AN ADVANCED APPROACH TO PREDICT CARDIOVASCULAR DISEASES USING DEEP LEARNING MODEL

A Thesis submitted

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
IN
COMPUTER SCIENCE AND ENGINEERING

By

Mr. DAMODHARAN D
Registration Number - 17SCSE103018

Supervisor

Dr. AMIT KUMAR GOEL
Professor

GALGOTIAS UNIVERSITY
UTTAR PRADESH, INDIA

July 2022

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis, entitled "**AN ADVANCED APPROACH TO PREDICT CARDIOVASCULAR DISEASES USING DEEP LEARNING MODEL**" in fulfillment of the requirements for the award of the degree of Doctor of Philosophy in the Faculty of Computer Science and Engineering and submitted in Galgotias University, Greater Noida is an authentic record of my own work carried out during a period from January 2018 to July 2022 under the supervision of **Dr. AMIT KUMAR GOEL**.

The matter embodied in this thesis has not been submitted by me for the award of any other degree of this or any other University/Institute.

Mr. DAMODHARAN D

This is to certify that the above statement made by the candidate is correct to the best of our knowledge.

Dr. AMIT KUMAR GOEL
Supervisor
School of Computing Science and Engineering

The Ph.D. Viva-Voice examination of _____ Research Scholar, has been held on _____.

Sign of Supervisor Sign of External Examiner

ABSTRACT

Cardiovascular disease is one of the most common types of chronic disorders that people are facing currently and this is due to the factors like lifestyle and food choices. Spending on medical and public health research is a major economic priority worldwide, especially for disease prediction. When performing routine screenings, echocardiography is the imaging method for evaluating the heart's chambers. But this imaging technique is not fruitful always so it is important to employ machine learning models that are used to predict cardiovascular diseases accurately with a minimum budget and time. The goal of this research work is to identify and create deep-learning models that can reliably detect cardiovascular disease in its earliest stages. This research work aims to develop a health informatics system for the classification and segmentation of heart disorders using machine learning and deep learning methods, with a focus on ultrasonic images. Two proposed models, the AWMYolov4+ method and the KSDSC method, are developed for this work. Both models train a deep learning architecture with an ultrasonic image dataset to improve the classification and segmentation of heart disease.

The KSDSC model used in the field of deep learning is composed of an input layer, hidden layers, and output layer. Initially, ultrasound images are collected by the input layer to be used as a data source for the model's parameters. Once the ultrasound image has been cropped, the Kushner-Stratonovich filter is applied as a preliminary step to further reduce noise. Once the initial stage is completed the preprocessed image is delivered to the next stage. In the second stage of the process, preprocessed images are segmented into a variety of sub-images according to the Sorensen-Dice image segmentation method. Then applies the Haar wavelet transformation into the segmented image to extract numerous features. The output layer receives the extracted features and uses them to implement the softmax activation function to match the extracted features with the disease to predict heart disease. Experimental evaluations of prediction accuracy, false positive rate, and prediction time are carried out on a variety of ultrasound images. Both qualitative and quantitative evidence showing that our proposed KSDSC model outperforms conventional approaches.

The second model, AWMYolov4+, combines the Adaptive Weighted Mean Filter (AWM) with the You Only Look Once Version 4 plus (Yolov4+) model. It has three phases of development, the first phase involves pre-processing the input data

given to the model, the second phase involves splitting up the noise in the images into two different processing paths namely noise detection and noise removal, and the third phase involves combining the results of the two processes. The images with noise are identified and eliminated by this module. Yolov4+ is a well-tested darknet53 networking model, with Mish activation, and it is used in the next phase of the CVD classification procedure. It uses two new classification and segmentation models to detect cardiomyopathy and heart valve disease, which results in a cutting-edge deep neural network methodology (Yolov4+ with Mish activation). The automatic identification of heart chambers in echocardiogram images is based on discriminative deep-learning algorithms. The precision, recall, and F1-score are compared with the alternative methods, in that the model loss value and the prediction time for the proposed method are significantly lower. For region-segmentation purposes, the proposed AWMYolov4+ model outperforms previous classifiers. There is a decrease in cost and an increase in accuracy. Area Under the Curve (AUC) graph values show 95.06 % accuracy for the ultrasound image dataset using the proposed AWMYolov4+ model, with a false positive rate of less than 7%, a short prediction time, and high sensitivity. CAMUS images are collected because they are of a higher quality and accuracy than those found in the dataset.

ACKNOWLEDGEMENT

Working as an Assistant Professor and doing research for the degree of Ph.D. at Galgotias University was a quite magnificent and challenging experience for me. In all these years, many people directly or indirectly contributed to shaping my career. It was hardly possible for me to complete my doctoral work without the precious and invaluable support of these personalities.

I would like to give my small tribute to all those people. Initially, I would express my sincere gratitude to my supervisor Dr. Amit Kumar Goel, Professor, School of Computing Science and Engineering for his valuable guidance, enthusiasm, and over-friendly nature that helped me a lot to complete my research work promptly.

I must owe a special debt of gratitude to Hon'ble Chancellor Mr. Suneel Galgotia, Mr. Dhruv Galgotia, CEO, and Hon'ble Vice-Chancellor Dr. K. Mallikharjuna Babu, Galgotias University for their valuable support throughout my research work.

I express my sincere thanks to Dr. Munish Sabharwal, Dean, School of Computing Science & Engineering, Dr. Sambath Kumar, Research Coordinator, SCSE, and Dr. Amrita Tayagi, Ph.D. Coordinator for their guidance and moral support during my research work and to all faculties of the School of Computing Science & Engineering who helped me a lot in my course of research work and all those who stood behind me.

I express my heartfelt gratitude to my parents Mr. R Dhashanamoorthy and Mrs. Padmini Dhashanmoorthy for their moral support and encouragement. I am happy to remember the prayers and moral support given by my beloved wife Mrs. Kalpana Damodharan and my son Mr. Kishan & my daughter Ms. Dhanshika Shree. I am thankful to my family members, friends, and well-wishers who have been constantly motivating me throughout the progress of this work.

DAMODHARAN D

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABBREVIATION		FULL FORM
AI	-	Artificial Intelligent
AID	-	Artificial Intelligence Decision
AMI	-	Acute Myocardial Infarction
ANNs	-	Artificial Neural Networks
AOR	-	Agent of Record
AP	-	Average Precision
AUC	-	Area Under Curve
AWM	-	Average Weighted Mean
AWMF	-	Adaptive Weighted Median Filter
AWMYolov4+	-	Adaptive Weighted Median Filter-YOLOv4 Plus
BMU	-	Body Mass Index
BoF	-	Bag of freebies
BoS	-	Bag of specials
BPNN	-	Back Propagation Nueral Network
CAMUS	-	Cardiac acquisitions for multi-structure ultrasound Segmentation
CETUS	-	Contrast enhanced thoracic Ultrasound
CIoU	-	Complete Intersection over union
CmBN	-	Cross Mini Batch Normalization procedure
CNNs	-	Convolutional neural networks
CRMPTN	-	Combined Reinforcement Multitask Progressive Time-series Networks
CSP	-	Cross Stage Partial
CT	-	Computed tomography
CVD	-	Cardiovascular Diseases
CV	-	Cardiac volume
DBNs	-	Deep Belief Networks
DCNN	-	Deep Convolutional Neural Network
DIoU	-	Distance Intersection over union
DNN	-	Deep neural networks
DT	-	Decision Tree
DTA	-	Decision-tree algorithm

DWM	-	Dynamic weighted median
DWMF	-	Dynamic weighted median filter
ECGs	-	Electrocardiography (or) electrocardiograms
EF	-	Ejection fraction
EM	-	Expectation Maximization
FCNs	-	Fully Convolution Networks
FN	-	False Negative
FP	-	False Positive
FPN	-	Feature Pyramid network
Fps	-	Frames per second
FPS	-	Frame per second
GANs	-	Generative Adversarial Networks
GBM	-	Glioblastoma Multiforme
GIST	-	Global Information Sharing Toolkit
GPU	-	Graphics processing unit
HIFU	-	High intensity focused ultrasound
IOU	-	Intersection over union
IVUS	-	Intravascular ultrasound
J-RDA	-	Jaya Algorithm with Red Deer Algorithm
KNN	-	K-Nearest Neighbors Algorithm
KS	-	Kushner-Stratonovich
KSDSC	-	Kushner Stratonovich Dice Segmented Haar Wavelet Deep Convolutional Neural Learning
KSFID	-	Kushner-Stratonovich Filtered Image Denoising
LSTMs	-	Long Short-Term Memory Networks
LV	-	Left Ventricle
MAE	-	Mean Absolute Error
MCD	-	Minimal change diseases
MCi	-	Mean Classification accuracy
MED	-	Median
ML	-	Machine Language
MLPs	-	Multilayer Perceptron’s
MOMP	-	Multi-objective mathematical programming
MR	-	Mitral regurgitation
MRIs	-	Magnetic resonance imaging
NB	-	Navy Bayes

NLP	-	Natural Language Processing
NMS	-	Non-Maximum Suppression
NN	-	Neural Network
OOB	-	Out-of-bag
PAC	-	Probably Approximately correct
PANet	-	Path Aggregation Network
PET	-	Positron emission tomography
RAE	-	Relative Absolute Error
RBF	-	Radial Basic Function
RBFNs	-	Radial Basis Function Networks
RBMs	-	Restricted Boltzmann Machines
RCNN	-	Region-based Convolutional Neural Network
RDIoU	-	Rotation Decoupled Intersection over union
ReLUs	-	Rectified Linear Units
RF	-	Radiofrequency
RGB	-	Red Green Blue
RMES	-	Root Mean Squared Error
RNNs	-	Recurrent neural networks
ROC	-	Receiver Operating Characteristic
ROI	-	Return on investment
RRSE	-	Root Relative Squared Error
SAM	-	Spatial Attention Module
SAT	-	Self-Adversarial Training
SDSI	-	Sørensen-Dice Image Segmentation
SOMs	-	Self-Organizing Maps
SPECT	-	Single-photon emission computed tomography
SPP	-	Spatial Pyramid Pooling
SVC	-	Superior vena cava
SVM	-	Support vector machine
TN	-	True Negative
TP	-	True Positive
UCI	-	University of California
US	-	Ultrasound
USI	-	Ultrasound images
VGGNet	-	Visual Geometry Group Network
VGVF	-	Vortical gradient vector flow

WHO	-	World Health Organization
WM	-	Weighted median
WML	-	White matter lesions
XGBoost	-	Extreme Gradient Boosting
Yolo	-	You only look once
Yolov3	-	You only look once version 3
Yolov4+	-	You only look once version 4 plus
ZCRn.	-	Zero Crossing rate
2D	-	Two Dimensional
3D	-	Three Dimensional

LIST OF PUBLICATIONS

INTERNATIONAL JOURNALS

1. Damodharan, D., & Goel, A. K. (2022) "An Implementation of Cardiovascular Disease Prediction in Ultrasonography Images Using AWMYOLOv4 Deep Learning Mode" International Journal on Recent and Innovation Trends in Computing and Communication, vol. 10, no. 9, Sept. 2022, pp. 40-52, doi:10.17762/ijritcc.v10i9.5669**. (Scopus)**
2. Damodharan, D., & Goel, A. K. (2022) "Cardio Vascular Diseases Detection Using Ultrasonic Image by Retaining Deep Learning Model" IJEER 10(3), 639-643. DOI: 10.37391/IJEER.100337. **(Scopus**)
3. Damodharan, D., & Goel, A. K. (2022) "Kushner-Stratonovich Dice Segmented Curvelet (KSDSC) deep convolutional neural learning for heart disease prediction" International Journal of Health Sciences, 6(S5), 5659–5672. https://doi.org/10.53730/ijhs.v6nS5.9881 **(Scopus).**
4. Damodharan, D., Goel, A., Kumar, T. (2019) "An improved enhancement technique for cardiovascular ultrasonic image analysis based on DCNN" International Journal of Advanced Trends in Computer Science and Engineering, 9 (5), pp. 7087-7091. **(Scopus)**

INTERNATIONAL CONFERENCES

1. A. Rai, D. Damodharan., A. S. Singh and K. Anandhan., "Enhanced Machine Learning Model for Design and Evaluation of Various Algorithms Applied in Clinical Cardiovascular Disease," 2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT), Sonepat, India, 2022, pp. 304-310, Doi: 10.1109/CCiCT56684.2022.00063. **(Scopus)**
2. D. Damodharan and A. Kumar Goel, "Comprehensive Study – A Deep Learning & Machine Learning classification Methods for Cardiogram Images" ISCMM 2021. International Symposium on Computer Vision and Machine Intelligence in Medical Image Analysis – ISCMM 2021. **(Scopus)**

3. D. Damodharan and A. Kumar Goel, (2021) "A Novel Approach on classification in Machine Learning model-based USG Cardiogram Images," International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), 2021, pp. 947-950, Doi: 10.1109/ICACITE51222.2021.9404752. **(Scopus)**

BOOK CHAPTER

1. Damodharan, D., Goel, A.K. (2023). Comprehensive Study - A Deep Learning and Machine Learning Classification Methods for Cardiogram Images. In: Gupta, M., Ghatak, S., Gupta, A., Mukherjee, A.L. (eds) Artificial Intelligence on Medical Data. Lecture Notes in Computational Vision and Biomechanics, vol 37. Springer, Singapore. https://doi.org/10.1007/978-981-19-0151-5_1.

Chapter 1

Introduction

Worldwide cardiovascular disease is one of the leading causes of death. World health organizing (WHO) survey reported that cardiovascular diseases (CVDs) are responsible for 32% of deaths worldwide in 2019, estimated at 17.9 million. As per this information, 85% of people have been affected by heart attacks and strokes that led to death [1, 2]. Cardiovascular diseases are conditions affecting the blood vessels of the heart. Four out of every five heart disease-related deaths are caused by stroke and heart attack. Diagnosis of heart disease is essential for fighting CVD disease, which requires analysis of cardiac morphology and functions from medical images. The function of the heart, especially the left ventricle (LV), is highly related to its shape, size, and the load exerted on blood during the myocardium shortening [3]. LV shape and volume changes with time, e.g., ejection fraction (EF) and cardiac volume (CV), are good indicators of some heart diseases like functional mitral regurgitation (MR) and acute myocardial infarction (AMI). To calculate these clinical parameters, the boundary of LV and myocardium need to be identified from low-level images using the techniques such as segmentation and tracking and reconstruct into 3D geometries.

The leading causes of cardiovascular disease are high blood pressure and cholesterol. In some cases, heart disease is caused by environmental factors, tobacco, and alcohol. One of the most significant organs in the human body is the heart. The heart is responsible for the functioning of the entire body. Several risk factors contribute to heart disease, including family history, aging, smoking, high cholesterol levels, high blood pressure, and an unhealthy diet [3]. Blood vessel dilation causes an increase in blood pressure. Cholesterol levels within the body can also cause heart disease because the fatty oxidized cholesterol can adhere to the artery wall, reducing blood flow and eventually leading to death. Another non-modifiable key factor influencing heart health is age. Smoking reduces the amount of oxygen in the bloodstream and causes blood vessels to constrict, accounting for approximately 40% of all CVD-related deaths.

Presently, a physician's intuition and knowledge are commonly used to interpret diagnostic lab tests. However, the medical industry collects massive amounts of patient data that may be used to uncover hidden trends and details to improve decisions and

anticipate problems before they occur [4]. The process of manual decision-making by healthcare professionals frequently results in accidental predispositions, considerable medical expenses, and lapses, all of which have an impact on the type of care patients receive.

