore them.
- Mirroring: We mirrored the transition frames between the labeled frames and placed them after the last labeled frame to create longer sequences. For instance, given a sequence $S = [f_ES, f_T1, \dots, f_TN, f_ED]$, where f_TN is the N-th transition frame, we extended it to $S' = [f_ES, f_T1, \dots, f_TN, f_ED, f_TN, \dots, f_ES, f_T1, f_ES]$. Sequences exceeding 128 frames were then randomly cropped to this length, ensuring the first frame was not always labeled.

These techniques align with standard practices for transformer models, ensuring all seen ES and ED frames are labeled correctly while maintaining spatio-temporal coherence without introducing empty frames.

- **Hyper Model (ResNetAE + BERT Encoders)**

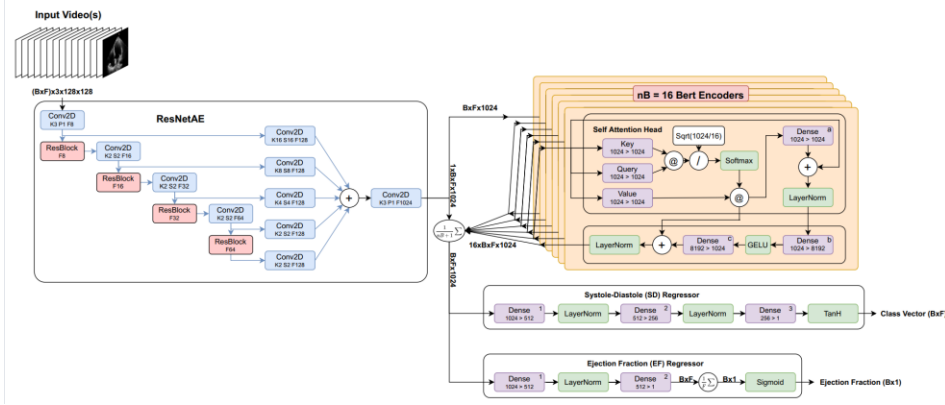


Fig.4.3 ResNetAE + BERT Encoders Architecture

As shown in Fig 4.3, This Hyper model Consist of 2 Main Parts:

1. ResNetAE is an encoder based on ResNet it consists of 4 Conv2D layers each of them followed by ResBlock and each result of them going to Conv2D layer then summed and going to Conv2D layer.
2. It encodes the feature effectively and reduces the size of the resulted encoded feature to reduce the computation of the followed transformer (BERT Encoders) Part.

The approach proposes using a new transformer model to interpret echo videos of varying lengths for accurate understanding and reasoning. The approach involves three main components: (a) an encoder for dimensionality reduction, (b) a Bidirectional Encoder Representations from Transformers (BERT)-based module for spatio-temporal reasoning, and (c) two regressors to predict key cardiac parameters (ES, ED frames and LVEF).

To manage computational complexity, the echo frames are first processed using a ResNetAE encoder to distill them into smaller embeddings. These embeddings are then fed into the transformer model, which incorporates both self-attention and attention mechanisms. The transformer, combined with the encoder, is trained end-to-end, optimizing the weights to enhance output accuracy [31].

Following feature extraction by the BERT encoders, the outputs are averaged with those from the ResNetAE and fed into two regressors. These regressors predict ES and ED frame indices and estimate LVEF. The LVEF estimation involves reducing the embedding dimensions for each frame, averaging predictions across frames, and employing regularization techniques to balance training objectives.

Overall loss is managed through Mean Squared Error (LMSE) and Mean Absolute Error (LMAE) metrics, along with a regularization term that emphasizes training focus on LVEF predictions. This comprehensive approach ensures robust performance in estimating cardiac parameters from echo videos.

4.3 Find the Binary Mask of Left Ventricle (LV)

Initially talked about how segmentation consists of classification and localization. Let's understand how a u net architecture performs these two tasks and why it is so apt for segmentation.

- **U-Net Network**

U-Net Architecture gets its name from its architecture. The “U” shaped model comprises convolutional layers and two networks. First is the encoder, which is followed by the decoder. With the U-Net, we can solve the above two questions of segmentation: “what” and “where.”

The encoder network is also called the contracting network. This network learns a feature map of the input image and tries to solve our first question- “what” is in the image? It is similar to any classification task we perform with convolutional neural networks except for the fact that in a U-Net, we do not have any fully connected layers in the end, as the output we require now is not the class label but a mask of the same size as our input image [32].

This encoder network consists of 4 encoder blocks. Each block contains two convolutional layers with a kernel size of 3×3 and valid padding, followed by a Relu activation function. This is inputted to a max pooling layer with a kernel size of 2×2 . With the max pooling layer, we have halved the spatial dimensions learned, thereby reducing the computation cost of training the model.

In between the encoder and decoder network, we have the bottleneck layer. This is the bottommost layer. As shown in Fig 4.4, It consists of 2 convolutional layers followed by Relu. The output of the bottleneck is the final feature map representation.

