Faculty of Cognitive Sciences and Human Development

A DEEP LEARNING APPROACH FOR HEART DISEASE CLASSIFICATION USING CONVOLUTIONAL NEURAL NETWORK

Arulmolly Annathurai

Bachelor of Science with Honours
(Cognitive Science)
2019

UNIVE: :AWAK

Grade: A-

Please tick one

Final Year Project Report ☒

Masters ☐

PhD ☐

DECLARATION OF ORIGINAL WORK

This declaration is made on the 28 day of MAY year 2019.

Student's Declaration:

I, ARULMOLLY ANNATHURAI, 55526, FACULTY OF COGNITIVE SCIENCES AND HUMAN DEVELOPMENT, hereby declare that the work entitled, A DEEP LEARNING APPROACH FOR HEART DISEASE CLASSIFICATION USING CONVOLUTIONAL NEURAL NETWORK is my original work. I have not copied from any other students' work or from any other sources with the exception where due reference or acknowledgement is made explicitly in the text, nor has any part of the work been written for me by another person.

28/05/2019

Arulmolly Annathurai (55526)

Supervisor's Declaration:

I, DR. ABDULRAZAK YAHYA SALEH AL-HABABI , hereby certify that the work entitled, A DEEP LEARNING APPROACH FOR HEART DISEASE CLASSIFICATION USING CONVOLUTIONAL NEURAL NETWORK was prepared by the aforementioned or above mentioned student, and was submitted to the "FACULTY" as a *partial/full fulfillment for the conferment of BACHELOR OF SCIENCE WITH HONOURS (COGNITIVE SCIENCE), and the aforementioned work, to the best of my knowledge, is the said student's work.

Received for examination by: _____ Date: 12/6/2019

(DR. ABDULRAZAK YAHYA SALEH AL-HABABI)

I declare this Project/Thesis is classified as (Please tick (√)):

☐ **CONFIDENTIAL** (Contains confidential information under the Official Secret Act 1972)*

☐ **RESTRICTED** (Contains restricted information as specified by the organisation where research was done)*

☒ **OPEN ACCESS**

I declare this Project/Thesis is to be submitted to the Centre for Academic Information Services (CAIS) and uploaded into UNIMAS Institutional Repository (UNIMAS IR) (Please tick (√)):

☒ **YES**
☐ **NO**

Validation of Project/Thesis

I hereby duly affirmed with free consent and willingness declared that this said Project/Thesis shall be placed officially in the Centre for Academic Information Services with the abide interest and rights as follows:

- This Project/Thesis is the sole legal property of Universiti Malaysia Sarawak (UNIMAS).
- The Centre for Academic Information Services has the lawful right to make copies of the Project/Thesis for academic and research purposes only and not for other purposes.
- The Centre for Academic Information Services has the lawful right to digitize the content to be uploaded into Local Content Database.
- The Centre for Academic Information Services has the lawful right to make copies of the Project/Thesis if required for use by other parties for academic purposes or by other Higher Learning Institutes.
- No dispute or any claim shall arise from the student himself / herself neither a third party on this Project/Thesis once it becomes the sole property of UNIMAS.
- This Project/Thesis or any material, data and information related to it shall not be distributed, published or disclosed to any party by the student himself/herself without first obtaining approval from UNIMAS.

Student's signature: _____ Supervisor's signature: _____

Date: 28/05/2019 Date: 12/6/2019

Current Address:
University Malaysia Sarawak (UNIMAS)

Notes: * If the Project/Thesis is **CONFIDENTIAL** or **RESTRICTED**, please attach together as annexure a letter from the organisation with the date of restriction indicated, and the reasons for the confidentiality and restriction.

A DEEP LEARNING APPROACH FOR HEART DISEASE CLASSIFICATION USING CONVOLUTIONAL NEURAL NETWORK

ARULMOLLY ANNATHURAI

This project is submitted in partial fulfilment of the requirements of the degree Bachelor of Science with Honours (Cognitive Science)

Faculty of Cognitive Sciences and Human Development

UNIVERSITY MALAYSIA SARAWAK

(2019)

The project entitled 'A Deep Learning Approach For Heart Disease Classification Using Convolutional Neural Network' was prepared by Arulmolly Annathurai and submitted to the Faculty of Cognitive Sciences and Human Development in partial fulfillment of the requirements for a Bachelor of Science with Honours (Cognitive Science).