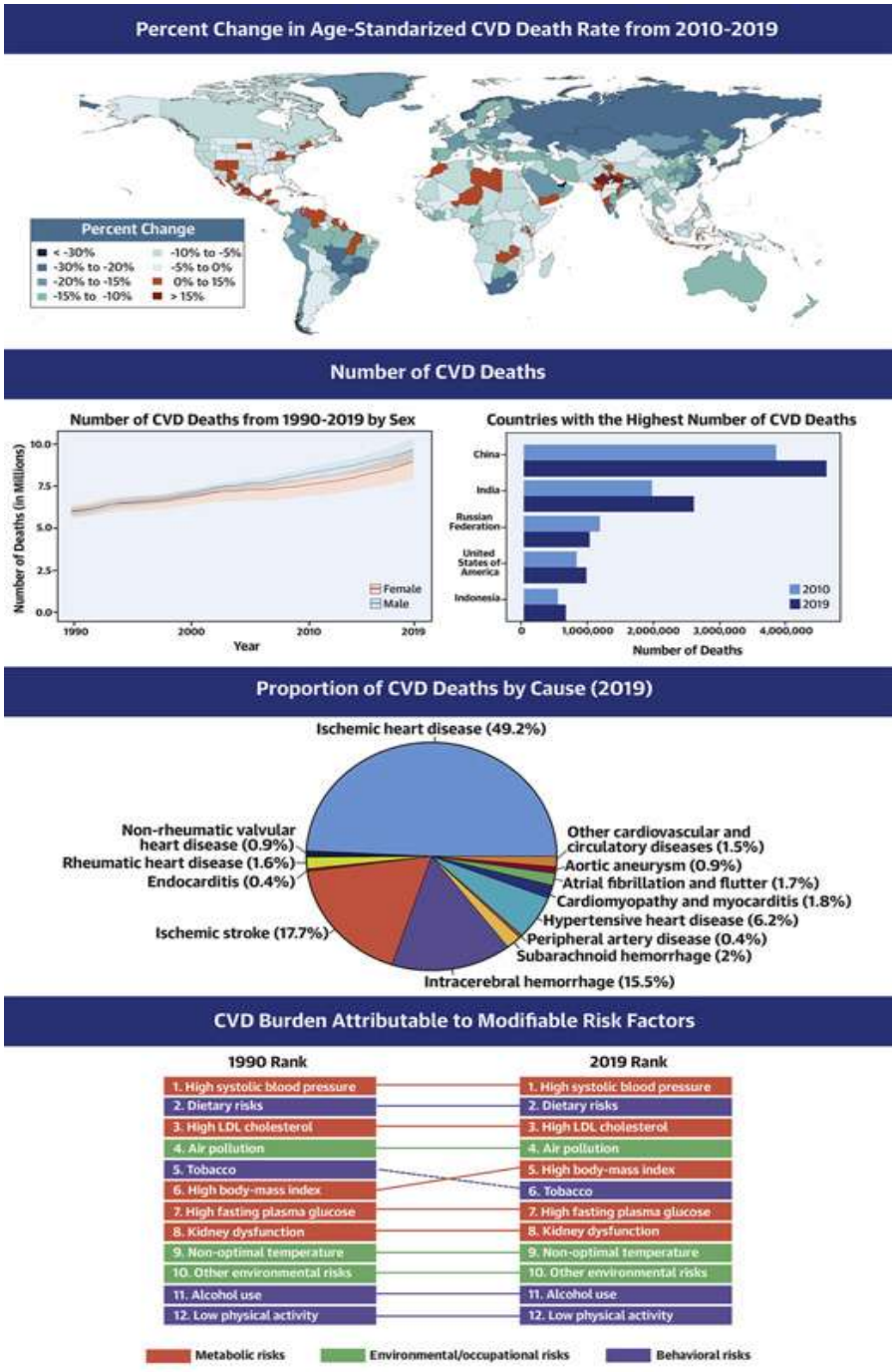

Figure 1.1 WHO Survey for 2019 heart diseases affected worldwide [1]

Figure 1.1 represents the survey report for the WHO in 2019 as the world level CVD spreading and affected country-level reports are shown in the graph and map.

1.1 Functioning of Heart

The heart functions as a piston, with the size of the hand, and sits somewhat to the left of the chest. The side halves of the heart are separated. Figure 1.2 represents the normal heart image view and cut for aortical and ventrical [5].

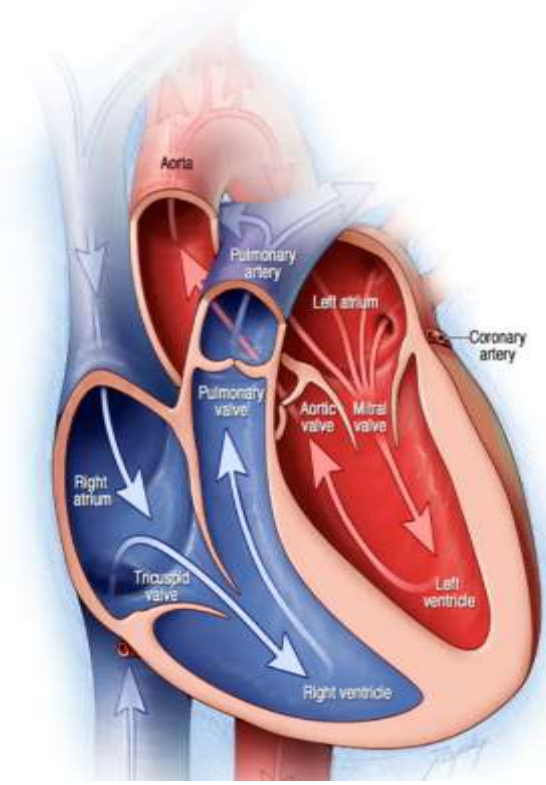

Figure 1.2 A Normal Heart view [5]

- The right aortic valve and ventricular valve are located on the right side of the heart. Via the pulmonary artery to the lungs, it gathers and pushes blood to the lungs.
- The lungs replenish the oxygenated blood. Co2, a by-product, is exhaled by the lungs.
- Oxygenated blood then flows into the heart's left side, along with the left chambers (atria & ventricles).
- The left atrium circulates blood via the aorta, the body's biggest arterial, to provide oxygenated blood to tissue cells.

1.1.1 About Cardio Vascular Diseases

Heart and blood vessel disorders are commonly referred to as CVDs. Some of them are

- CORONARY CVD
- CEREBROVASCULAR CVD
- PERIPHERAL ARTERIAL CVD
- RHEUMATIC CVD
- CONGENITAL CVD
- DEEP VEIN THROMBOSIS AND PULMONARY EMBOLISM

Cardiovascular diseases are generally sudden occurrences triggered by an obstruction that stops blood to flow to the heart or the skull [6]. The growth of fatty plaques on the interior lining of the circulation arteries that supply the blood to the heart or brain is the most prevalent cause. And also affected as brain by strokes are caused by blood blockages or haemorrhages from a blood artery in the brain.

1.1.2 Common symptoms of CVDs

The fundamental illness of the coronary arteries often has no signs. A cardiac arrest event might be the initial symptom of a more serious condition. The person may have breath-inheld issues or breathlessness, stomach problems, light-headedness or dizziness, chilly perspiration, and paleness [7]. Breathlessness, vertigo, puking, and backbone or mouth discomfort are much more prevalent in females than in males. Immediate paralysis of the face, arm, or leg, usually on one side of the body, is the most typical sign of a stroke.

Some signs are an immediate and unexpected presence of

- numbness of the face, arm, or leg, especially on one side of the body.
- ambiguity, speech problems, or comprehension of voice.
- trouble noticing by one eye or both.
- struggle moving, light-headedness, and/or loss of coordination.
- extreme pain in the head for an unknown reason.
- drowsiness or loss of consciousness.

Those who suffer from these signs should visit a doctor right once.

1.1.3 Artery disease-causing factors

The much more incident of coronary artery disease is the development of lipid hardening of the arteries. Atherosclerosis is associated with poor lifestyle factors such as a bad diet, insufficient physical activity, being obese, and smoking.

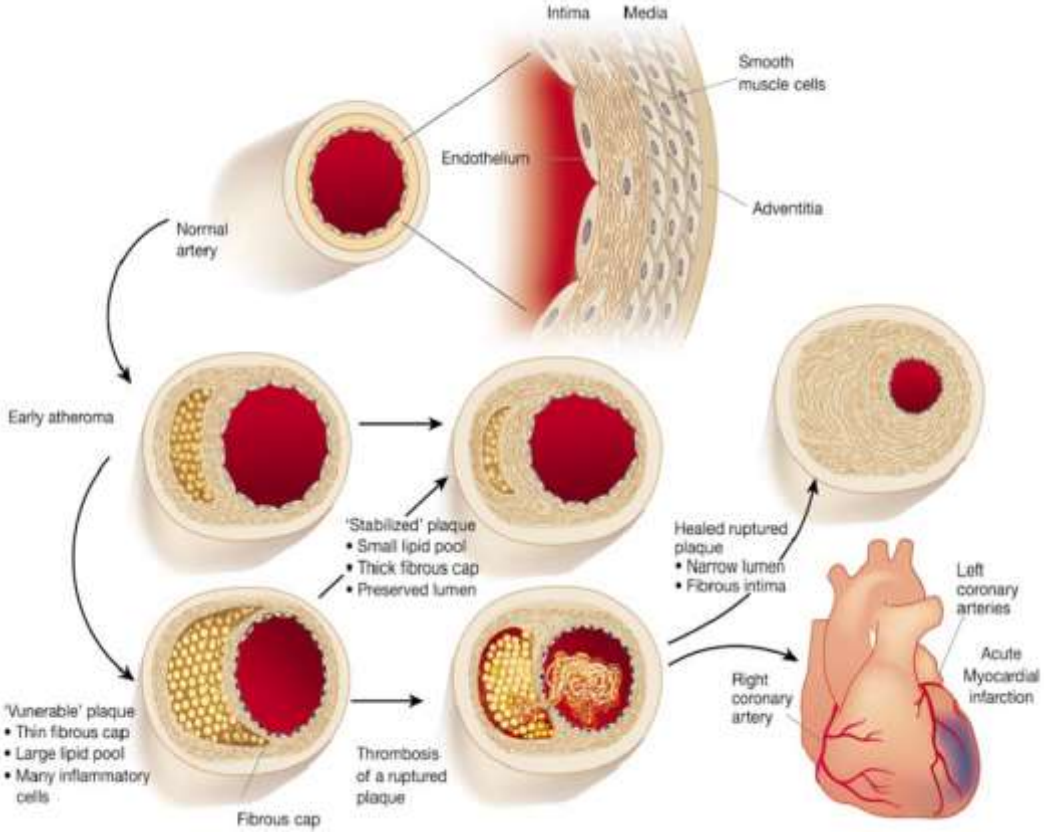

Figure 1.3 Inflammation in Atherosclerosis [8]

Figure 1.3 denotes blood vessels' bad cholesterol can cause heart disease. Plaque ruptures cause blood clots. Plaques and clots can block arteries. The illustration above shows the evolution of atherosclerotic plaques and suggests two different types of evolution. Slowly expanding plaques grow over time as a result of lipid build-up in foam cells, smooth muscle cell migration, and proliferation. These plaques don't easily rupture and tend to stabilize. The fatty plaques create due to improper food habits, lack of exercise, cigarette usage, and hazardous binge drinking are the major lifestyle risk factors for CVDs and stroke. People might well have symptoms such as hypertension, hyperglycaemia, high blood lipids, and weight gain as a result of lifestyle factors [8]. It is critical to diagnose cardiac illness as quickly as practicable so that treatment may start with counseling and medications. CVDs are indeed affected by a combination of

contributing causes [9]. Unemployment, anxiety, and genetic reasons are additional predictors of CVDs.

1.2 Medical imaging overview

Anatomical images can be generated using medical imaging techniques in living being [10]. Organs and tissues are revealed structurally and functionally through the generated images that help in clinical diagnoses and interventions. In most cases, medical imaging does not require any invasive procedures. Therefore, there should not be any cuts or openings in the skin of the patients [11]. This thesis frequently uses ultrasonography (ultrasound) images as a method of medical imaging.

1.2.1 Imaging modalities

Ultrasound imaging techniques have, no doubt, demonstrated their capability in the measurement of adipose tissue deposits. These measurements are indeed accurate and precise. Particularly, two-dimensional (2D) echocardiography has been recognized, globally, as the safest and most accurate imaging technique for the diagnosis of cardiac issues [11]. Ultrasound (US) imaging has been in use for a long time, and it is prevalent due to the various advantages it offers over other imaging modalities. The ultrasound imaging technique is easy to use, safe, non-invasive, inexpensive, and easily portable.

Ultrasonography produces images in 2D or 3D by using high-frequency sound waves. Ultrasound probes generate and transmit pulses of ultrasound waves into the body. Interactions between tissues absorb sound waves. Ultrasound probes record the reflected soundwaves and reconstruct images [12]. Ultrasound imaging is widely used to examine unborn children, abdominal organs, and the heart of pregnant women since it is non-harmful to the mother and has a low cost. Real-time ultrasound images are also possible. As a result of noise, artifacts, and shadows, ultrasound images are typically poor in quality. This ultrasound image is shown in figure 1.4.

Echocardiography, Ultrasound imaging of the heart has been proven to provide an excellent way to accurately measure the amount of adipose tissue deposited around the heart. This technique provides visual information about the tissues based on the reflections of the ultrasonic sound waves that are transmitted to the patient using a transducer probe [13]. Various studies have demonstrated the use of echocardiography in tissue characterization based on visual data. A study investigated the texture analysis of echocardiographic images to identify or characterize myocardium and has

successfully been able to identify myocardium using a wavelet transform to analyze the tissue texture in the visual data [14-16]. However, visual data analysis methods are time-consuming and are limited to offline processing only. Moreover, visual data obtained from the reflection of the ultrasonic sound waves only contain the intensity information related to the tissue structure which can be determined from the amplitude of the reflected signal.

Table 1.1 Properties of soft tissue within the body as well as the properties of various ultrasound Images

Serial number	Medium	Density (kg/m^3)	Speed of ultrasound (m/s)	Acoustic Impedance (kg/ (m^2. s))
1	Air	1.3	330	429
2	Water	1000	1500	1.5 x 10^6
3	Blood	1060	1570	1.66 x 10^6
4	Fat	925	1450	1.34 x 10^6
5	Muscle(average)	1075	1590	1.70 x 10^6
6	Bone(varies)	1400-1900	4080	5.7 x 10^6 to 7.8 x 10^6
7	Barium titanate (transducer material)	5600	5500	30.8 x 10^6

Table 1.1 denotes the echocardiography testing method to find the issues in the body and the ability to access the frequency information of the images along with the visual data. Many works on tissue characterization have demonstrated the potential frequency content in their respective studies. Apart from the regular B-mode images, the radiofrequency (RF) data that can be obtained during echocardiography has been known to contain information related to the tissue microstructure [17]. In addition to access to the information related to tissue structure properties, the frequency data also helps in classifying the tissue types with a higher resolution than conventional B-mode images [18-20].

Over the years, many studies have emerged that investigated the use of ultrasonic frequency data in tissue characterization. These studies have used spectral analysis of the RF data to identify the information related to the targeted tissue types. Three different studies investigated the use of ultrasound integrated backscatter in segregating the myocardium and have suggested the tissue structure of the myocardium changes due to pathologies that reflect in the integrated backscatter [21-23]. These studies have been able to successfully identify the morphological changes within the myocardium using RF backscatter data. Various other studies on myocardial damage due to toxicity, myocardial infarction, and myocardial viability have also shown the potential of ultrasonic tissue characterization techniques [24]. One particular study investigated the use of ultrasound tissue characterization techniques to identify an abnormal increase in the myocardial wall echo density in diabetic patients [25].