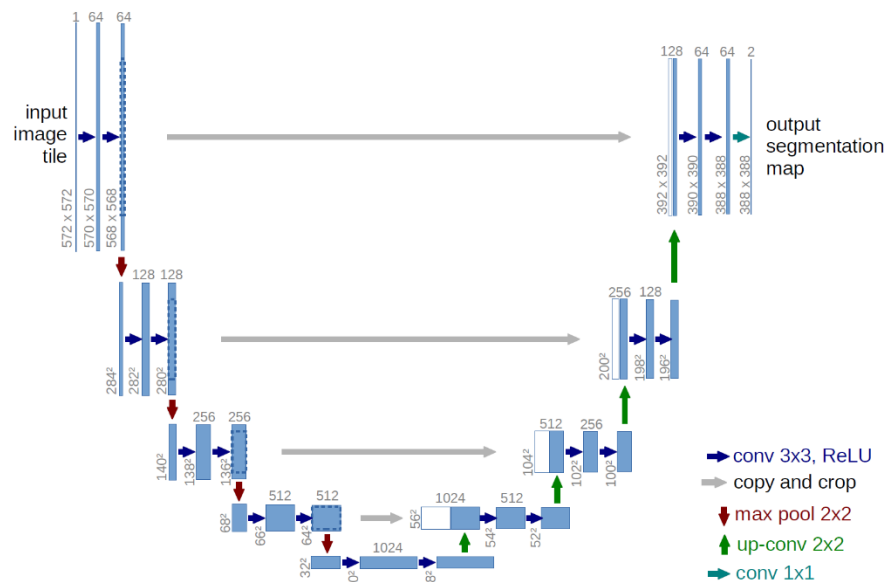


Fig.4.4 U-Net Network Architecture

Now, what makes U-Net so good at image segmentation is skip connections and decoder networks. What we have done till now is similar to any CNN. The skip connections and decoder network separate the u net architecture from other CNNs.

The decoder network is also called the expansive network. Our idea is to up sample our feature maps to the size of our input image. This network takes the feature map from the bottleneck layer and generates a segmentation mask with the help of skip connections. The decoder network tries to solve our second question-“where” is the object in the image? It consists of 4 decoder blocks. Each block starts with a transpose convolution (indicated as up-conv in the diagram) with a kernel size of 2×2 . After concatenating this output with the corresponding skip layer connection from the encoder block, the network utilizes two convolutional layers with a kernel size of 3×3 , followed by a Relu activation function.

Skip connections are indicated with a grey arrow in the model architecture. Skip connections help us use the contextual feature information collected in the encoder blocks to generate our segmentation map. The idea is to use our high-resolution features learned from the encoder blocks (through skip connections) to help us

project our feature map (output of the bottleneck layer). This helps us answer “where” is our object in the image. A 1×1 convolution follows the last decoder block with sigmoid activation which gives the output of a segmentation mask containing pixel-wise classification. This way, it could be said that the contracting path passes across information to the expansive path. And thus, we can capture both the feature information and localization with the help of a U-Net.

4.4 Left Ventricle Landmark Detection

1. Identifying the Key Vertical Line

- **Find the Upper Point:**
Identify the maximum y-coordinate in the binary mask. This point will be the uppermost point of the vertical line.
- **Find the Bottom Points:**
Determine the minimum (x, y) coordinate within the binary mask to get the leftmost bottom point.
Find the point with the maximum x-coordinate that has the same y-coordinate as the minimum (x, y) point to get the rightmost bottom point.
Get the midpoint from these two points.
- **Ensure All Points are Within the Binary Mask:**
Verify that all identified points (upper point, leftmost bottom point, rightmost bottom point, midpoint) are part of the binary mask.

2. Identifying the Key Horizontal Lines

- **Divide the Vertical Line into 20 Discs:**
Split the vertical line obtained in the first step into 20 equal segments.
- **Add Perpendicular Lines:**
At each division point, draw a line perpendicular to the vertical line.
Ensure these lines extend to the edges of the binary mask.

3. Adding Landmark Points

- **Identify Edges of the Binary Mask:**
At each perpendicular line, determine the intersection points with the binary mask edges.
- **Mark Landmark Points:**

Mark the intersection points from the previous step as landmarks. This will result in 21 lines or 42 points (including the vertical line's endpoints and the intersections of the perpendicular lines with the mask edges).

By following these steps, you can accurately determine the Volume of the left ventricle from a binary mask using Simpson's method. This method involves precise identification of key points and careful division of the mask into segments for accurate volume calculation to get the EF from Simpson's Method by get ED and ES volume.

4.5 Left Ventricle Volume Calculation

In our case, only A4C measurements are available; hence, the monoplane Simpson's method is utilized, which uses measurements obtained from A4C only. To extract features from the segmented images, the contour points (volume tracings) are derived from the segmented mask representing the left ventricle at the end-diastolic and end-systolic frames, respectively. The localization process identifies the mitral valve and apex points, which enable the determination of the major axis of the left ventricle [33].

The major axis is the length between these two points. Subsequently, the disks are positioned orthogonal to the major axis, spaced at the positions obtained by dividing the major axis into 20 equal parts. At each disk position, a region of interest (ROI) is defined as a disk shape centered at that position, and the pixels within this ROI are extracted. To ascertain the diameter of each disk, the maximum distance between any two points within the ROI is computed. Fig 4.5 show the segmented images and their corresponding feature extraction at the end-systole and end-diastole, respectively.