Received for examination by:

(DR.ABDULRAZAK YAHYA SALEH AL-HABABI)

Date:

12/6/2019

Grade
A-

ACKNOWLEGMENT

I thank God for blessing me to complete this work. I thank Him for giving me the strength, courage, and wisdom. I am grateful to Him for giving me a stellar team of lecturers who have laid the necessary foundation for me to complete my thesis successfully.

To my supervisor, Dr. Abdulrazak Yahya Saleh Al-Hababi, I appreciate the effort, and guidance that he has invested in me throughout this process. Supervisor Dr. Razak, thank you for believing in me and for encouraging me throughout this journey with your inspiring words. You have nurtured the potential in me so patiently all these months.

To my parents, Annathurai and Manimala, thank you for your unfailing love and support. To my siblings, Malani and Tanasilan, thank you for your love. To my friends, I appreciate your prayers and timely words of encouragement. I am so blessed to have everyone in my journey of completion.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Cardiovascular disease is the leading cause of death, and specialists estimate that roughly half of all heart attacks and strokes happen in individuals who have not been flagged as 'at risk.' Hence, there's an urgent need to progress the exactness of heart infection diagnosis. To this end, we explore the potential of utilizing information examination, and in specific the plan and utilize of convolutional neural networks (CNN) for classify heart disease based on ultrasound heart images. Our fundamental commitment is the plan, assessment, and optimization of CNN models of expanding profundity for heart disease classification. Moreover, a system with high precision and accuracy is required to analyse heart disease classification. This study utilized ultrasound scanned heart images which are collected from local hospital. A total of 50 images which includes 20 normal and 30 abnormal heart images. The model has been run successfully without any errors and produce a high accuracy of 96% after running for 46 epochs. Furthermore, a comparative study has been done between SVM and CNN using the similar dataset to analyse performance based on the accuracy.

ABSTRAK

Penyakit kardiovaskular merupakan punca utama kematian, dan pakar-pakar menganggarkan bahawa kira-kira separuh daripada semua serangan jantung berlaku pada individu yang telah flagged sebagai 'pada risiko.' Justeru, terdapat keperluan segera untuk kemajuan diagnosis jangkitan jantung. Untuk tujuan ini, kami meneroka potensi dan kepentingan maklumat peperiksaan, dan khusus rancangan dan menggunakan rangkaian neural convolutional (CNN) bagi penyakit jantung diklasifikasikan berdasarkan imej ultrasound jantung. Komitmen asas adalah pelan, penilaian, dan pengoptimuman model CNN untuk memperluaskan kedalaman untuk pengelasan penyakit jantung. Selain itu, sistem yang berketepatan tinggi dan ketepatan yang diperlukan untuk menganalisis pengelasan penyakit jantung. Ultrasound kajian yang digunakan ini satu imej yang dikumpul dari hospital tempatan yang diimbas. Sejumlah 50 imej yang merangkumi 20 biasa dan 30 gambar jantung yang tidak normal. Model yang telah dijalankan berjaya tanpa sebarang ralat dan menghasilkan ketepatan yang tinggi sebanyak 96% selepas berjalan untuk 46 epochs. Selain itu, satu kajian perbandingan telah dilakukan antara SVM dan CNN menggunakan dataset serupa untuk menganalisa prestasi berdasarkan ketepatan.

CHAPTER 1

INTRODUCTION

1.1 Introduction

In 2016, there were 162,201 individuals that died due to heart diseases in Malaysia (World Health Organization, 2016). In addition, heart diseases are consistently ranked the leading causes of death globally. Due to the aging population and expanding predominance of hazard components such as diabetes and weight, the number of heart diseases related deaths are anticipated to extend by 30% which is 23.3 million a year in 2030 (Murray, 2012). Pointing at earlier diagnosis and disease administration, a few non-invasive imaging choices have been designed for the evaluation of heart diseases.