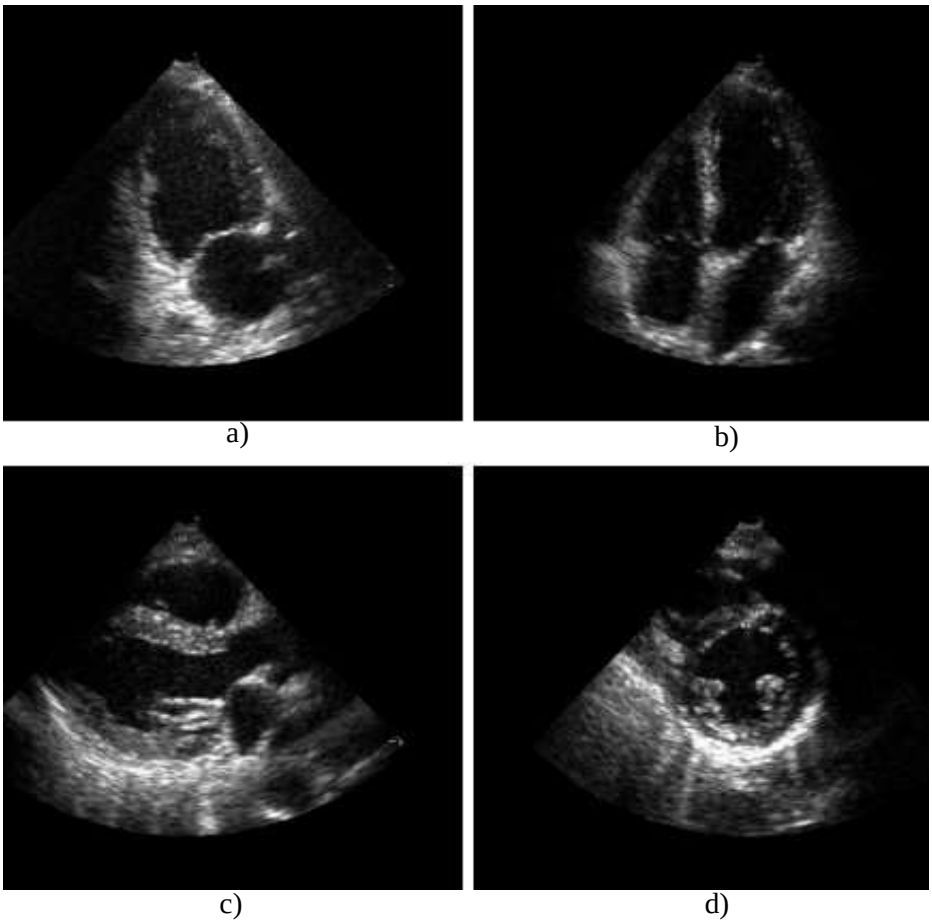

Figure 1.4 An ultrasound heart images (a) long axis 2-chamber (b) long axis 4-chamber (c) Short-axis base (d) Short-axis mid [26]

Figure 1.4 represents the ultrasonic heart images in the different views. Apart from the identification of myocardium, numerous studies have emerged that investigated the use of tissue characterization in identifying or diagnosing different diseases related to the heart. Studies on carotid plaques, atherosclerosis, HIFU cardiac lesions, vulnerable plaques, thin cap fibroatheromas, and calcification have been successful in demonstrating the potential of ultrasound radio frequency data in tissue characterization [26, 27]. Numerous studies have also been conducted on the characterization of coronary plaques using intravascular ultrasound (IVUS) backscatter data [28-31]. These studies have suggested that IVUS data is important in identifying the different types of coronary plaques with improved resolution and provides an accurate measurement of the thickness of the plaques. Studies have also investigated the use of RF data in detecting non-cardiac tissues and pathologies such as fatty liver disease, skin lesions, and intercostal blood vessels [32]. All these studies have also signified the potential of ultrasonic tissue characterization techniques. Although there has been a lot of research in identifying or characterizing myocardium and myocardial damage using ultrasound techniques, studies of the characterization of CVD using backscatter data are limited. The study presented here aims to characterize CVD and other cardiac tissues using ultrasound tissue characterization techniques to help diagnose CVD-related health risk conditions.

1.2.2 2D Echocardiography

Cardiologists commonly use two-dimensional echocardiography (2D echo) to evaluate heart disease among other imaging techniques. Two-dimensional echo is the most widely used non-invasive method because of the following advantages. Initially, 2D echo generates high temporal resolution images (50-250 frames per second (fps)) which contain more accurate LV motion information within the heart cycle [33]. Also, 2D echo needs a short acquisition time which allows real-time analysis [34]. Standardized scanning locations for both long-axis and short-axis that can act as the criteria planes facilitate the clinical analysis from the 2D echo images [35]. Finally, 2D echo is relatively inexpensive and does not need advanced preparation for the testing. 2D echo provides a grayscale image with anatomical features, e.g., LV cavity, LV myocardium, and heart valves. Calculating cardiac functions requires these variables. The first step to quantification of the cardiac functions from the 2D echo images is the segmentation of the left view of the heart starting from the grayscale 2D echo images.

Various studies that involve ultrasonic tissue characterization have also involved the use of different spectral analysis techniques. Traditionally, Fourier transform techniques were used to compute the power spectrum of the backscatter data [36, 37].

One weakness is the trouble in locating a location on the patient's body from which the ultrasound waves from the probe can reach the heart uninterrupted. Ultrasound waves can be obstructed by objects such as bones or gases. The heart is normally visible from places such as between the ribs, beneath the breastbone, and above the abdomen. These locations where the heart can be seen result in images of the heart from various perspectives and with different heart structures visible. These views have specific names and are referred to as standard echocardiographic points of view. Each ultrasound heart image falls into one of these categories such Parasternal long/short axis (PLAX or PSAX), Apical four/five/two chambers, Apical long axis, subcostal four-chamber, and Suprasternal view.

In this study, multiple sub-views at various levels in the parasternal short axis mid-LV classes are included. If necessary, these points of view could be classified separately. These images are created using ultrasound machines, which are manufactured and sold by various manufacturers. Even though they all work in the same way and have similar features the images they produce can differ. They can differ in image quality, pixel density, and graphical interface elements, in addition to the presence of patient information in the images. Some ultrasound machines incorporate patient data into the images, while others do not. These devices can extract single images or image sequences. Normally, these images are captured and visible on the ultrasound machine's screen. They can be monochromatic or color images [38]. Because they record what is displayed on the screen at any given time, the images may contain other graphical elements related to different measurements made on the image when exported. Color doppler measurements of blood flow, for example, are common. The doppler effect is used in this case to measure the velocity of the blood relative to the ultrasound probe and is displayed in images with different colors.

An expert in doppler ultrasound can perform several measurements required for diagnosing any heart problems. Instances of these measures include the interventricular septum thickness, the diameter of the left ventricular outflow tube, the diameter of the sinuses of valsalva, the blood flows, and the volume of the various heart chambers. The core principle behind ultrasound imaging is that most waves are taken in by the body, and some waves are reflected at the boundaries between various tissue densities. The

transducer not only produces ultrasonic waves but also detects reflected waves, processes the information, and shows it as images. Therefore, the prospect of automating any such measures would be made possible by the ability to automatically classify echocardiographic images. This would make the ultrasound imaging modality an even more helpful diagnostic tool by lowering the amount of expert knowledge required to interpret and evaluate ultrasound cardiac pictures.

Other imaging methods namely Magnetic resonance imaging (MRI), Positron emission tomography (PET), single-photon emission computed tomography (SPECT), echocardiography, endoscopy, tactile imaging, computed tomography, X-ray, and optical imaging are among the other medical imaging modalities used in diagnosis and intervention. Because these techniques are not used in this thesis, they are not discussed in depth here.

1.3 Machine learning-based medical images

In healthcare, machine-learning applications can be classified based on clinical activities and also created based on data types. Demographic data, medical records, pathology slides, skin lesions, retinal photographs, electrocardiograms, vital signs, electronic recordings, physical examinations, and images from clinical laboratories are some examples of the types of data that can be collected. A variety of data types are used to train AI methods, which are typically compared with physicians' assessments to verify their suitability. The Receiver Operating Characteristic (ROC) is calculated by plotting the true-positive versus false-positive rates and is used to determine the area under the curve [39]. The purpose of this section is to provide some examples of how AI is being used in healthcare. Consequently, the following sections are arranged by types of data and are divided into different applications of AI. The AI healthcare literature is most interested in biomedical images as one of the forms of medical data, due to the large volume of medical imaging scans collected in the Patient portal systems of healthcare institutions. The second component relates to the development of computer vision algorithms, particularly convolutional neural networks.

The major challenge that healthcare organizations facing currently is the rendering of high-quality services at a reasonable price. Quality service entails appropriate diagnosis of the disease of patients and dispensing efficacious therapies. Unqualified medical assessment can have disastrous effects, and therefore cannot be tolerated. In hospitals, clinical testing expenses are reduced by Incorporating machine

learning and deep learning into the healthcare industry's data analytics and prediction systems [40-44]. So, designing a system that could diagnose & predict future coronary disease risk becomes imperative to help people at risk before the risk becomes fatal. Various methods, including Decision Trees, KNN, Neural Networks, and Nave Bayes, have been used to predict heart disease. KNN comprises a range of K values to determine the values of cardiovascular disease risk factors. SVM is used to calculate the svc scores for different SVM kernels to find the most suitable kernel for the necessary prediction [45-48]. A decision tree algorithm is used to calculate the efficiency at a different number of maximum features selected for classification for CVD prediction. Random Forest is a technique that is used to calculate the precision with a different number of estimators.

The Naive Bayes method is used to predict heart disease through probability. XGBoost is a rather new algorithm and is used to calculate the scores for the different number of estimators. In all the above-mentioned techniques and methodologies, the patient records are classified and predicted continuously. Changes in different attributes will affect the outcome and future risk of coronary disease, if the risk comes to be positive, the patient can be informed and necessary steps can be taken to avoid a risk of death due to CVDs. In this approach, clinicians can predict heart illness at an early stage and use machine learning to facilitate and improve the process [49]. This research work gave about different machine learning and deep learning algorithms and compares them in different ways to find out which one gives us the most accurate and efficient results.

Using the huge amount of data available in various formats, such as medical images, electrical signals, and digital health records, artificial intelligence systems could also be able to make accurate and timely diagnoses. Although computer vision applications can be useful in the healthcare industry, medical images are not the primary tool for diagnosing cardiovascular disease [50]. In particular present research on the use of MRI for the diagnosis of chronic myocardial infarction. In addition, the applications of electrocardiography (ECG) are discussed before the discussion of biomedical electrical signals. Interestingly, elevated blood pressure, glucose, lipids, and other signs like overweight and obesity could typically be displayed in electronic health records as a diagnosis of CVDs. This has led to the development of various prediction systems by researchers using these backgrounds.

Automated medical image analysis has been using machine learning techniques for decades. Historically, computational resources were limited. The site also contains

a few images. Premium quality attributes, like scale-invariant feature transforms and key points, were used to train machine learning models [51, 52]. In supervised and weakly supervised learning, this same method of assessment results (such as labels or images), while in unsupervised learning, the models produced intrusion within the data themselves.

1.4 Machine learning

There are three primary benefits to automatic medical image analysis. To begin, automated algorithms, especially those developed on deep learning technologies like CNNs can process images much more quickly than humans might. Example. the proposed models KSDSC and AWMYOLOv4+ are used in the research instantly segmenting myocardial infarction (mi) lesions and taking a few seconds to run [53, 48]. Further, in difficult segmentation problems like WML segmentation in CT images, inter- and intra-raster consistency of human experts can be low [54]. Segmentations provided by a panel of specialists can be used to derive automated methods, which are particularly useful for low signal-to-noise ratio or artifact-ridden images. Therefore, automated methods may provide more trustworthy results than human experts. The final point is that expert annotating of medical images requires a lot of time and money, while automated analysis can be scaled up. Automatic medical image analysis has made use of machine learning techniques for many years [55]. Until recently, there were only a small number of computers available. In addition, only a handful of pictures have been released so far. The collection of images, annotations, and computational resources has increased significantly in recent years. Machine learning models, particularly deep neural networks, may take image features as input. The performance is considerably enhanced by the deep neural network's ability to learn visual characteristics and subsequent tasks in their entirety.

1.4.1 Introduction

Machine learning, classification, and segmentation networks are concepts and terms covered in this chapter. Research on heart disease prediction using various machine learning approaches such as Nave Bayes (KNN), Decision Tree (SVM), Yolo, etc. is examined, focusing on what was done, how it was done, classification technique employed, the data set required for implementation, and the tools applied.

Instruction data is provided to machines to start the process. Algorithms can accept any data and operate on it without our assistance once these models have been established [56]. The machine could eventually learn to distinguish unlabelled data.

Approximating a function, such as f (), through the use of computational models M and empirical data X is the focus of machine learning (ML). The empirical data X = x_1, x_2, ..., x_m is made up of m instances, and each instance x_i = x_1, x_2, ..., x_n, i = 1, 2, ..., m, can be thought of as a vector of n feature attributes. In most cases, the target activity is extraordinarily subtle and intricate. ML methods can be broken down into two groups, supervised learning, and unsupervised learning, depending on whether or not the training data have associated labels Y available. While Y = y_1, y_2,..., y_m in supervised learning, Y = X in unsupervised learning indicates that the data is unlabelled. Weakly supervised learning refers to a learning process in which only some of the data (mp/m instances) are annotated as Y = y_1, y_2, ..., y_{mp}. Approximating the function f: X 7 Y is a common goal in machine learning.

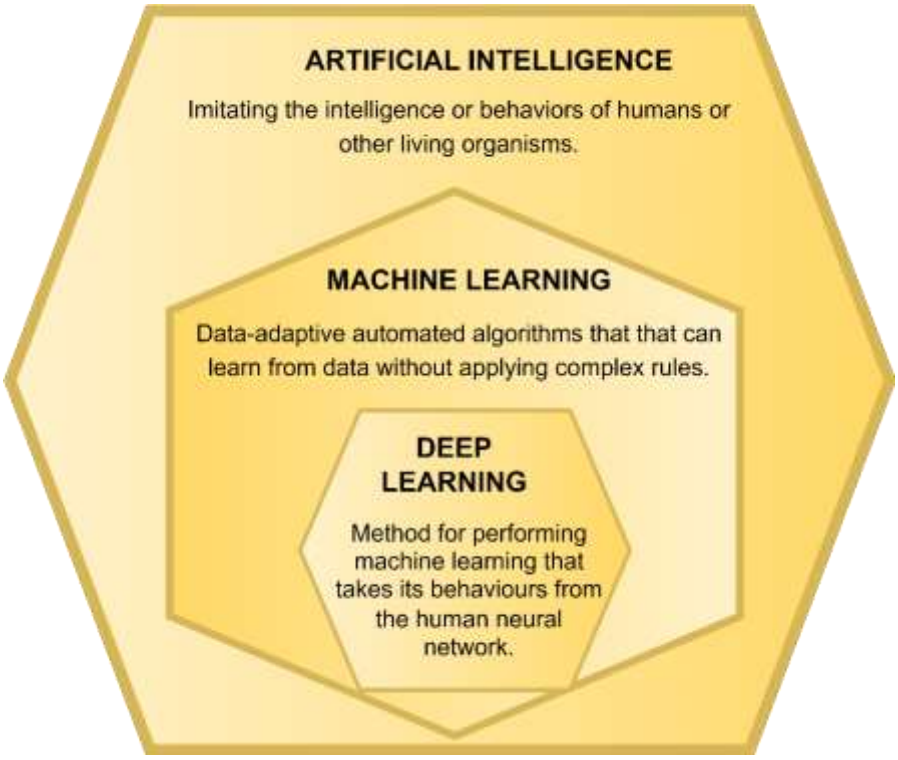

Figure 1.5 Relation between AI, Machine learning & Deep Learning model

Figure 1.5 represent the subset of artificial intelligence is machine learning that is based on deep learning. The concept of the outer layer AI was the first to emerge in this field, and the middle layer flourished later the AI and inner layer the expansion of the middle layer.