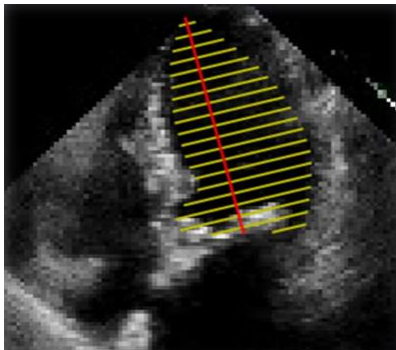


Fig.4.5 Left ventricle landmarks

$$LV \text{ Volume} = \sum \frac{\pi d_i^2}{4} \cdot dx \text{ (where } dx: \text{ verticle distance between any 2 disks)} \quad (4.1)$$

4.6 Left Ventricle Ejection Fraction (LVEF) Calculation

Left ventricular (LV) ejection fraction (EF) is one important metric for evaluating cardiac function. EF is defined by the percent volume change between the end diastolic volume (EDV) and the end systolic volume (ESV) as shown in Eq. 1 and describes what percent of the left ventricle volume is pumped out every heartbeat and is interpreted as a measure of the heart's efficiency. Low EF, even marginally reduced EF, has been shown to be highly indicative of heart disease [34].

$$LVEF\% = 100 * \frac{EDV - ESV}{EDV} \quad (4.2)$$

4.7 Classification

We used simple thresholding Normal ranges for two-dimensional echocardiography obtained LVEF as per the American Society of Echocardiography and the European Association of Cardiovascular Imaging [35].

LVEF (%) among the *male* population:

- 52% to 72% normal range
- 41% to 51 mildly abnormal
- 30% to 40% moderately abnormal
- Less than 30% severely abnormal

LVEF (%) among the *female* population:

- 54% to 74% normal range
- 41% to 53 mildly abnormal
- 30% to 40% moderately abnormal
- Less than 30% severely abnormal

Chapter 5

Experimental Results and User Interface

Chapter 5: Experimental Results

5.1 Datasets

Echo-Net dataset is standard full resting echocardiogram study consists of a series of videos and images visualizing the heart from different angles, positions, and image acquisition techniques, as shown in Fig 5.1. The dataset contains 10,030 apical-4-chamber echocardiography videos from individuals who underwent imaging between 2016 and 2018 as part of routine clinical care at Stanford University Hospital. Each video was cropped and masked to remove text and information outside of the scanning sector. The resulting images were then down-sampled by cubic interpolation into standardized 112x112 pixel videos.

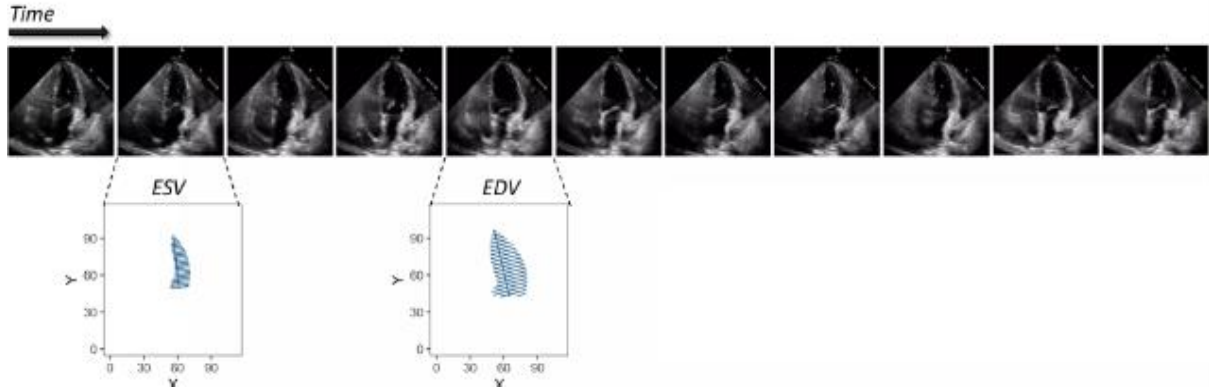


Fig. 5.1 Samples of Echo-Net dataset

In this dataset, for each video, the left ventricle is traced at the endocardial border at two separate time points representing end-systole and end-diastole. Each tracing is used to estimate ventricular volume by integration of ventricular area over the length of the major axis of the ventricle. The expert tracings are represented by a collection of paired coordinates corresponding to each human tracing. The first pair of coordinates represent the length and direction of the long axis of the left ventricle, and subsequent coordinate pairs represent short axis linear distances starting from the apex of the heart to the mitral apparatus. Each coordinate pair is also listed with a video file name and frame number to identify the representative frame from which the tracings match, as shown in Table 5.1.