Cardiovascular Magnetic Resonance (CMR) imaging has demonstrated to be of specific incredible value in heart disease diagnosis and administration. A combination of variables such as need of ionizing radiation, excellent delicate tissue differentiation and high reproducibility have made it the favored imaging modality within the measurement of ventricular volumes, myocardial work and scarring (Salerno,2017). Increasing clinical utilisation has resulted in an expanded application of CMR in large cohort studies (Medrano,2015). This expansion of restorative imaging datasets will increase the requirements for automated tools, making machine learning for enormous imaging information a really promising field.

A few propels in deep learning have empowered machines to beat people in image classification if given with a database of millions of images as within the ImageNet challenge (Russakovsky,2015). Within the restorative field, exceedingly assignment particular databases of restorative photos have shown capable of exact determination (Esteva,2017). For instance, utilized dermatology images to train a show that recognizes skin cancer, and utilized

retinal fundus images for detection of diabetic retinopathy (Gulshan,2016) Both scored on standard with certified clinical specialists in their respective field. Whereas the basic information is exceptionally diverse from cardiac volumetric information, the wide range of application appears that deep learning can indeed be utilized for conclusion within the medical field.

Current machine learning based strategies for automated cardiac determination focuses on detection and segmentation of the heart, followed by the extraction of highlights that are at that point utilized for diagnosis. This approach is reflected within the 2017 Automated Cardiac Diagnosis Challenge (ACDC) where the point is to consequently perform fragmentation and diagnosis on a 4D cineCMR scan. According to Bernard (2008) all but one member within the segmentation portion of the challenge utilized profound learning and scoring on standard with clinical experts, whereas none of the members within the classification part did. Instead they performed classification utilizing support vector machines (SVM) and random forests (RF) on handcrafted highlights extricated from gotten segmentation maps (Bernard,2018).

As the handcrafted features characterize clinical diagnosis of the pathology, extraction of those features could be a sensible approach. Subsequent diagnosis utilizing machine learning based methods shows the request for adaptability within the current demonstrative prepare that cannot easily be captured in rule-based strategies. Deep learning can give the desired adaptability and perform accurate segmentation and classification as appeared in state-of-the-art strategies (He,2018). However, no attempts have been made at end-to-end learning for demonstrative purposes in cardiology. One explanation for usually inadequately information, a repeating articulation for deep learning in medical image analysis (Zreik,2018).

In this study a Convolutional Neural Network (CNN) model is developed to classify heart disease from ultrasound scanned heart images. These cardiac ultrasound images are specifically obtained from hospital in Malaysia to represent the Malaysian's cardiac data. This CNN model classifies heart disease with high accuracy by scanning a group of data in one go. It is able to classify the images into normal and abnormal heart images. Moreover, the system can classify heart disease according to the abnormalities.

1.2 Background of Study

Disease prediction utilizing patient treatment history and health information by applying data mining and machine learning procedures are continuous battle for the past decades. Many works have been applied information mining methods on pathological information or medical profiles to forecast specific diseases. These approaches attempted to foresee the reoccurrence of illness. Moreover, a few approaches try to do forecast on control and movement of the disease.

The recent achievement of deep learning in disparate regions of machine learning has driven a move towards machine learning models that can learn rich, progressive representations of raw information with small pre-processing and deliver more accurate results (Long,2015). Numbers of papers have been published on a few data mining strategies for diagnosis of heart disease such as Choice Tree, Credulous Bayes, neural arrange, part density, automatically characterized bunches, stowing calculation and back vector machine appearing different levels of exactness's in illnesses forecast.

There has not been much research work done on utilizing Convolutional Neural Networks (CNNs) (A. Krizhevsky,2017) to analyse Magnetic Resonance Image (MRI) scans of the heart. Probably, this challenge represents one of the outranging attempts at this.

However, there is a significant amount of writing on applying machine learning models on radiology issues. According to Chen (2017) almost all machine learning algorithms are useful for the medical image division, brain function or action examination, and content-based image retrieval systems for Computer Topography (CT) or MRI images are additionally a less recent, but valuable perusing for this reason.