A. Methods of machine learning

Machine learning is divided into three categories namely reinforcement learning, unsupervised learning, and supervised learning. Learn more about them now. Machine learning is a broad and constantly emerging field. Machine learning keeps subdividing more and more into subfields [57, 58]. In the 21st century, ML is a crucial component of any successful business or study. Using algorithms and neural network models can help software programs gradually improve their functionality with minimal human intervention.

Computers can learn in the same way that people can't do anything useful when they are first born and have to be taught how to do things. To give a simple example, if you want a computer program to learn how to tell if an animal is a dog or a cat, you would show it a lot of pictures of dogs and cats. There is hope that in the future, the computer system will have the ability to apply statistical models built on prior data to distinguish between cats and dogs when looking at a picture. Algorithms used in machine learning are frequently divided into several categories.

Machine learning can be broadly classified into four broad categories

- Supervised ML
- Unsupervised ML
- Semi-supervised ML
- Reinforcement ML

1.4.2 Supervised machine learning

Classification and regression are the main challenges in supervised learning. while y_i, i = 1, 2..., m is typically continuous in regression problems, it is discrete and categorical in classification tasks. Clustering and dimensionality reduction are the primary uses for unsupervised learning. The result of the learning process in clustering problems is a collection of clusters. in image compression problems, however, the classification of the original data instances in a lower dimensional space is the result of

the learning process. Using unlabelled data to strengthen supervised learning is the goal of weakly supervised learning [59-61]. The purpose of reinforcement learning is to learn to maximize rewards with the help of an optimal policy. The problem with policy learning is that there are no "instance-label" pairs.

A classification of topics into medical specialties of image pixels into different classes according to tissue types or organs represents the most common but also effective approach in medical image processing. Supervised learning algorithms can be used when every image is labeled [62]. But in classification problems, there are a lot of images that are not labeled because the costs of achieving this are too high. To make use of these unannotated images, weakly supervised techniques are used.

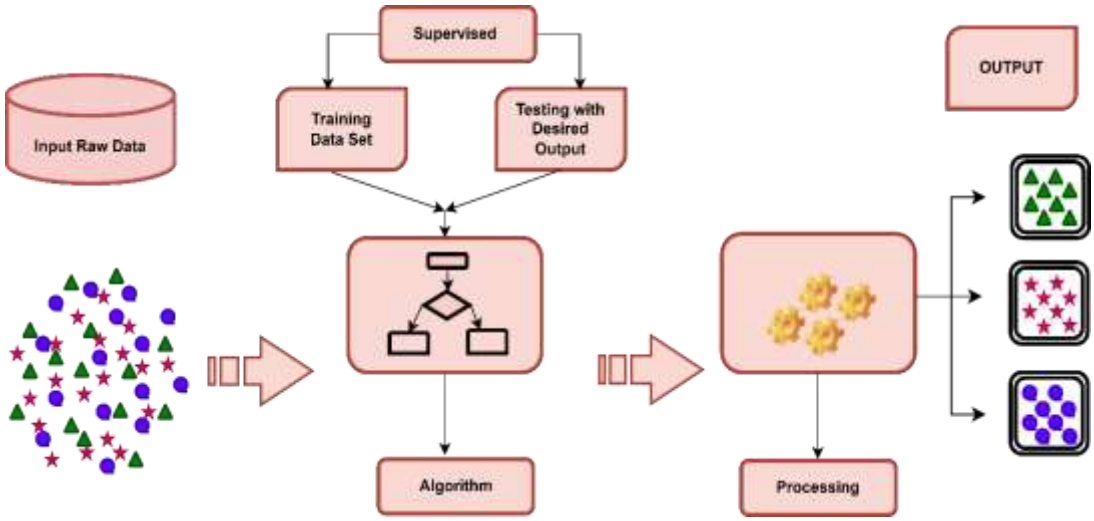

Figure 1.6 Supervised learning process work flow diagram

Figure 1.6 shows the symbols classified based on the input data and label the machine learning can train the supervised algorithm and predicted the result. This kind of algorithm can predict what will happen in the future by comparing information it has already learned with information it has just received. As a result, the system does not learn how to arrive at specific results; rather, it learns the process by which certain results might be obtained. For example, in supervised learning, you can perform tasks like regression and classification. An observation or piece of data can be classified by assigning it to one of several possible groups. Classification makes use of historical information and observations to determine the best organizational scheme for newly acquired facts [63]. A set of labels are applied to the new data. A good illustration of this is the task of identifying the species depicted in an image (a dog, a cat, etc.). Medical diagnoses, targeted advertising, and targeted banking are just some examples of current

uses. Predicting a continuous numerical value is the goal of regression models. Regression is often used to predict the price of a car based on factors like horsepower, size, etc. A variety of supervised machine learning methods, such as polynomial and random forest models as well as decision trees and logistic models are all examples of this sort of regression.

A. Classification of learners based on their learning styles

Lazy learners do not actively process the training data; instead, they save it until they're later given a test on what they've learned. The classification process uses only the most important information from the training data. Comparatively, they have a better sense of timing than excited learners. k-nearest neighbour and case-based reasoning are some examples. Eager learners develop a classification model using the training data before receiving data for predictions, which is a standard procedure. One hypothesis that works across the board must be able to commit to the project as a whole. These models require a lot of practice before they can make an accurate forecast, thus training takes a while. For example, a decision tree, a naive Bayes model, or an artificial neural network.

In machine learning, classification refers to the use of supervised learning techniques to divide a dataset into distinct categories. Common examples of classification problems include speech recognition, facial recognition, handwriting recognition, document categorization, and others [62]. Depending on the circumstances, a "binary classification problem" or a "multi-class problem" can be used. It is possible to use a variety of machine learning methods to classify data in machine learning.

1.4.3 Logistic regression

Using one or more independent variables, this machine learning classification technique determines the output. A simple binary variable is used to evaluate the progress, so there can be only two values either true or false, yes or no, 0 or 1. Finding the optimal relationship between a dependent variable and a group of independent variables is the focus of logistic regression. Comparatively speaking, it performs better than "nearest neighbour" and other binary classification algorithms because it provides numerical estimates of the contributing factors to classification [63]. The logistic regression and the naive Bayes classifier are two examples of linear as well as probabilistic models that are used as building blocks in supervised learning algorithms

for classification tasks [64]. In the context of linear modeling, z is expected to be a linear function of x_i.

$$Z = W^T X_j + b \tag{1.1}$$

In this equation (1.1) case, w and b are the linear model's trained parameters. Using the equipment step response, the predicted variable z could be transformed into a nominal value y_i.

$$y_i = \begin{cases} 0, & z < 0; \\ 1, & z \geq 0, \end{cases} \tag{1.2}$$

The transfer function, moreover, is discontinuous. Since this is the case, the logistic function is frequently substituted for the traditional step function.

$$y_i = \frac{1}{1+e^{-1}} \tag{1.3}$$

Logistic regression attempts to predict the probabilities of each class rather than the classes themselves, which can be more useful in many situations. This linear model discriminates between X and Y, directly estimating their posterior probability P(Y | X). Conceptual models can also be used for this purpose. Estimating the joint distribution of X and Y, P(X, Y), is a common goal of unsupervised learning.

$$P(X|Y) = \frac{P(X,Y)}{P(X)} \tag{1.4}$$

A. According to Bayes' theorem,

$$P(X|Y) = \frac{P(Y)P(X|Y)}{P(X)} \tag{1.5}$$

The preliminary occurrence of Y is denoted by P(Y), the likelihood of X being a member of the Y class by P(X | Y), and similarity observations by P(X). However, estimating the likelihood of P(X | Y) from insufficient training instances is challenging [64]. Because of this, the naive Bayes classifier treats all characterizes as independent if and only if Y is known. Therefore,

$$P(X|Y) = \frac{P(Y)}{P(X)} \tag{1.6}$$

Classification is a common task for logistic regression, and it can help you investigate the impact of a set of independent variables on a dependent one. The logistic regression algorithm has one major flaw. It can only be used to make predictions when the variable in question takes on the values 1 or 0. It also assumes that there are no missing data in the records and that the predictors don't affect each other.

1.4.4 K-nearest neighbour

It is a lazy learning algorithm that stores all training data instances in a space with n dimensions. It's a lazy learning algorithm because it doesn't try to build a general internal model. Instead, it stores examples of training data. Classification is done by taking the simple majority of the votes of each point's k nearest neighbours. It is watched over and uses a bunch of points that have already been labeled to classify other points. To label, a new point looks at the points that have already been labeled that are closest to it. These are called their nearest neighbours. It lets the neighbours vote, and the new point gets the label that most neighbours have. The number "k" tells how many neighbours it checks. This algorithm is easy to use, and it works well even with noisy training data [65]. Even if there is a lot of data to train on, it works well. The only bad thing about the KNN algorithm is that the user has to figure out what the value of K is, and it takes a long time to run when compared to other algorithms.

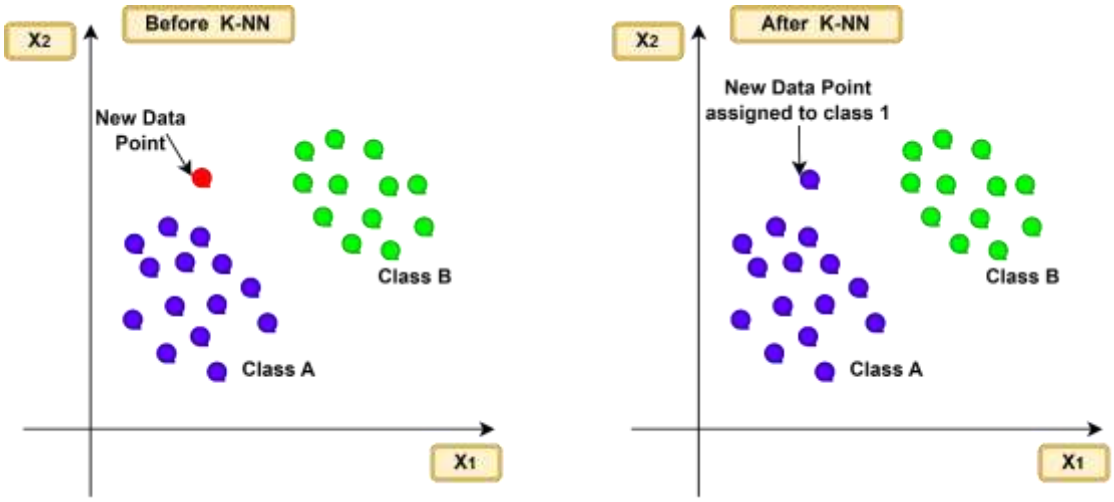

Figure 1.7 (a) KNN before classification & (b) After KNN classification methodology

Figure 1.7 (a) represents the two classes to analyze and calculate the new point belonging before the KNN classifier. (b) represents the KNN classifier's calculation and

identification of the new point's membership in class A. x1, and x2 labels are located to calculate the distance based on the new data point.

The KNN algorithm is a lazy way to learn that doesn't use parameters. The K-nearest neighbour method is used to predict the values of new data points based on how closely they seem to be related to points in the training dataset. It also means that the amount of the data point being looked at will depend on how closely it is related to the points in the training dataset [66, 67]. Classifier looks for the classes of a data point's K closest neighbours and gives that point a class based on which class is most common. On the other hand, the number of neighbours can be changed. The test score was determined in each situation where the number of neighbours was increased or decreased. Then, we looked at the different results for each value of K.

1.4.5 Support Vector Machine

Support vector machine (SVM) and other classification methods can use a maximally wide gap to represent training data as points in space divided into categories. This is done by anticipating which category each point falls into and which space it belongs to. The decision function only uses a small number of training points, so it is easy to recollect and works well in high-dimensional spaces [68, 69]. The only drawback to using a support vector machine is that the underlying method does not provide precise confidence intervals. SVMs are powerful and flexible supervised machine learning (ML) techniques [62]. SVMs have a different way of being put into action than other ML algorithms. SVM model can be thought of as the descriptions of different classes in a hyperplane, which is a multidimensional space. To reduce errors, SVM creates the hyperplane by going through a series of steps. By dividing datasets into classes, SVM tries to find the highest peripheral support vectors.

A Kernel is used in SVM to transform the input data space into a structured sequence. A "kernel trick" is used by SVM to convert a low-dimensional input space into a higher-dimensional area. This improves the robustness, adaptability, and precision of SVM [68]. The SVM kernels utilized in our research are as follows

- **Linear:** Any two observations can be utilized as a dot product using linear kernel. The linear kernel formula is as follows

$$X(l, l_i) = sum(l * l_i) \tag{1.7}$$

It can be seen that the sum of the multiplication of each pair of input values is the product of two vectors, say $l\&l_i$.

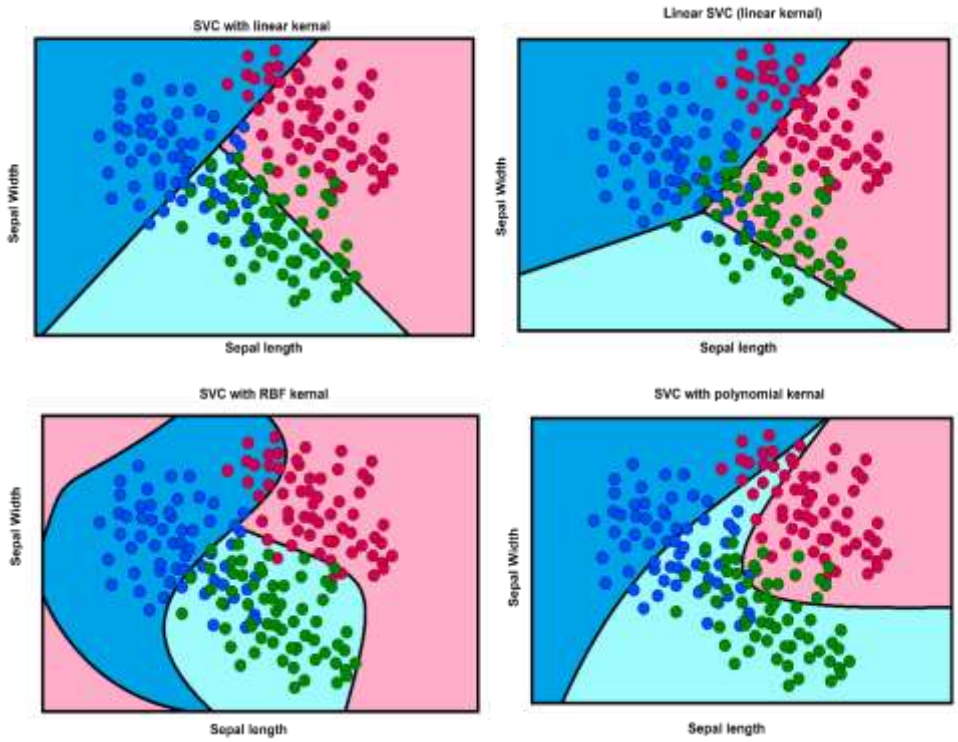

Figure 1.8 Classification using different SVC kernels

Figure 1.8 shows the four SVM classifiers with different kernels, representing the classifiers for SVC, Linear-SVC, Radial basis with SVM kernel, and polynomial SVM based on axis sepal width and length. The image has three features namely blue, green, and pink. Based on the feature with kernels the points are separated.

- **Polynomial:** A polynomial kernel is a more extended version of a linear kernel that can discriminate between curved and nonlinear input spaces. The formula for the polynomial kernel is as follows

$$x(K, K_i) = 1 + sum(K * K_i)a \tag{1.8}$$

Here a is the polynomial degree, which we must explicitly define in the learning method.