Table 5.1 Echo-Net dataset label variables

Variable	Description
FileName	Hashed file name used to link videos, labels, and annotations
EF	Ejection fraction calculated from ESV and EDV
ESV	End systolic volume calculated by method of discs
EDV	End diastolic volume calculated by method of discs
Height	Video Height
Width	Video Width
FPS	Frames Per Second
NumFrames	Number of Frames in whole video
Split	Classification of train/validation/test sets used for benchmarking

In addition to the video itself, each study is linked to clinical measurements and calculations obtained by a registered sonographer and verified by a level 3 echocardiographer in the standard clinical workflow. A central metric of cardiac function is the left ventricular ejection fraction, which is used to diagnose cardiomyopathy, assess eligibility for certain chemotherapies, and determine indication for medical devices. The ejection fraction is expressed as a percentage and is the ratio of left ventricular end systolic volume (ESV) and left ventricular end diastolic volume (EDV) determined by $(EDV - ESV) / EDV$.

5.2 Experimental Results

Experiments have been conducted for both recognition and classification approaches. The primary objective of this work is not only to achieve higher Dice-Sørensen coefficient and lower Mean Square Error.

The Dice-Sørensen coefficient (DSC) is a statistic used to gauge the similarity of two samples. The formula for DSC is:

$$DSC = \frac{2 * |A \cap B|}{|A + B|} \quad (5.1)$$

where $|X|$ and $|Y|$ are the cardinalities of the two sets (i.e. the number of elements in each set). The Sørensen index equals twice the number of elements common to both sets divided by the sum of the number of elements in each set [37].

Mean Square Error (MSE) is a measure of the quality of an estimator. As it is derived from the square of Euclidean distance, it is always a positive value that decreases as the error approaches zero. measures the average of the squares of the errors that is, the average squared difference between the estimated values and the actual value. The formula for MSE is:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (5.2)$$

This section presents the results obtained using applied approach for detect the value of ejection fraction and classification results. The primary parameter varied in these experiments was the number of epochs. The MAE and DSC were used to evaluate the segmentation left ventricle as shown in Table 5.2.

Table 5.2 Left Ventricle Detection Results

Segmentation	Dice-Sørensen coefficient	Mean Square Error
Segmentation Diastolic	92.7	1.88
Segmentation systolic	90.9	1.51
Total Average	91.8	1.69

DSC is higher than the anther in the related papers. indicating transformer better classification of two data classes: End Diastolic, End-Systolic, as shown in Fig 5.2.

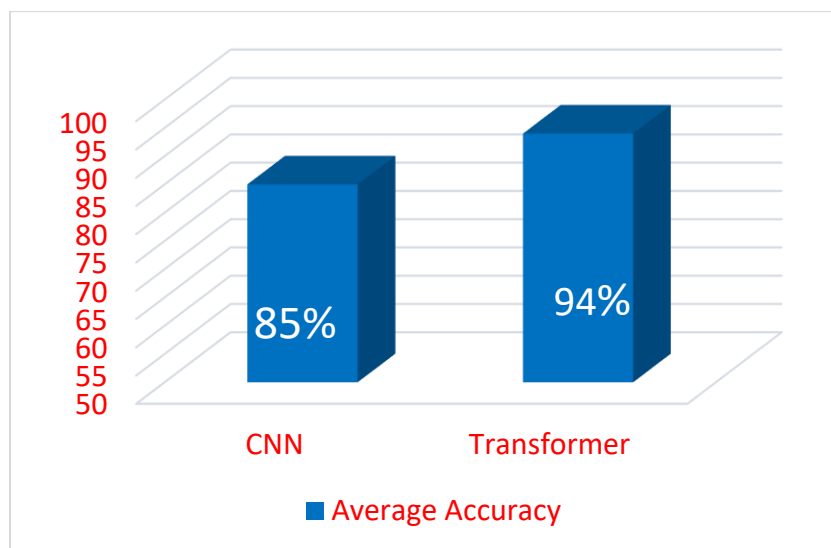


Fig. 5.2 Frames Detection Result

5.3 User Interface

A graphical user interface (GUI) is a visual interface that enables users to interact with electronic devices or software applications. It utilizes graphical elements such as icons, buttons, and windows to represent and manipulate data. GUIs enhance user experience by providing an intuitive and user-friendly environment for navigation and interaction.

- **Starting window**

As shown in Fig 5.3, The Initialization window when opening the Application.

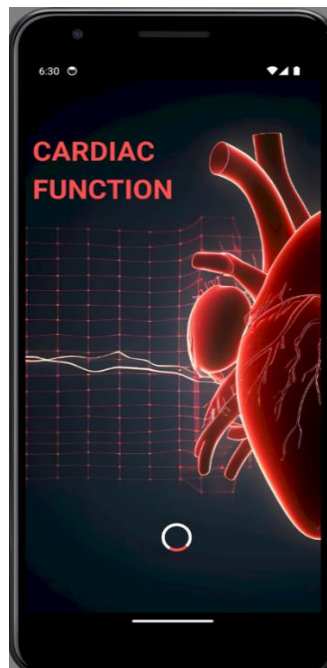


Fig. 5.3 Starting window

- **Main window**

If have an account: Here user can login to our application by adding his Email Address in the first text field which has hint text "Email Address" and password in the first text field which has hint text "Password." Then press on the blue button "Login." as shown in Fig 5.4.