1.3 Problem Statement

According to Datuk Dr Ahmad Khairuddin, the National Heart Institute (IJN) consultant cardiologist, Malaysians are developing heart disease at a younger age compared with their peers in other countries such as Thailand and Indonesia. Cardiovascular diseases are the number one cause of death globally (WHO, 2016). For those affected, early defection is basic for both management and treatment. One of the leading symptomatic tools in this range is cardiac imaging—including attractive reverberation (MR), ultrasound, and computed tomography (CT) (Wetstone,2018).

Cardiac imaging can evaluate both heart life systems and work and help within the discovery of different heart-related pathologies, such as coronary artery infection, cardiac masses, and inherent heart disease (Medical Informed Consent, 2013). Computerized approaches to processing these images are in high request to lighten the burden on radiologists, trusting to move forward both diagnostic precision and efficiency.

Radiologists today confront an ever-increasing sum of medical images to review. In a sense, the medical imaging community is a victim of it own success; this blast in images are fundamentally due to advancements in medical imaging technologies, which have implied both more images per person check and a more prominent request for scans as a diagnostic tool (Litjens, 2017). Many researched in this field on deep learning approach for heart

disease classification. For instance, Emad (2015), who conducted a research on automatic localisation of left ventricle on MRI images.

However, only fewer studies found utilizing local (Malaysia) data. This study is fully focused on local data that is gathered from local hospitals. In addition, a system with high accuracy and fast processing is needed to analyse a huge sum of data in a short period.

1.4 Research Questions

1. How to pre-process the data of scanned heart images?
2. How to apply deep learning for heart disease classification system?
3. How effective is the algorithm and its performance accuracy?

1.5 Research Objectives

1. To pre-process the data of scanned heart images.
2. To apply deep learning for heart disease classification system.
3. To evaluate the effectiveness of algorithm based on the performance accuracy.

1.6 Significance of Study

This study is focused on classifying heart diseases or heart abnormalities by using Convolutional Neural n\Network (CNN). Numerous studies have been published worldwide in this similar area in the past years and applying deep learning techniques in this field has been increasing annually (Emad, 2015). However, these study that I mentioned earlier uses dataset from foreign continentals like America, Australia and Europe. Only a number of studies had been done using Asian people's CMR.

This study is fully based on data that have been collected from local hospitals on heart abnormalities. A series of ultrasound scanned images of the heart have been gathered to pre-process the data accordingly. Referring to all the previous studies in this field, I used the Convolutional Neural Network (CNN) technique to classify heart images into normal and abnormal classes. The gathered dataset consist of normal and abnormal heart scanned images of patients. The abnormal class contain three major types of heart disease which are pericardial effusion, dilated cardiomyopathy and left ventricle hypertrophy.

1.7 Definition of Terms

Table 1.1

Definition of Terms

Term	Conceptual definition	Operational definition
Heart disease	Heart disease generally refers to conditions that involve narrowed or blocked blood vessels that can lead to a heart attack, chest pain (angina) or stroke. Other heart conditions, such as those that affect your heart's muscle, valves or rhythm, also are considered forms of heart disease.(Mayo Foundation for Medical Education and Research (MFMER)	The major causes for heart failures and heart attacks. For example, abnormal size of heart chambers.
Prediction	Say or estimate that (a specified thing) will happen in the future or will be a consequence of something.(Oxford Dictionary)	What will happen next and the possible outcome of scanning the heart images.

Table 1.1

Definition of Terms (Continued)

Term	Conceptual definition	Operational definition
Cardiac Magnetic Resonance (CMR)	Known as cardiac MRI, is a medical imaging technology for non-invasive evaluation of the work and structure of the cardiovascular framework. (Lee,2018)	The system that usually used in foreign country to produce a 3D images of heart.
Convolutional Neural Network (CNN)	It is a feed-forward artificial neural networks that use a variety of multilayer perceptrons planned to require minimal preprocessing.(Le Cun,2013)	The technique that used in this study to learn through scanning and image processing.
Deep Learning	Deep learning is part of a broader family of machine learning strategies based on learning data representations, as restricted to task-specific calculations. (Bengio,2015)	The theory that has derived this study to include some machine learning methods.
Pre-process	Subject (data) to preliminary processing (Oxford Dictionary)	It refers to methods like data preparations and segmenting heart images.
Accuracy	The degree to which the result of a measurement, calculation, or specification conforms to the correct value or a standard. (Oxford Dictionary)	The result or outcome obtained in the end of this research regarding its exactness and preciseness.