- **Radial Basic Function:** This kernel maps input space in indefinite dimensional space. The formula for RBF Kernel is below.

$$X(k, k_i) = exp\left(-gamma * sum\left(k - k_i^2\right)\right) \tag{1.9}$$

Gamma here varies from 0 to 1. In the algorithm, we must explicitly define it.

- **Sigmoid:** The Sigmoid Kernel is like a two-layer perceptron neural network model that artificial neurons use as an activation function.

1.4.6 Decision Tree

The classification model is constructed in the form of a tree structure by the decision-tree algorithm (DTA). It uses if-then rules, which are both complete and cannot be used together to classify anything. As the process goes on, the data is broken up into smaller pieces and eventually linked to a decision tree that grows over time. The finished structure looks like the branches and leaves of a tree. The rules are sequentially taught using the training data [70, 62]. When a rule is captured, the corresponding data items are removed. This procedure repeats itself until the target is achieved on the training set. It is a top-down, recursive, divide-and-conquer process that generates the tree. Each leaf of a classifier represents a different categorization or verdict. Decision trees specialize in processing both quantitative and qualitative information. The root node is the node at the top of the tree that corresponds to the best predictor.

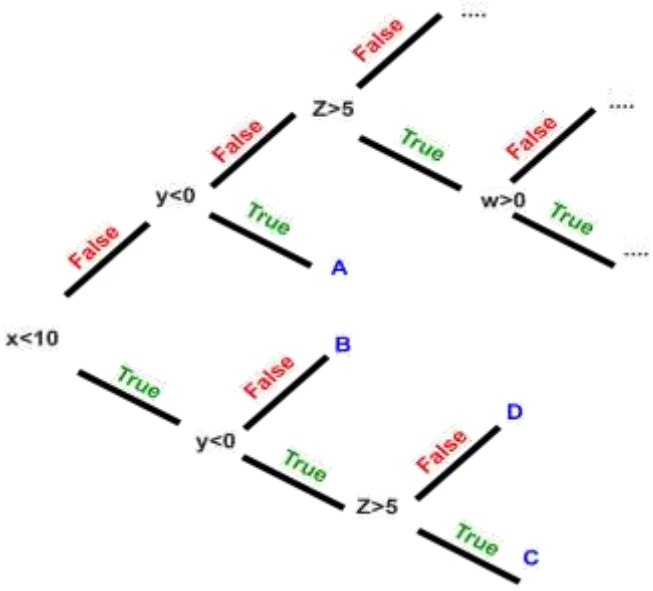

Figure 1.9 Decision tree classifier

Figure 1.9 represents a decision tree that is easy to understand the process, and the user does not have to do much to get the data ready for it either. The decision tree has the limitation of theoretically producing overly complex trees, which are then difficult to classify. Because a seemingly inconsequential shift in the data can modify the entire tree, they are not always reliable.

Decision trees are a specific kind of supervised machine learning technique in which the data is always split up based on a specific parameter. Explaining the inner workings of a decision tree requires discussing two concepts, Leaves and Decision nodes. The outcomes or options are represented by the leaves. The decision nodes deviate from the data streams [71, 72]. Each data point is allocated to a segment after the classifier constructs a single decision tree. The maximum number of features considered by the model can be altered.

1.4.7 Random Forest

Team training can be accomplished through the use of random decision trees or random forests, which can be used for classification, regression, and other tasks. The method works by training with a large number of decision trees and then returning a class that represents either the mean of all classes or the mean prediction (regression) of all trees [71].

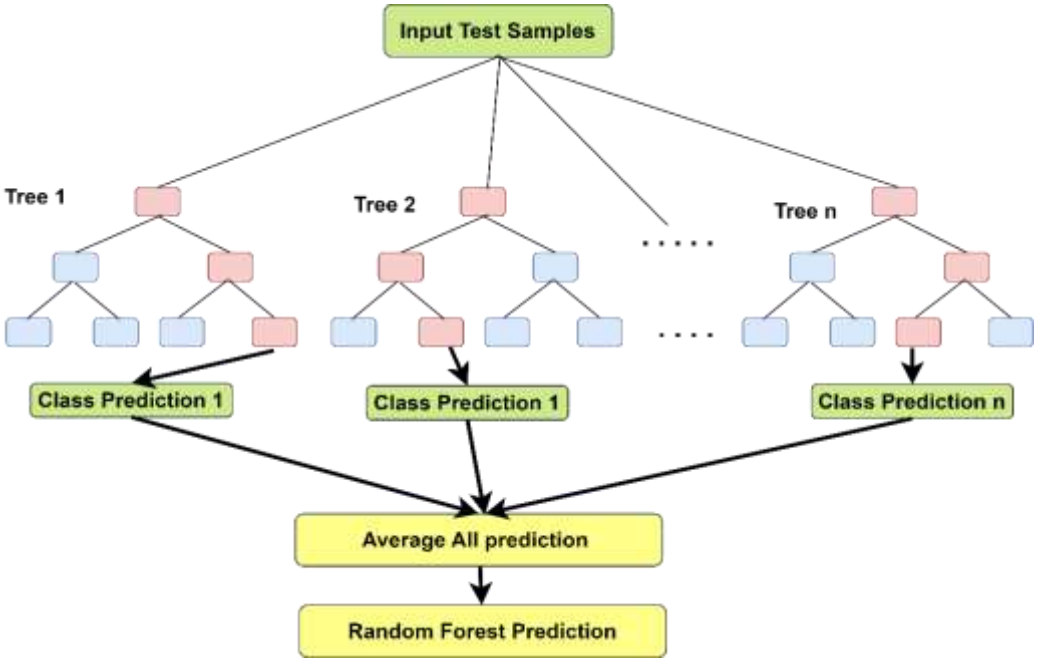

Figure 1.10 Random Forest Classifier with tree structure

Figure 1.10 represents the random forest trees as pink and blue box are represented separate classes and to connecting line as the decision boundary. It is a structured machine-learning technique that is frequently used to rectify classification and regression problems. The name comes from real life, where a forest is made up of

trees, and the more trees there are, the stronger the forest is. In the same way, the RF method takes a sample of data and uses it to make multiple decision trees.

Depending on the results of these trees, it selects the most appropriate solution. It is viewed as superior to DT because averaging the results eliminates the possibility of over-fitting. This classifier improves upon the DT framework. It builds a diverse forest of trees with a wide range of characteristics by randomly selecting features from the entire dataset. Constantly adapting the number of trees used to make a class guessable. Ten trees, one hundred trees, two hundred trees, five hundred trees, and a thousand trees were used to generate test scores for our analysis. To improve the model's prognostic accuracy, a random forest uses multiple trees, each tailored to fit a different subset of the data, and then takes an average [72]. While input and subsample sizes are always the same, substitutions are frequently used when drawing samples. Instead of trying to fit the data too well, the random forest achieves better accuracy. The main drawback of using a random forest classifier is that it can be time-consuming to set up and can become quite sluggish when making predictions in real-time.

1.4.8 Naive Bayes Classifier

An approach to categorization is predicated on the idea that predictors are unrelated to one another, as stated by bayes' theorem. Simply put, a Naive Bayes classifier assumes that the inclusion of one feature in a class is separate from the existence of any other feature. Even if some of the features depend on others, they all add to the probability on their own [73]. The naive bayes model is easy to make and works best when there are a lot of data points. Even though naive bayes is a simple method, it is known to work better than most classification methods in machine learning. Here is how to use the naive bayes theorem with the bayes theorem.

$$P(C_i|x_1, x_2, \dots, x_n) = \frac{P(x_1,x_2,\dots,x_n|C_i)\cdot P(C_i)}{P(x_1,x_2,\dots,x_n)} \; for\; 1 < i < k \qquad (1.10)$$

A relatively small amount of training data is required for the naive bayes classifier to estimate the parameters required to obtain the results. In comparison to other classifiers, they move very quickly [74]. The only bad thing about them is that they are not very good at estimating. Bayes' Theorem provides the foundation of such a family of classification models known collectively as naive bayes, with the assumption of estimator separation as its central component.

$$P(\alpha|\beta) = \frac{P(\beta|\alpha)\cdot P(\alpha)}{P(\beta)} \quad (1.11)$$

As conditional probability is the focus of Bayes' Theorem, its formula (1.11).

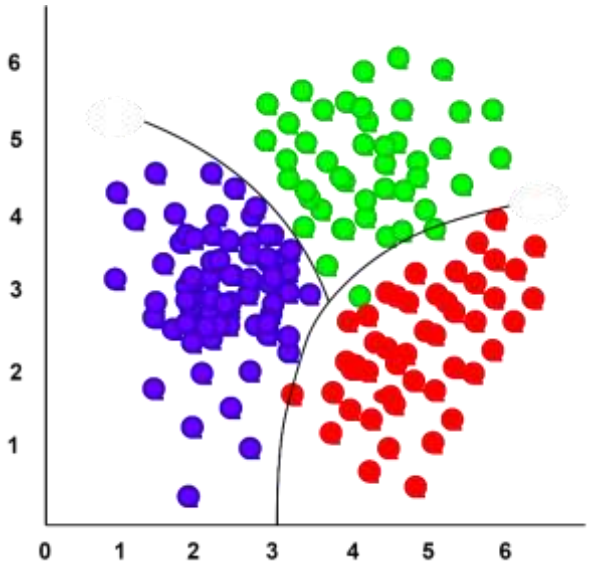

Figure 1.11 Naïve Bayes Classification

Figure 1.11 represents the Naive Bayes model works very well when there are a lot of data points, and it is easy to make. It's not simply one algorithm but several that share the same functionality, because it is so simple, nave bayes often does a better job of classifying than the most advanced methods [75]. The naive bayes classifier estimates the likelihood that a data point is part of a specific class based on its characteristics. Gaussian naive bayes in this used in the research. Attributes of a continuous characteristic are typically assumed to follow a Gaussian (or Normal) distribution across classes.

1.4.9. XGBoost

The gradient boosting algorithm has been upgraded to become the XGBoost technique. XGBoost is just a gradient descent-based tree-based technique used to teach techniques that are otherwise deemed to be at a disadvantage. The fundamental design is enhanced by XGBoost, which enhances the system's functionality and techniques [76]. As part of the distributed ML Community, there is a package known as XGBoost. Using the gradient boosting method and decision trees, it is a ML approach. ANNs are thought to be better than any algorithm when it comes to making predictions based on unstructured data [77]. But DT-based algorithms are known for how well they work

with medium-sized to small tabular data. XGBoost allows for a parallel implementation phase, allowing for sequential tree construction to take place. The loops employed in the construction of such fundamental students can be repurposed in this manner.

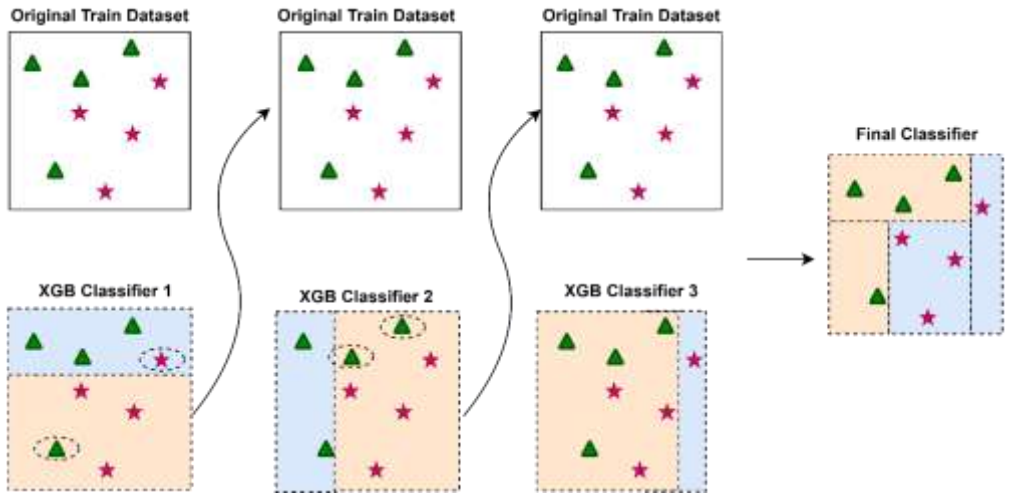

Figure 1.12 XGBoost Classification methodology

As shown in figure 1.12 above, classifier 1 receives the training dataset. The classifier predicted a triangle on the background in blue, and a star on the background in pink. One triangle and one star were incorrectly predicted by the classifier 1 model. A circle is used to denote them. These data points that were incorrectly predicted are sent to the following classifier with increased weights. To classifier 2, then. The two triangles are correctly predicted by classifier 2 while classifier 1 was unable to do so. However, classifier 2 also commits a few other mistakes. After more iterations, we have a final combined classifier that accurately predicts every data point.

In the CVD framework, the main idea behind tree-splitting greedy stopping criteria is the need for a loss at the split point. XGBoost uses the "max depth" variable instead of the "criteria" variable first and then prunes the trees in the wrong order. This method is made so that it makes good use of the available hardware resources. Cache awareness is used to do this, and each thread makes its buffers to hold gradient statistics [78]. Some of the algorithmic advances used include regularisation, sparsity awareness, weighted quantile sketch, and cross-validation.

1.5 Unsupervised Machine Learning

Unsupervised learning looks at datasets that have not been labeled and group them. These machine-learning algorithms help find patterns or groups of data that were

not obvious before. Image recognition is a common way that unsupervised learning is used. Clustering, neural networks, finding outliers, and other things are all examples of unsupervised learning models [79]. This algorithm is an example of a machine learning (ML) algorithm, in which a computer is given access to a dataset and instructed to find connections and patterns within that dataset. In unsupervised machine learning, the information is not sorted or labeled like it is in supervised machine learning [80]. This system is not used to find the right answer. Instead, it looks at a set of data and concludes on its own.

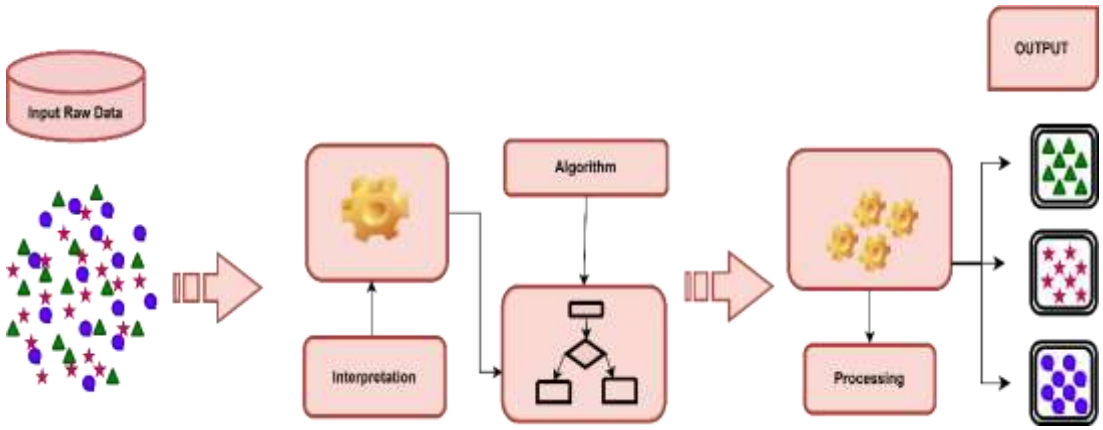

Figure 1.13 Unsupervised machine learning process

Figure 1.13 shows the symbols classified based on the input data and but not contain any label the machine learning can train unsupervised algorithms and predicted the result in unsupervised machine learning, the following are some of the most common algorithms. K means, KNN, Gaussian mixture, Principal Component Analysis (PCA). This algorithm is used for organizing unsupervised data into subgroups that share more characteristics than any clusters. Statistical analysis is widely used in fields such as pattern recognition, image analysis, data compression, and bioinformatics. There is no single algorithm that does cluster analysis. Alternatively, it is a job that can be accomplished by a variety of algorithms [81]. Each algorithm has its way of defining a cluster, and those ways can vary widely. Methods vary in their effectiveness at finding and grouping clusters based on the development company or system performance.