If doesn't have an account: He can press on blue text at the bottom left of the screen "Register" To create a new account.

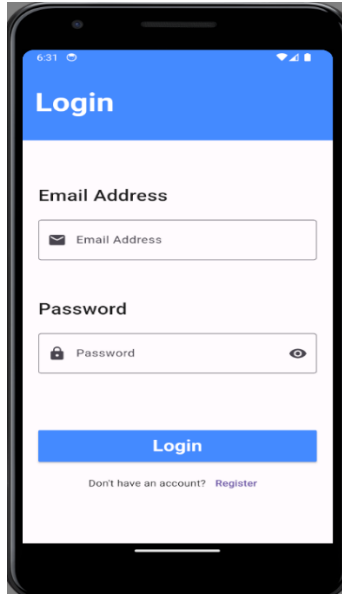


Fig. 5.4 Main window

- **Sign Up Window**

Here user can create new account by adding his name in the text field which has hint text” **User Name**”, adding his phone number in the text field which has hint text “**Email Address**”, password in the text field which has hint text “**Password**”, and write password again in “**Confirm Password**” to make sure that it is password which he wants to confirm as shown in Fig 5.5, Then check out box to agree our terms, then press on blue button “**SUBMIT**”

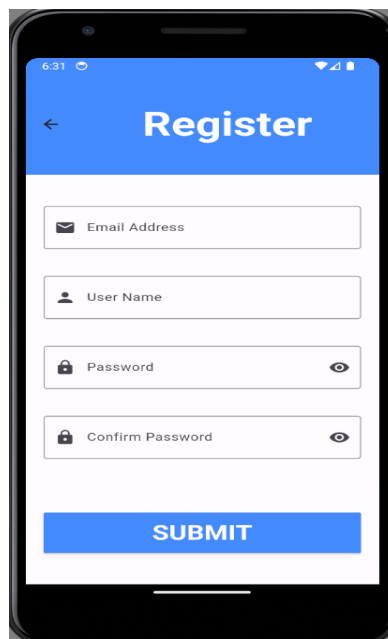


Fig. 5.5 Sign Up window

- **Upload Screen**

Here user can select video from his gallery by pressing on Upload button as shown in Fig 5.6, After Uploading the Video, “Echo Of The Heart” press The Button “Analyzing The Video”.

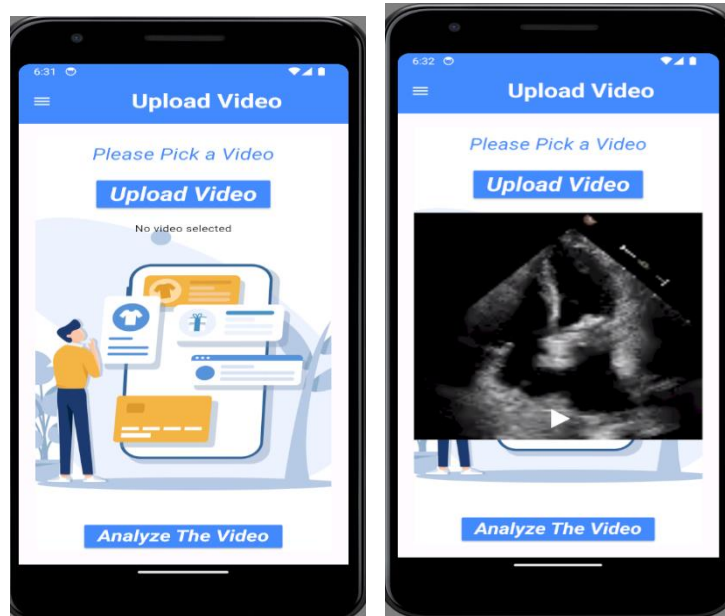


Fig. 5.6 Upload window

- **6.5 Loading Screen**

As shown in Fig 5.6, Here user waiting until Assessment process has been completed.

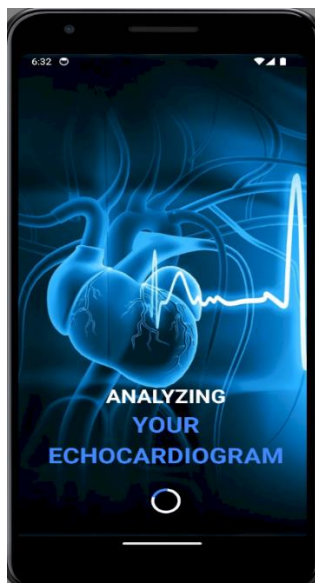


Fig. 5.7 Loading window

- **Result Screen**

As shown in Fig 5.8, Here user can see the result of assessment the cardiac function, the result of the End Systolic frame and the End diastolic frame from the video, and the segmentation of the left ventricle in the systolic and diastolic frames.