1.8 Scope of the study

The population used for this study is people from states of Selangor and Kuala Lumpur to represent the whole Malaysia's heart disease data. The hospital of Kuala Lumpur has provided dataset of ultrasound scanned images of heart. However, this dataset may vary a bit for each state due to the environmental factors.

In addition, as this study is only using ultrasound scanned images, it may only able to classify 2D images. As the facility of producing CMR images is not available in the local hospital, we are forced to use this 2D image for this study. Thus, this system is only designed for 2D images of the heart.

1.9 Summary

This chapter generally explains about the objective of this study which is to apply deep learning for heart disease classification system. This chapter also briefly describs the importance of having a system with high accuracy and precision. This study utilized local (Malaysia) data which is collected from the local hospitals.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview

This chapter discusses about the previous studies that have been done on cardiac field and deep learning. These previous researches are important to study and understand how the system or algorithms works to produce an efficient system. Moreover, this literature reviews concludes the previous approaches that have been done on this field.

2.2 Heart Disease

The heart is a strong organ in most creatures, which pumps blood through the blood vessels of the circulatory system. Cardiac structures, see Figure 1, is part in a left and right half, both containing an atrium and ventricle. Left ventricle (LV), left atrium (LA), right ventricle (RV), and right atrium (RA). The right half pumps blood through the lungs for oxygenation, taken after by dispersion through the body by the left half of the heart. The ventricles are dependable for the pumping work of the heart whereas the atria anticipate stasis of venous blood flow amid systole (i.e. withdrawal of the ventricles). A valve is found at each conclusion of the ventricles to anticipate backflow of blood (Gerald,2018). Two surfaces are characterized within the heart. The epicardial surface depicts the external surface of the heart while the endocardial surface refers to the lining on the interior of a chamber. In imaging, the epicardial form usually refers to the external form of the LV myocardial (e.g. muscle) tissue (Myo), though mostly anatomically off base. Three coronary supply routes start straightforwardly after the aortic valve (i.e. the valve between LV and aorta). that navigate the epicardial surface of the heart to supply the myocardium with oxygen and other

supplements. The axis of the heart is defined as the line between the apex and base). Long-axis planes are parallel to this pivot whereas short-axis planes are characterized perpendicular to the axis (Gerald,2018).