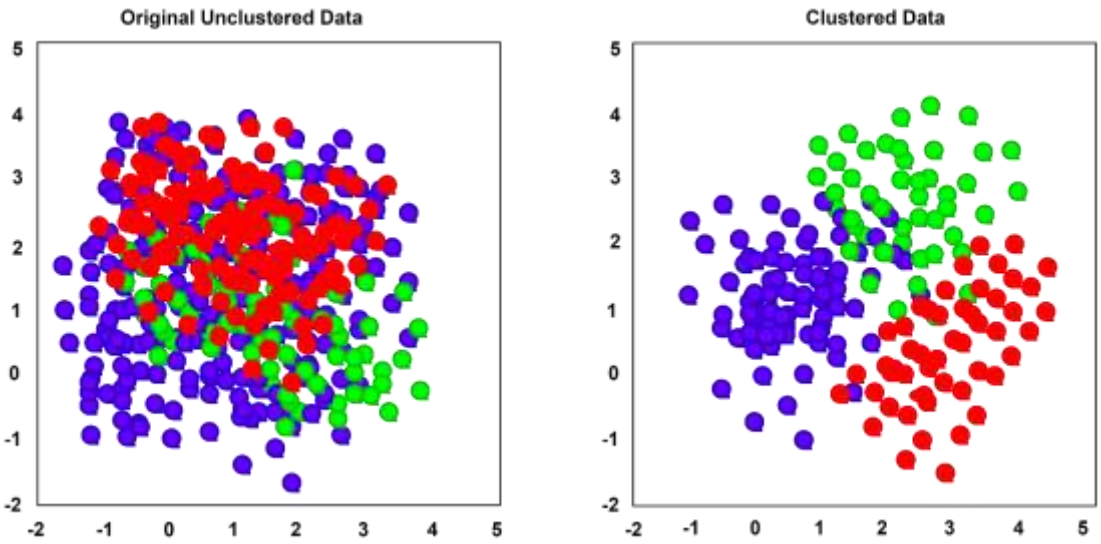

Figure 1.14 (a) Non-clustered data, b) Cluster data based on the color.

Figure 1.14 shows the cluster analysis representation The first image is displayed using non-clustered data, which includes a mixture of red, blue, and green objects. Based on the same cluster property, such as color, the objects are grouped in the second image.

1.5.1 Machine learning with partial supervision

This approach to machine learning falls somewhere between supervised and unsupervised methods. This would be because when training, they implement both labeled and unlabelled information, with the labeled data typically comprising a small percentage. Machine learning accuracy can be significantly improved by combining unlabelled data with a little amount of labeled data.

1.6 Reinforcement machine learning

With reinforcement learning, you teach models to make a set of decisions in a certain order. Think of this as a game of try and see. The machine will be given rewards or punishments based on what it does to get it to do what the user wants. The user wants it to learn how to get the most rewards in the end [82, 83, 60]. In the real world, a good example of this is Facebook's Horizon, which uses reinforcement learning to do things like make suggestions more relevant to each user and send them more meaningful notifications.

This method of learning interacts with its environment by doing things and then finding mistakes or new ideas. Data is collected and the best course of action is determined through a method of trial and error in machine learning. For the machine to figure out what the best thing to do is, it needs simple reward feedback. Reinforcement learning works best for problems where you have to choose between long-term and short-term rewards. Since it does not require any pre-existing data, it is also useful when there is little to no background information on the issue at hand [84]. Researchers have looked into this because it is one of the most important things going on in the world today. Hear more than one method to look at the problem and compare how well they worked.

1.7 Deep learning

AI-based neural networks combined with representation-based training are the foundation of deep learning. Learning can happen with or without supervision. The term "Deep Learning" refers to a subfield of computer vision that employs networks of artificial neurons. Just the same as machine learning, deep learning can be used in a supervised, semi-supervised, or unsupervised setting. Recurrent neural networks (RNNs), deep neural networks (DNNs), and convolutional neural networks (CNNs) are just a few examples of deep learning architectures that have observed a wide variety of uses (CNNs) [91]. In artificial intelligence, machine learning is the most common method. Based on what we said in the last article, machine learning is an algorithm that keeps getting better by constantly analyzing patterns and new information. Machine learning includes deep learning as one of its parts.

1.7.1 Introduction

Artificial neural networks are used in deep learning to perform numerical computations on large datasets. To put it simply, it is a form of machine learning inspired by the structure and operation of the neural network. By simulating human behaviour patterns, deep learning algorithms train computers to perform tasks. Applications of deep learning are widespread and include everything from medicine and retail to the arts and marketing. AI is a set of algorithms and intelligence intended to imitate human intelligence. Among these is machine learning, of which deep learning is a subset [92]. The goal of deep learning in artificial intelligence is to create systems that can analyze large amounts of data and draw conclusions based on those analyses.

Artificial intelligence's deep learning method employs self-teaching networks to make sense of unlabelled or unstructured data.

Artificial neural networks are used in a hierarchical structure in deep learning, a type of machine learning. In conventional software, the data is analyzed linearly. In contrast, deep learning employs a hierarchical function to teach computers to conduct nonlinear analyses of data. Computer models can be trained to perform a wide variety of tasks, including the recognition of specific images, words, and sounds. As accurate, or even more so, than humans, deep learning can be a real possibility. To train deep learning algorithms, a large dataset of labeled data and a network of neural nodes with many layers are used. Self-driving vehicles are the primary application of deep learning currently in the automotive industry [93]. Researchers in the automotive sector use deep learning to develop systems that can detect pedestrians, traffic lights, and stop signs automatically. Algorithms are continuously collecting new data, and the system is constantly learning from its previous experiences.

1.7.2 Deep learning methodology

To replicate the biological brain's information processing capabilities, deep learning algorithms implement ANNs and self-learning characterizations. During the training phase, algorithms enhance the blank spaces in the input distribution to extract features, classify objects, and discover meaningful data patterns. This occurs on various tiers, with algorithms serving as the foundation upon which models are constructed. This is very much like training a self-learning machine. Deep learning models incorporate some different algorithms. While there is no such thing as a perfect network, there are some algorithms that perform certain tasks more efficiently than others. You need to have a solid understanding of each of the major algorithms to make the right choices. To solve complex problems, deep learning algorithms require a large amount of computing power and data.

Machine learning includes deep learning as one of its parts. Deep learning models are made to look at data in a way that is logically similar to how people make decisions and come to conclusions. These models are based on the human brain, and they let data move between nodes that look like neurons [94]. They layer algorithms to make an artificial neural network (ANN) that can learn and make decisions on its own. Deep learning models are smarter than regular machine learning models because of how they are built. Deep learning systems usually need a lot of data to work well, but once

they have the data, they can show results right away. Once it is set up, very little human help is needed. Transfer learning, which uses models that have already been trained, is a big step forward in the field of deep learning. These models that have already been trained help meet the need for large training datasets.

1.7.3 Artificial neural networks

Each neuron in a neural network obtains a vector of input data and generates a corresponding output value. Each neuron takes in data, processes it with a function often a nonlinear one, and then passes on the result to the next layer [95, 91]. A neural network is a collection of neurons organized in layers, each of which receives an input vector and outputs an appropriate response. In this process, each neuron takes in information, applies a function to it that is often not linear, and then sends the result to the next layer.

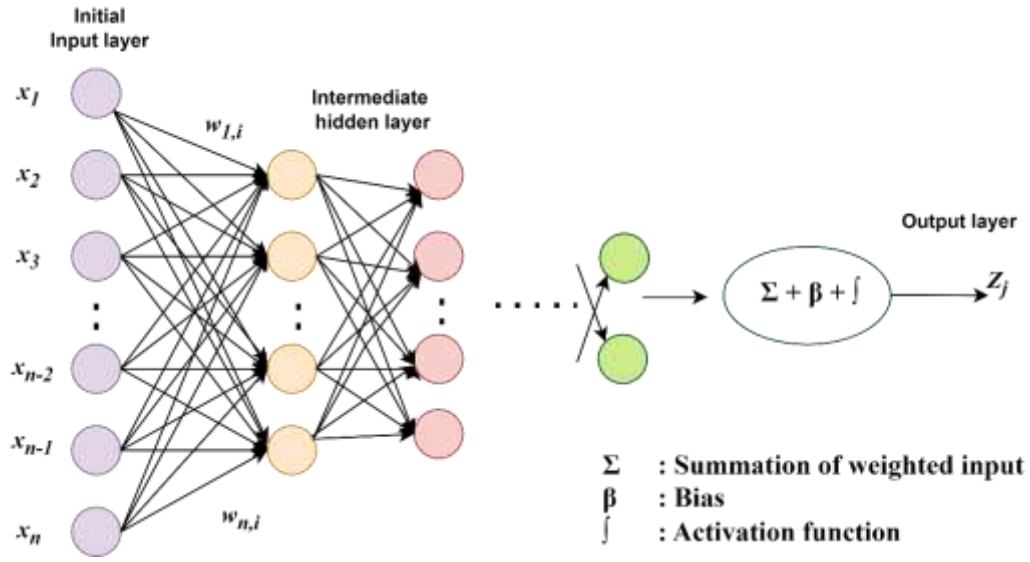

Figure 1.15 Artificial Neural Network workflow structure

Figure 1.15 shows the artificial neural network workflow architecture it starts with inputs as x_1, $x_{2,}$ and x_n intermediate hidden layer and output. In general, we assume a sigmoid relationship between the input variables and the activation rate of hidden nodes or between the hidden nodes and the activation rate of output nodes. And the weights of the linkages can be denoted as $w_{i,j}$. Most networks are designed to be feed-forward, meaning that the output from one layer is sent to the next without any feedback being sent back to the layer that came before it. Signals that move from one layer to the

next are given weights, and throughout the training, a neural network's effectiveness for a given problem statement is improved by adjusting these weights.

It can handle noisy data well and can recognize patterns that have not been taught to it. It works best with inputs and outputs that have continuous values. The problem with artificial neural networks is that they are not as easy to understand as other models.

- Handwriting analysis
- Colorization of black-and-white images
- Computer vision processes
- Captioning photos based on facial features

1.7.4 Convoluted neural networks (CNNs)

This kind of neural network has more than one layer. These layers look at the data and pull-out features. CNNs are mostly used for computer vision, processing images, and finding objects. Convolutional networks (or ConvNets) are multi-layered models typically employed in the fields of image processing and object detection [96]. The original convolutional neural network (CNN) was created by Yann LeCun in 1988 under the name LeNet. Zip code and number recognition were among its primary applications.

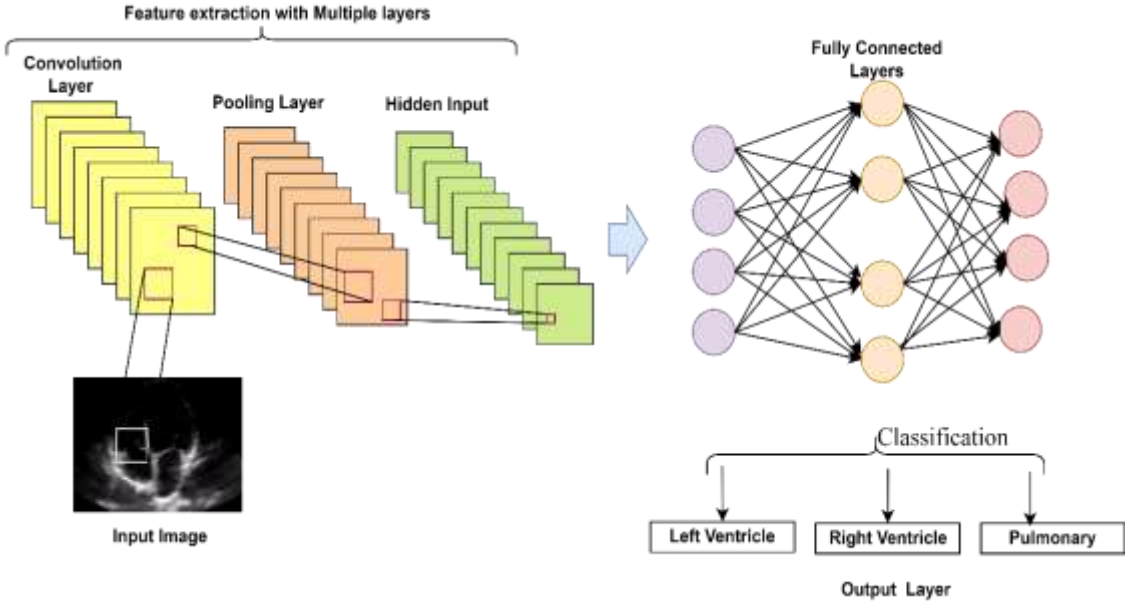

Figure 1.16 Convolution neural network architecture layer diagram

Common applications of CNNs include object recognition in satellite images, medical image processing, time series forecasting, and anomaly detection. Here is an example of an image that CNN has changed.

Figure 1.16 represents the CNNs have been useful to initial phase of CNNs information and data extract features at multiple levels. CNN's convolution layer, which contains many filters, is responsible for the actual convolution operation [97, 59]. Second processing data into CNNs use a layer called Rectified Linear Units (ReLUs) to perform operations on data points. A refined feature map is a result. As soon as the feature map has been corrected, it is fed into a pooling layer. The dimensions of a feature map can be reduced in two ways, through pooling (down sampling) and through flattening (generating two-dimensional arrays from the pooled feature map) [98]. When the pooling layer's flattened matrix is used as an input, we have a fully connected layer. Images are categorized and labeled at this layer.

1.7.5 Back-to-back neural networks (RNNs)

To solve ordinal and temporal problems, RNNs employ sequential data or time-series data. They learn from training data. Google Translate, Siri, and image captioning are all common ways that RNNs are used.

There are three layers of these nodes

- Input layer
- Hidden layer(s)
- Output layer

Autoencoders

Neural networks are used by autoencoders to learn how to represent things. These models are used to solve problems requiring unsupervised learning by copying information from an input layer to an output layer. They are accustomed to tasks such as image processing and drug discovery [99]. Nodes, which are analogous to human neurons, are the building blocks of a neural network that is structured similarly to the brain.

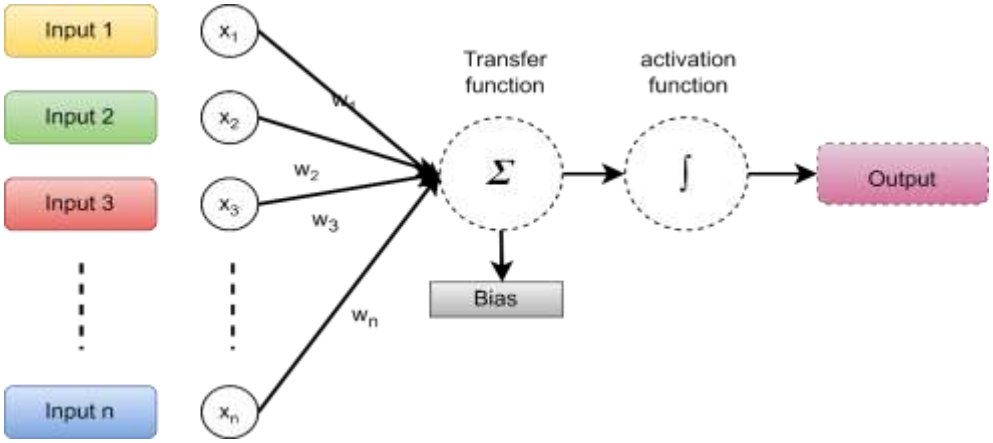

Figure 1.17 General layer diagram for the basic layer

Figure 1.17 represents the basic neural layer works and data serve as input variables to base stations, providing them with information. Using random weights, the node multiplies the inputs computes the outputs and then applies a bias. Finally, optimization algorithms special kinds of non-linear functions used to select the neuron that will be activated.