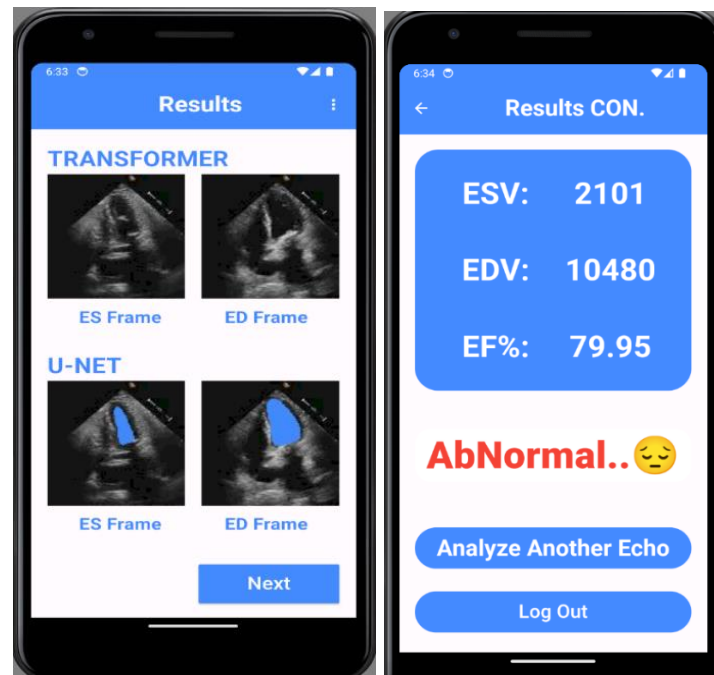


Fig. 5.8 Result window

Here the result of the End Systolic Volume and End Diastolic Volume and The Percentage of The Ejection fraction Value. Which define the patient state if he is Normal (50 – 70 %) or not. Additionally, for analyzing another video press the button “Analyze Another Echo”. And for going the main window and login with another account press the button “Logout”.

Chapter 6

Conclusion and Future work

Chapter 6: Conclusions and Future work

6.1 Conclusions

The percentage of blood pumped out of the left ventricle of the heart during each contraction or heartbeat is measured by a vital medical metric called left ventricular ejection fraction, or LVEF. Accurate measurement of left ventricular ejection fraction (LVEF) is crucial for the treatment of patients with heart problems. Furthermore, LVEF is a useful marker for preventing adverse outcomes in individuals suffering from illnesses like congestive heart failure, after a heart attack, or after revascularization treatments. A person's left ventricular ejection fraction can reveal if they have a cardiac condition or are normal. Quickly determining the left ventricular ejection fraction measurement can help medical professionals diagnose and treat the condition in its early stages.

Using deep learning models (DL), the suggested system in this work effectively ascertains whether an individual has a cardiac condition by recognizing the ejection fraction of a heartbeat from the echocardiography film. The system first does the video analysis stage of the echocardiography footage, splitting it into many frames and determining whether the heart is contracting or not in each frame. Next, use the binary mask to find the left atrium's location. In addition to estimating the heart's size based on the left atrium's specifications. Finally, but just as importantly, the method can determine the heart's ejection fraction by using its size in both diastole and systole.

Finally, a threshold method for the heartbeat's ejection fraction value is used to classify a person as either normal or ill. Effective experimental findings are obtained by the suggested system, which produced a mean square error of 1.69 and a dice coefficient similarity of 91.8. The suggested approach is embodied in an application that is easily navigable and satisfies the requirements of physicians worldwide, assisting them in accurately diagnosing heart failure in patients. The application might have a beneficial effect on the diagnosing procedure.

6.2 Future Work

In the future, the emphasis will be on improving performance through experimentation with diverse classification models, exploring various preprocessing methods, enriching datasets with additional images, and expanding the range of recognized classes. Moreover, there are plans to introduce new features or objects for recognition. Following these enhancements, the application can be deployed to aid charitable associations and institutions catering to special needs in their diagnostic processes.

References

1. Lang, R. M. et al. Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *Eur. Heart J. Cardiovasc.Imaging* 16, 233–271 (2015).
2. McMurray, J. J. et al. ESC guidelines for the diagnosis and treatment of acute and chronic heart failure 2012. *Eur. J. Heart Fail.* 14, 803–869 (2012).
3. Loefer, L. R., Rosamond, W. D., Chang, P. P., Folsom, A. R. & Chambless, L. E. Heart failure incidence and survival (from the Atherosclerosis Risk in Communities study). *Am. J. Cardiol.* 101, 1016–1022 (2008).
4. Bui, A. L., Horwich, T. B. & Fonarow, G. C. Epidemiology and risk profile of heart failure. *Nat. Rev. Cardiol.* 8, 30–41 (2011).
5. Roizen, M. F. Forecasting the future of cardiovascular disease in the United States: a policy statement from the American Heart Association. *Yearbook Anesthesiol. Pain Manage.* 2012, 12–13 (2012).
6. Sengupta PP, Huang YM, Bansal M, et al. Cognitive Machine-Learning Algorithm for Cardiac Imaging: A Pilot Study for Differentiating Constrictive Pericarditis from Restrictive Cardiomyopathy. *Circ Cardiovasc Imaging.* 9(6), (2016)
7. Ouyang, David & He, Bryan & Ghorbani, Amirata & Yuan, Neal & Ebinger, Joseph & Langlotz, Curtis & Heidenreich, Paul & Harrington, Robert & Liang, David & Ashley, Euan & Zou, James. (2020). Video-based AI for beat-to-beat assessment of cardiac function. *Nature.* 580. 10.1038/s41586-020-2145-8.
8. Duffy G, Jain I, He B, Ouyang D. Interpretable deep learning prediction of 3d assessment of cardiac function. *Pac Symp Biocomput.* 27:231-241. (2022)
9. Liu, Xin & Fan, Yiting & Li, Shuang & Chen, Meixiang & Li, Ming & Hau, William & Zhang, Heye & Xu, Lin & Lee, Alex. Deep learning-based automated left ventricular ejection fraction assessment using 2D echocardiography. *American Journal of Physiology-Heart and Circulatory Physiology.* (2021)