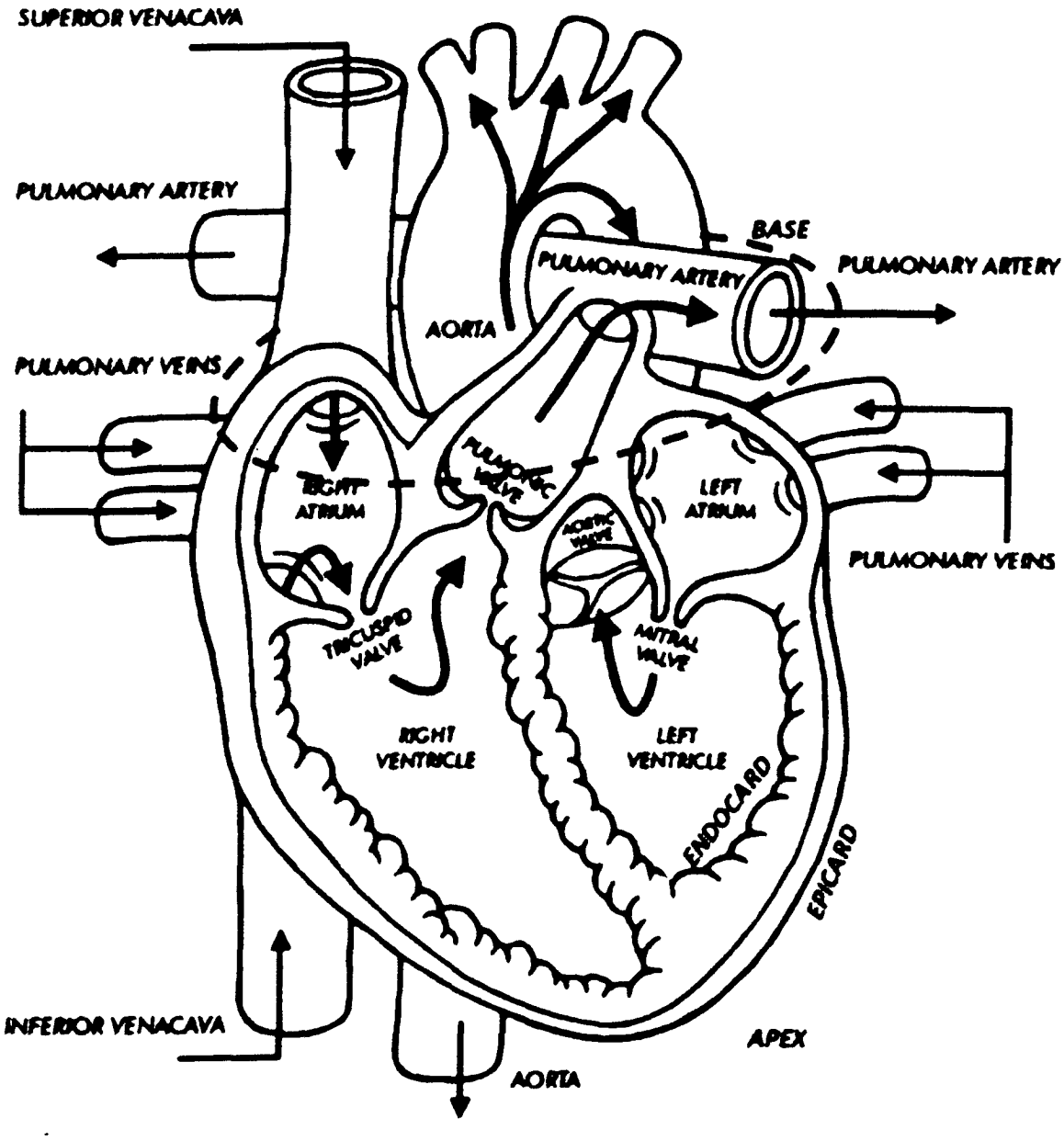

Figure 2.1. Anatomy of heart

Heart illness portrays a range of conditions that influence our heart. Illnesses beneath the heart illness umbrella incorporate blood vessel diseases, such as coronary course disease;

heart beat issues (arrhythmias); and heart absconds you're born with (innate heart absconds), among others. The term "heart disease" is regularly utilized interchangeably with the term "cardiovascular disease." Cardiovascular illness by and large alludes to conditions that include contracted or blocked blood vessels that can lead to a heart attack, chest torment (angina) or stroke. Other heart conditions, such as those that influence your heart's muscle, valves or cadence, too are considered shapes of heart disease.

Administration of cardiac pathologies typically relies on various cardiac imaging modalities, which incorporate echocardiogram, computerized tomography, and magnetic reverberation imaging (MRI). The current gold standard is to use non-invasive cine MRI to quantitatively analyse worldwide and regional cardiac work through the induction of clinical parameters such as ventricular volume, stroke volume, ejection fraction, and myocardial mass. Calculation of these parameters depends upon precise manual outline of endocardial and epicardial forms of the clear out ventricle (LV) and right ventricle (RV) in short-axis stacks (Gerald,2018).

2.2.1 Dilated Cardiomyopathy

Dilated cardiomyopathy (DCM) is the foremost common type, happening for the most part in grown-ups 20 to 60. It influences the heart's ventricles and atria, the lower and upper chambers of the heart, respectively. Frequently the disease begins within the left ventricle, the heart's fundamental pumping chamber. The heart muscle starts to expand, meaning it extends and gets to be thinner. Subsequently, the interior of the chamber broadens. The problem often spreads to the correct ventricle and then to the atria (American heart association, 2018).

As the heart chambers widen, the heart muscle doesn't contract ordinarily and cannot pump blood very well. As the heart gets to be weaker heart failure can happen. Common