1.7.6 Long short-term memory networks

LSTMs are a type of Recurrent Neural Network (RNN) that could also learn and remember complex connections over time. The tendency to retain information for a long time is built in. Long short-term memories (LSTMs) can hold information for later recall. They can be helpful for time series prediction since they retain information from previous inputs. LSTMs are organized like a chain, with four layers that talk to one another in their special ways [100].

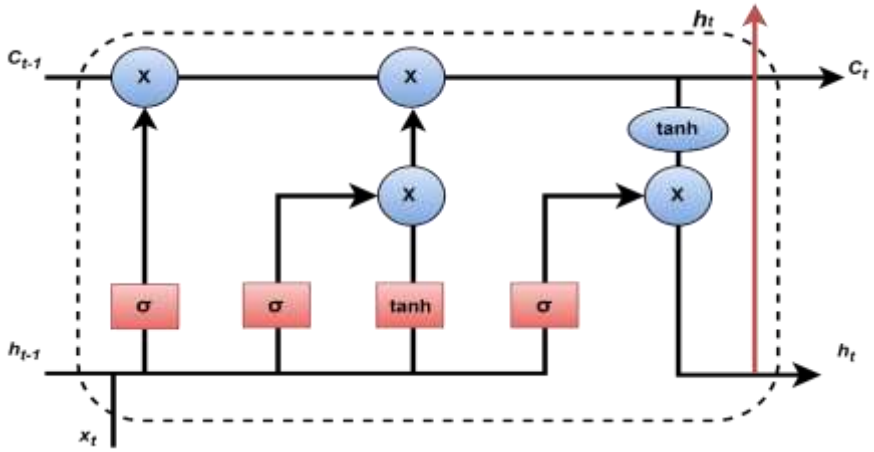

Figure 1.18 Workflow diagram for Long Short-Term Memory network

Figure 1.18 represents the cell state C_{t-1}, C_{t+1} input, and output, the h_{t-1}, h_{t+1} hidden state of the input and output timestamps steps present in the LSTM model. The cell state after the hidden state goes to the tanh function and gets multiplied by the sigmoid function of the output state.

Other common applications of LSTMs beyond time series prediction include speech recognition, music composition, and drug discovery. Initially, LSTM does not remember irrelevant features of the previous condition, and then they only update some of the previous cell values. Ultimately, they alter the cell's state's output. An explanation of how LSTMs work is depicted in the following diagram.

1.7.7 Recurrent neural networks

Through connections that produce directed cycles, LSTM outputs can be fed into the active phase of RNNs. The current phase receives the LSTM's output. As a result of the current phase and the LSTM's ability to store information, previous inputs can be recalled. Preview images, analyzing time series, processing natural language, recognizing writing style, and translating theory between alphabets are just some of the common uses of RNNs [101,102]. Figure1.19 An unfolded RNN looks like this,

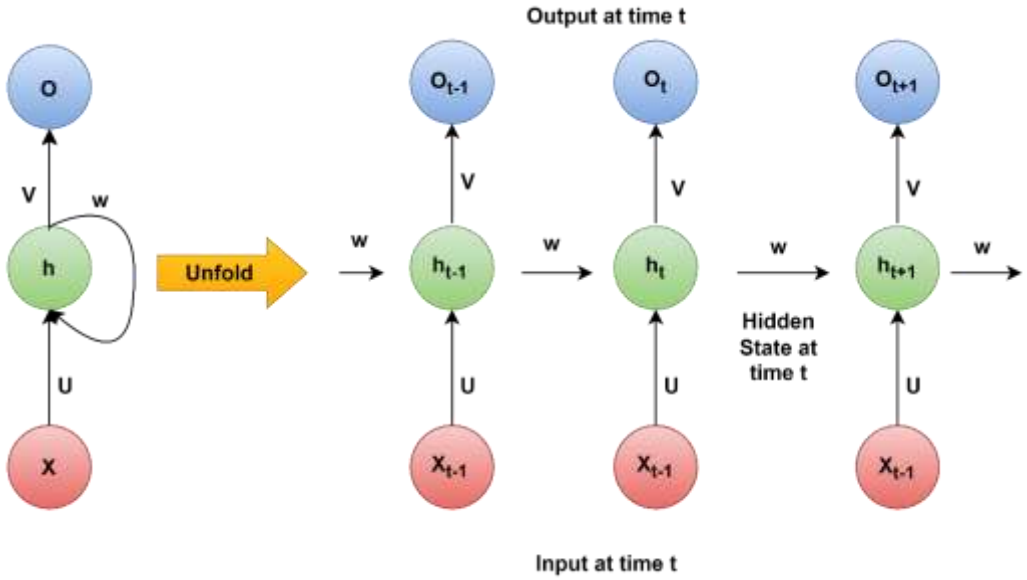

Figure1.19 Workflow diagram for Recurrent Neural Networks

Figure 1.19 represents the value U, V, and W are the weighting matrices of the input-to-hidden connection, hidden-to-output connection, and hidden-to-hidden connection. Both the (t-1) and (t) output are fed back into the (t) input, and the (t+1)

input is used to drive the t+1 output. Any input length is fine for RNNs. The model's size remains constant regardless of the size of the input, and the computation takes the past into account.

1.7.8 Generative adversarial networks

Generative Adversarial Networks (GANs) are a type of deep learning technique that can generate new information that is similar to the training data. GAN is composed of a generator that discovers how to create fake data and a discriminator that discovers how to spot and avoid false examples. Over time, more and more people have been using GANs. They can be used to improve pictures of the night sky and to study dark matter by simulating gravitational lensing [103]. Image training is used by game developers to recreate low-resolution 2D textures from old games in higher resolutions. GANs can be used to make realistic pictures and cartoon characters, take pictures of people's faces, and show what 3D objects look like. Below is a diagram of how GANs operate.

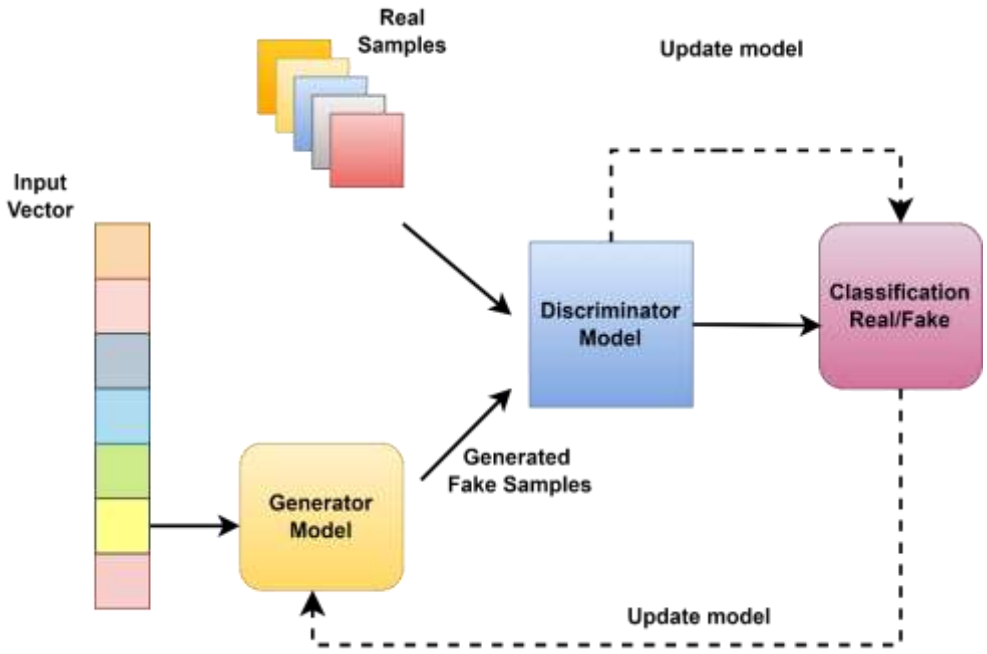

Figure 1.20 Generative adversarial networks workflow architecture diagram

Figure 1.20 represents the GAN input of some random noise (n) used by the generator to try and produce fake pictures. It is the goal of discriminator D to distinguish G's generated images from those in the real-world dataset (x), for which it has been trained. The discriminator accepts both authentic and spoofed images as input,

calculating the likelihood that a given sample is drawn from the authentic image dataset rather than the spoofed one generated by G. The discriminator is trained to recognize fake data from the generator during the initial training phase. The GAN communicates its findings to both the generator and the discriminator to refine the model. The differentiators can tell the difference between real and fake data based on your input.

1.7.9 Radial basis function network

RBFN, which stand for radial basis function neural network, are a subset of feedforward neural network. Such machine learning is employed for numerous jobs, such as classification, regression, and time-series prediction. Similarity to having trained set examples is used by RBFN to make classification decisions. The input layer of RBFN is fed data in the form of a vector [103, 104].

All the neurons in that layer are RBF types. Each type of information is represented by a unique output node, as well as the function finds the weighted sum of the inputs. The hidden layer's Gaussian performance parameters produce outputs that decrease as one moves away from the neuron's center. Input radial-basis functions are linearly combined with neuron parameters to produce the network's performance. See this figure 1.21 of an RBFN

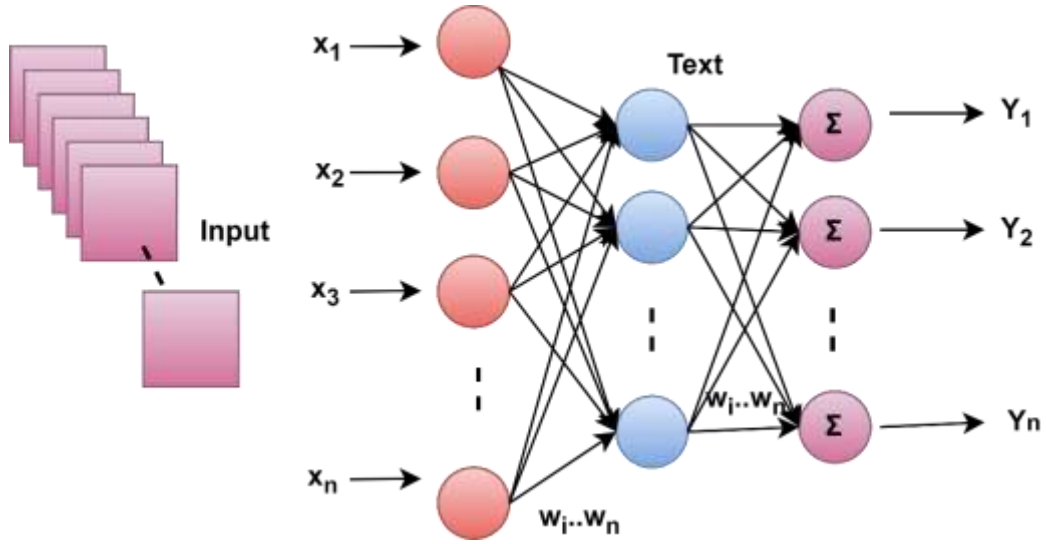

Figure 1.21 Workflow diagram for radial basis function network

1.7.10 Multilayer perceptron

Multilayer perceptron (MLP) are specific kinds of neural networks that feedforward and use activation functions across multiple layers of the perceptron. In

MLPs, the connection between the input and output layers is continuous [105]. On the other hand, they can be used to create speech-recognition, image-recognition, and machine translation software with the same number of layers of input and output layers.

The layers of neurons are graph-connected so that the signal can only travel in one direction. MLPs determine input based on the weights connecting the input and hidden layers. To determine which nodes to activate, MLPs rely on activation functions. Examples of activation functions include the ReLU, sigmoid, & tan h functions. Multi-layer perceptron (MLP) uses a data set for training to learn and practice the model of the interdependencies and correlations between the independent variable and the target variable [106]. Observe this illustration of a multi-level marketing-planning structure. figure 1.22 The diagram determines the input layer of x1, x2, …, xn, the intermediate hidden layer, and a final output layer with necessary weights and biases and applies the appropriate activation functions to determine whether or not an input image is of a star or triangle or circle.

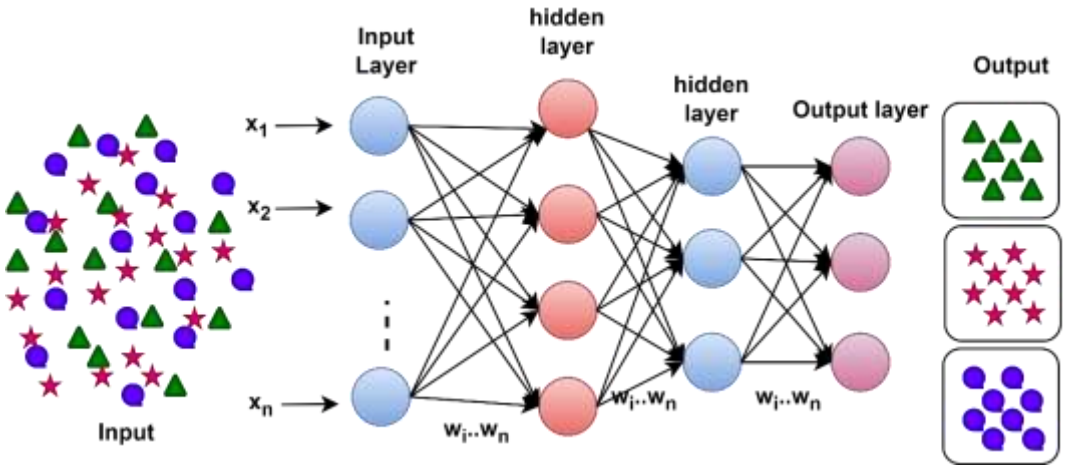

Figure 1.22 Workflow architecture diagram for multilayer perceptron

1.7.11 Self-organizing maps

Using self-organizing artificial neural network (SOMs), Professor Teuvo Kohonen developed a way to minimize the dimensionality of data. The goal of data visualization is to make high-dimensional data easier to understand and interact with for humans. The purpose of a SOM is to facilitate the comprehension of such data. Exactly the SOMs work, In SOMs, an irregular vector is selected first from the classification algorithm to serve as the initial weight with each node [108,109]. The most probable input vectors are indeed the weights assigned to the nodes. The winning

node is the one with the best possible matching unit. Eventually, the number of BMU neighbours will decrease as word of the neighbourhood spreads among SOMs.

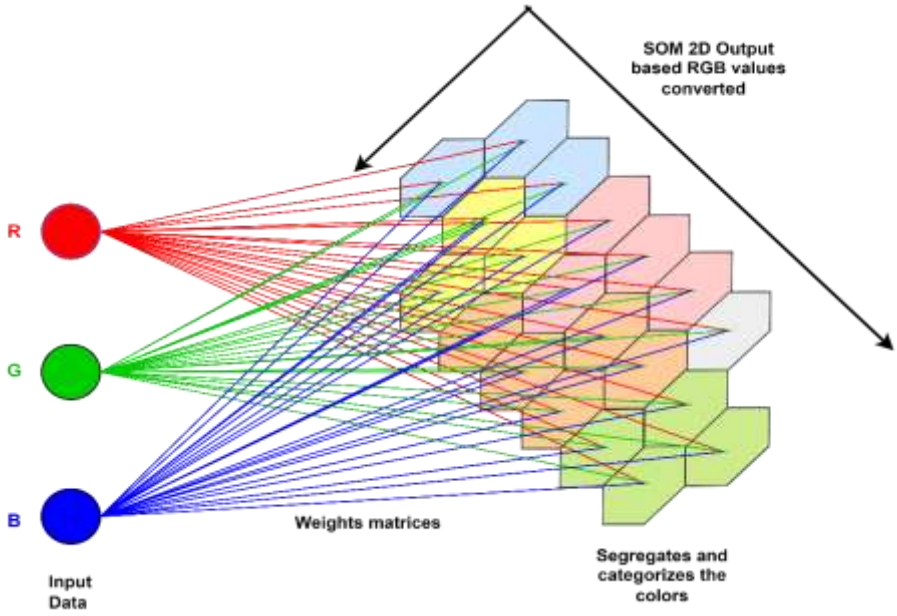

Figure 1.23 Self-Organizing Maps Workflow diagram

In SOMs, the sample vector gets an advantage. There is a correlation between a node's proximity to a BMU and the rate at which its weight changes. The farther away the neighbour from the BMU, the lesser it learns. When running a SOM, the second step is repeated N times. A colorful representation of a vector input is depicted in above figure 1.23. A SOM processes this data and outputs 2D RGB values. At last, it categorizes the hues into distinct piles.