10. Ashley P. Akerman, Mihaela Porumb, Christopher G. Scott, Arian Beqiri, Agisilaos Chartsias. Automated Echocardiographic Detection of Heart Failure with Preserved Ejection Fraction Using Artificial Intelligence, *JACC Adv.* 2 (6), (2023)
11. Anesth Analg, Mathis MR, Engoren MC, Joo H, Maile MD, Aaronson KD, Burns ML, Sjoding MW, Douville NJ, Janda AM, Hu Y, Najarian K, Kheterpal S. Early Detection of Heart Failure with Reduced Ejection Fraction Using Perioperative Data Among Noncardiac Surgical Patients: A Machine-Learning Approach. (2020)
12. Francesc Formiga, Julio Nuñez, María José Castillo Moraga, Marta Cobo Marcos, María Isabel Egocheaga, Concha F. García-Prieto, Angel Trueba-Sáiz, Arantxa Matalí Gilarranz & José María Fernández Rodríguez. Diagnosis of heart failure with preserved ejection fraction: a systematic narrative review of the evidence. *Heart Failure Reviews.* 29, pp. 179–189, (2024)
13. Bhattarai, Sunita & Dzikowicz, Dillon & Xue, Ying & Block, Robert & Tucker, Rebecca & Bhandari, Shilpa & Boulware, Victoria & Stone, Breanne & Carey, Mary. Estimating Ejection Fraction from the 12 Lead ECG among Patients with Acute Heart Failure. *medRxiv: the preprint server for health sciences.* (2024)
14. Feng, Bing & Sitek, Arkadiusz & Gullberg, Grant. Calculation of the left ventricular ejection fraction without edge detection: Application to small hearts. *Journal of nuclear medicine: official publication, Society of Nuclear Medicine.* 43. 786-94, (2002).
15. Reynaud, Hadrien & Vlontzos, Athanasios & Hou, Benjamin & Beqiri, Arian & Leeson, Paul & Kainz, Bernhard. Ultrasound Video Transformers for Cardiac Ejection Fraction Estimation. (2021).
16. McDonagh, T.A.; Metra, M.; Adamo, M.; Gardner, R.S.; Baumbach, A.; Böhm, M.; Burri, H.; Butler, J.; Celutkienė, J.; Chioncel, O.; et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur. Heart J.* 2021, 42, 3599–3726.
17. Heidenreich, P.A.; Bozkurt, et al. 2022 Guideline for the Management of Heart Failure: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation* 2022, 145, e895–e1032.

- 18.Lang, R.M.; Badano, L.P.; Mor-Avi, V.; Afilalo, J.; Armstrong, A.; Ernande, L.; Flachskampf, F.A.; Foster, E.; Goldstein, S.A.;Kuznetsova, T.; et al. Recommendations for cardiac chamber quantification by echocardiography in adults: An update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *Eur. Heart J. Cardiovasc. Imaging* 2015, 16, 233–270.]
- 19.Nauta, J.F.; Jin, X.; Hummel, Y.M.; Voors, A.A. Markers of left ventricular systolic dysfunction when left ventricular ejectionfraction is normal. *Eur. J. Heart Fail.* 2018, 20, 1636–1638.
- 20.Potter, E.; Marwick, T.H. Assessment of Left Ventricular Function by Echocardiography: The Case for Routinely Adding GlobalLongitudinal Strain to Ejection Fraction. *JACC Cardiovasc. Imaging* 2018, 11, 260–2746. Wallis, C. How Artificial Intelligence Will Change Medicine. *Nature* 2019, 576, S48.
- 21.Litjens, G.; Ciompi, F.; Wolterink, J.M.; de Vos, B.D.; Leiner, T.; Teuwen,J.; Išgum, I. State-of-the-Art Deep Learning in Cardiovascular Image Analysis. *JACC Cardiovasc. Imaging* 2019, 12, 1549–1565.
- 22.Davis, A.; Billick, K.; Horton, K.; Jankowski, M.; Knoll, P.; Marshall, J.E.; Paloma, A.; Palma, R.; Adams, D.B. Artificial Intelligence and Echocardiography: A Primer for Cardiac Sonographers. *J. Am. Soc. Echocardiogr.* 2020, 33, 1061–1066.
- 23.LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* 2015, 521, 436–444.
- 24.Krittanawong, C.; Johnson, K.W.; Rosenson, R.S.; Wang, Z.; Aydar, M.; Baber, U.; Min, J.K.; Tang, W.H.W.; Halperin, J.L.; Narayan,S.M. Deep learning for cardiovascular medicine: A practical primer. *Eur. Heart J.* 2019, 40, 2058–2073.
- 25.Nolan, M.T.; Thavendiranathan, P. Automated Quantification in Echocardiography. *JACC Cardiovasc. Imaging* 2019, 12, 1073–1092.
- 26.Khamis, H.; Zurakhov, G.; Azar, V.; Raz, A.; Friedman, Z.; Adam, D. Automatic apical view classification of echocardiograms using a discriminative learning dictionary. *Med. Image Anal.* 2017, 36, 15.

- 27.Madani, A.; Ong, J.R.; Tibrewal, A.; Mofrad, M.R.K. Deep echocardiography: Data-efficient supervised and semi-supervised deep learning towards automated diagnosis of cardiac disease. NPJ Digit. Med. 2018.
- 28.Madani, A.; Arnaout, R.; Mofrad, M.; Arnaout, R. Fast and accurate view classification of echocardiograms using deep learning.NPJ Digit. Med. 2018.
- 29.Narang, A.; Bae, R.; Hong, H.; Thomas, Y.; Surette, S.; Cadieu, C.; Chaudhry, A.; Martin, R.P.; McCarthy, P.M.; Rubenson, D.S.; et al.Utility of a Deep-Learning Algorithm to Guide Novices to Acquire Echocardiograms for Limited Diagnostic Use. JAMA Cardiol.2021, 6, 624–632.
- 30.Leclerc, S.; Smistad, E.; Pedrosa, J.; Ostvik, A.; Cervenansky, F.; Espinosa, F.; Espeland, T.; Berg, E.A.R.; Jodoin, P.-M.; Grenier, T.; et al. Deep Learning for Segmentation Using an Open Large-Scale Dataset in 2D Echocardiography. IEEE Trans. Med. Imaging 2019, 38, 2198–2210.
- 31.Smistad, E.; Salte, I.M.; Ostvik, A.; Leclerc, S.; Bernard, O.; Lovstakken, L. Segmentation of apical long axis, four-and twonchamber views using deep neural networks. In Proceedings of the IEEE International Ultrasonics Symposium, Glasgow, UK, 6–9 October 2019.
- 32.Leclerc, S.; Smistad, E.; Ostvik, A.; Cervenansky, F.; Espinosa, F.; Espeland, T.; Rye Berg, E.A.; Belhamissi, M.; Israilov, S.; Grenier, T.; et al. LU-Net: A Multistage Attention Network to Improve the Robustness of Segmentation of Left Ventricular Structures in 2-D Echocardiography. IEEE Trans. Ultrason. Ferroelectr. Freq. Control. 2020, 67, 2519–2530.
- 33.Ouyang, D.; He, B.; Ghorbani, A.; Yuan, N.; Ebinger, J.; Langlotz, C.P.; Heidenreich, P.A.; Harrington, R.A.; Liang, D.H.; Ashley, E.A.; et al. Video-based AI for beat-to-beat assessment of cardiac function. Nature 2020, 580, 252–256.
- 34.Asch, F.M.; Poilvert, N.; Abraham, T.; Jankowski, M.; Cleve, J.; Adams, M.; Romano, N.; Hong, H.; Mor-Avi, V.; Martin, R.P.; et al.Automated Echocardiographic Quantification of Left Ventricular Ejection Fraction without Volume Measurements Using a Machine Learning Algorithm Mimicking a Human Expert. Circ. Cardiovasc. Imaging 2019, 12, e009303.
- 35.Zhang, J.; Gajjala, S.; Agrawal, P.; Tison, G.H.; Hallock, L.A.; Beussink-Nelson, L.; Lassen, M.H.; Fan, E.; Aras, M.A.; Jordan, C.; et al. Fully

Automated Echocardiogram Interpretation in Clinical Practice. *Circulation* 2018, 138, 1623–1635.

36. Salte, I.M.; Østvik, A.; Smistad, E.; Melichova, D.; Nguyen, T.M.; Karlsen, S.; Brunvand, H.; Haugaa, K.H.; Edvardsen, T. Lovstakken, L.; et al. Artificial Intelligence for Automatic Measurement of Left Ventricular Strain in Echocardiography. *JACC Cardiovasc. Imaging* 2021, 14, 1918–1928.
37. Quer, G.; Arnaout, R.; Henne, M.; Arnaout, R. Machine Learning and the Future of Cardiovascular Care: JACC State-of-the-Art Review. *J. Am. Coll. Cardiol.* 2021, 77, 300–313.
38. Tokodi, M.; Shrestha, S.; Bianco, C.; Kagiya, N.; Casacang-Verzosa, G.; Narula, J.; Sengupta, P.P. Interpatient Similarities in Cardiac Function: A Platform for Personalized Cardiovascular Medicine. *JACC Cardiovasc. Imaging* 2020, 13, 1119–1132.