1.7.12 Deep belief network

DBN may be multi-layered generative models that rely on a large number of hidden, random variables. The "hidden units" are the binary values of the latent variables. DBNs are constructed from a network of interconnected Boltzmann Machines arranged in a stack [110]. When using an RBM, each layer can communicate with the layers above and behind it. Images, videos, and motions can all be identified with the help of DBNs. Below is an example of DBN architecture

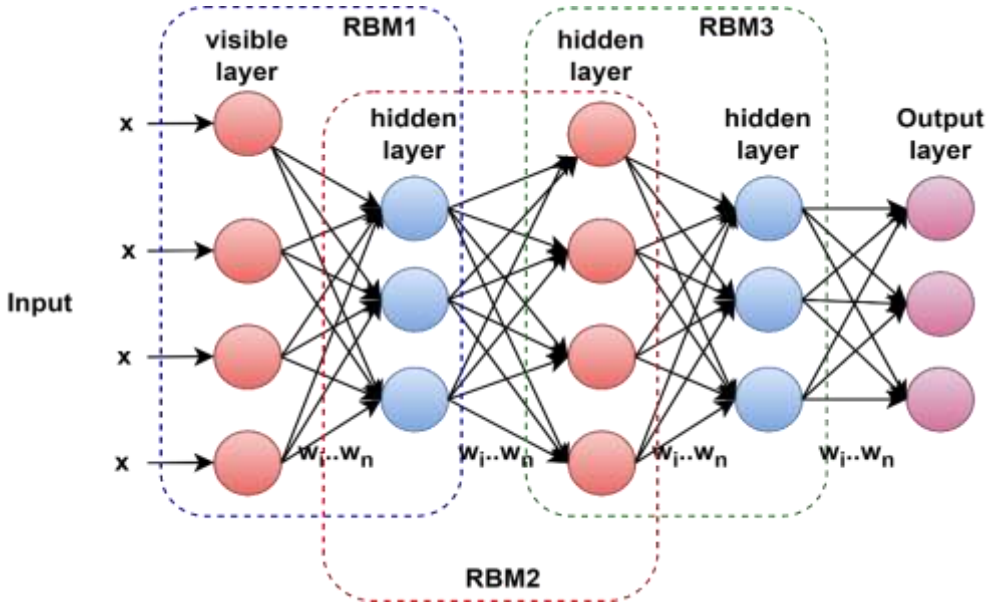

Figure 1.24 Workflow diagram architecture for Deep Belief Network

Figure 1.24 shows the deep belief network value x represents the input layer and size in the case of the classification DBN, the number of classes in the representative dataset is included in each record type, along with the RBMs that make up the layers of the network. In contrast to the DBN record, which depicts a model composed solely of stacked RBMs, the CDBN record includes the CRBM record that serves as the top-level associative memory.

To summarise, greedy learning techniques are used to train DBN. The top-down, generative weights are learned incrementally by the greedy learning algorithm. it is implemented in Gibb's sampling procedures on the first two hidden layers [111]. In this step, the RBM is sampled using the first two hidden layers. DBNs collect a sample from the observable units by performing a single pass of ancestral sampling through the rest of the model. At each layer, a single, bottom-up pass enables DBNs to deduce the latent variable values.

1.7.13 Restricted Boltzmann machine

RBM first explored by Geoffrey Hinton, are a type of stochastic neural network that can acquire knowledge from an unpredictable assortment of inputs. Dimensionality reduction, classification, regression, content-based filtering, pattern recognition, feature extraction, topic modeling, and more are all possible thanks to its implementation. On top of RBMs, DBNs are constructed [112]. RBMs consist of double layer Units that are

visible and masked units. There is no way to distinguish between visible and invisible components. Instead of having output nodes, RBMs have a bias unit that is connected to both the units that are on display and the ones that are not.

A schematic of how RBMs work is shown below

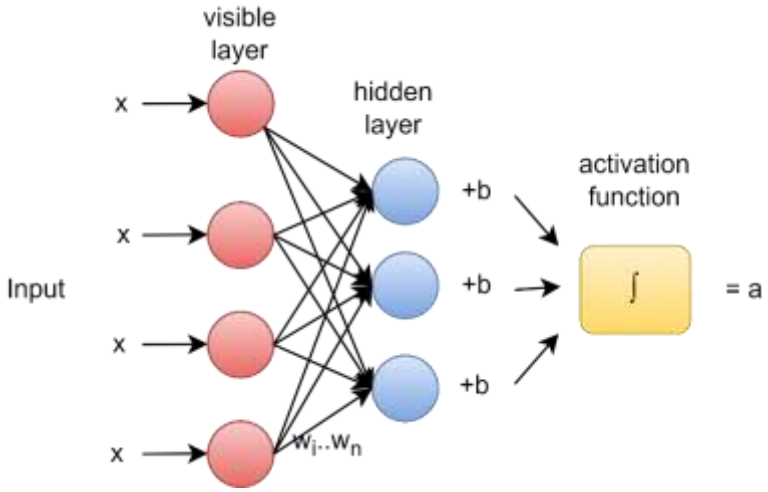

Figure1.25 Restricted Boltzmann Machines Workflow diagram

When operating an RBM, there are two distinct phases forward pass and reverse pass. Figure 1.25 shows the x as inputs are encoded using RBMs during the forward pass. RBMs combine all data into a single bias that is then weighted. The results of the algorithm are passed on to a hidden layer for analysis [113]. In the RBM's backward pass, these numbers are used to reconstruct the inputs. RBMs reconstruct images by combining the output of the visible layer with a weighted and partial mixture of the activations. RBM evaluates the quality of the output at a perceptual layer by comparing reproduction to the original data.

The Cleveland data source is a popular example of a well-known collection. The data set at issue in this thesis. The efficiency of prediction methods has been evaluated by researchers using the cleveland dataset [114]. The information can be used by other specialists in conjunction with the data they already have from local hospitals. The outcomes and methods they employ are dissimilar. As a result, this part will attempt to summarise the research that has been done with this data set. Three distinct phases make up clinical image analysis. Find the slices in the ultrasound image that contain the lesions first. Mark out the lesions or other structures that will serve as our points of

focus in the following step. In a conclusion, the structures of interest can be used to derive quantitative or qualitative measurements (e.g., size, shape, or texture).

Chapters 2 & 3 discussed various ML and DL networks and how it is working principles are seen and followed those methods for classifying and segmenting medical images are developed using a variety of supervised and unsupervised learning algorithms as well as deep learning network models. In Chapter 3, a framework based on the Kushner-stanovich filter and curvelet transform is suggested for segmenting US images. AWM Filter-based Yolov4+ with Darknet53 Network and Mish active function are heavily utilized in Chapter 4 to create methods for segmentation and classification issues.

1.8 Dataset preparation

Diagnostic investigations from 500 patients were collected online from the University Hospital of St. Etienne (France) for this research project, with their identities concealed under guidelines established by the hospital's local ethical committee. The data collection processes were fine-tuned to maximize left ventricular ejection fraction readings. In the interest of maintaining a high level of clinical realism, no data selection nor any other prerequisite procedures have been carried out [115]. As a side benefit some cases were hard to trace, and there is a lot of variation in the selected settings in the dataset, in some cases, the images did not show the full wall.

In some instances, it was impossible to acquire a rigorous four-chambers view consistent with the probe orientation suggestion, so a five-chambers view was obtained instead. This resulted in a dataset that is highly diverse in terms of both image quality and pathological cases, which is indicative of data collected during routine clinical practice. The data set can be accessed by the public. online access, there are two parts to the dataset 450 patients used for training, with accompanying manual references based on the analysis of a single clinical expert; and 50 patients used for testing. The raw/mhd file format is used to supply the input images.

1.8.1 Study population

The majority of people in the dataset have what is considered a pathologically low left ventricular ejection fraction (less than 45 percent) (beyond the uncertainty of the measurement). In addition, one expert believes that 19% of the images are of unsatisfactory quality, meaning that for this group, it is not possible to reliably estimate

clinical indices or identify the endocardium and epicardium of the left ventricle. Poor-quality images are typically excluded from the dataset in conventional analysis because of their limited clinical utility. As a result, these records were not developed during the project's computation of the various metrics, but rather served as training and validation sets for deep learning methods to investigate their impact.

1.8.2 Involved systems

In total, a GE M5S probe and a GE Vivid E95 ultrasound scanning (GE Vingmed Ultrasound, Horten Norway) were used to collect the data (GE Healthcare, US). No special procedures were instituted outside of the regular clinical routine. Echo PAC analysis tool extracted sequences of 2D apical four-chamber and two-chamber views for every patient [115, 116]. To estimate left ventricular ejection fraction values using Simpson's biplane technique of discs, these common cardiac viewpoints were selected for this analysis. Each exported file is a group of polar-coordinated B-mode images. The entire sequences were expressed in cartesian coordinates using the same linearization process, with a grid resolution of $\lambda/2 = 0.3$ mm along every x-axis (the axis parallel to the probe) and $\lambda/4 = 0.15$ mm along the z-axis (the axis perpendicular to the probe). The cardiac cycle was acquired for each patient in each view, allowing for labeled training data of cardiac frameworks.

1.9 Thesis Establishment

The thesis is divided into six chapters, each of which focuses on a distinct aspect of the research. Here is a summary of all the chapter has to say.

Chapter 1 is an introduction to finding heart diseases and a more in-depth look at the problems. In addition, the research objectives and goals of this study are described. In addition, it provides an outline of machine learning and Deep learning. This chapter also talks about the reasons why this research was done and what the thesis has to offer. In addition, the chapter contains information regarding the validation of the technique. provides a full analysis of the results and implementation of an integrated.

Chapter 2 Perform a comprehensive literature review of work done using a variety of Deep Learning methods and modeling techniques, as well as describing every existing series.

Chapter 3 Describes the first proposed method, which uses deep learning to automatically detect, categorize, and partition heart diseases. The many parts of the strategy that were deduced from this research are outlined. The built structures have also been evaluated to demonstrate the effectiveness and validity of the proposed method. Analyses the research procedures and infrastructure of the network.

Chapter 4 provides an overview of the first proposed approach, which makes use of deep learning to automatically detect, categorize, and partition heart diseases. All the various components of the strategy that can be inferred from this analysis are detailed. The efficacy and validity of the proposed method have also been demonstrated through the evaluation of the constructed structures. Examines the network's infrastructure and research processes. The methods for preparing the ultrasonography heart disease dataset in the ultrasound images and the analytical resources available to researchers were also described.

Chapter 5 Research results are briefly discussed. This chapter summarizes the thesis's basic point and discusses the many ways in which the research has been instrumental. Additional emphasis is placed on suggestions for further study. makes a final declaration after completing the investigation.

Chapter 6 Through "Conclusion & Future Work," we briefly recap the study's overall framework contribution and outline some potential directions for future study.

Chapter 2

Literature Survey and Background

Deep learning networks enhanced the performance of computer vision algorithms for image classification and object recognition recently. Deeper and more complicated networks have been analyzed here. DarkNet53 with curvelet transform and ResNet [117] are providing similar comparable results for a variety of computer vision tasks. Curvelet transform includes the directional parameters with needle-shaped elements and high directional specificity. Curvelet coefficients are obtained from the scaling and windowing function. Darknet contains multiple convolutional phases with pooling layers that decrease the pixel density of the convolutional features while increasing the image quality of the neurons. Deeper layers of the network acquire more accurate results along a route when compared to the localized information gained during the early convolutional stages [117]. Darknet53 adds shortcuts at each layer, which allows improved gradient flow across the structure and overcomes gradient vanishing problems efficiently when compared to VGGNet & ResNet.

2.1 Introduction

The application of deep learning is majorly based on the diagnostic challenges and, Demographic information, medical history, scan images, vital signs, skin lesions, retinal images, electronic recordings of electrocardiograms (ECGs), photographs of diagnostic procedures, and micrographs of clinical laboratories all qualify as data.

Ultrasonic cardiovascular images are used in this research work for clinical purposes thus, the training dataset and testing dataset were used along with existing methods and techniques to discover the research gap to create the proposed model [118]. Detailed descriptions of existing methods and models are provided in this chapter. Area Under the Curve (AUC) of true-positive characters is compared with false-positive variation rates, also known as the (ROC) Receiver Operating Characteristic Curve, to assess the efficiency of the proposed deep learning model.

Many prediction systems have been developed for the detection and prediction of various diseases using various DL methodologies. Some of these diseases are cancers, diabetes, heart disease, flu, the common cold, diseases caused by uterine

fibroids, and so on. Since the 1990s, people have been able to guess who will get heart disease [119]. The majority of articles have implemented many simple techniques for the machine learning models such as decision tree (DT), naïve Bayes (NB), neural network (NN), and support vector machine (SVM), with varying degrees of precision. Some articles utilized datasets from the University of California (UCI) library [120], but others utilized patient data. They presented a new model that improved the accuracy of the decision tree in predicting heart disease.

In lower-level procedures, there is no pre-built model to guide them, so they need to use the data directly. As well as excessive flexibility, there is also excessive noise, artifacts, and imprecise boundary shapes that may affect the result of classification and segmentation. The guidance is based on model-based approaches that depend on previously acquired knowledge. It is not uncommon to segment cardiac images automatically or semi-automatically using these techniques. For simplicity, we refer to any probabilistic or deformable model that has been considered.

2.2 Pre-processing

The actual data contains a large amount of incomplete and noisy data, to eliminate these defects the data must be pre-processed. The acquired data often contains noisy and null values due to deleting or changes in categories, repeated groups, and outlier data. To fit the data into a particular range, it may be necessary to normalize and scale it. Data transformation is the method of transforming processed data into a form or structure that is more suitable for model creation as well as information retrieval in particular.

The earlier stage of predicting disease is saving many people's life and the physician will tack the decision and start the treatment. The proposed method is having various stages such as Data collection, data pre-processing, and data transformation model to train the algorithm [121]. The method has extracted the feature with given input to select filter and wrapper classification techniques for the CVD classification.

2.3 Classification

In classification, we generally predict the result of anything in the form of discrete outputs. The target class in the data for each case is easily predicted with full accuracy, which is referred to as one of the important goals of classification [122]. For example, in giving loans to the applicants the bank keeps the records of high, medium,

and low risk of credit by classifying all the aspects of a loan given to the deserving applicant. So, in this way the classification is used everywhere to get a successful result.

For the random forests contain classification trees that are supervised by machine learning techniques. The classification trees partition the data iteratively until the unknown sample is identified as belonging to a particular class. The trees start with a subset of features and samples and partition the data accordingly. Random forests contain several such trees, and the final prediction of the forest is the majority vote of all the trees in it. Several research studies have incorporated random forest classifiers for tissue characterization and had shown successful classification rates [123-125]. The study presented in this paper generated random forest classifiers with fifty classification trees. The number of trees was selected to minimize the out-of-bag error of the classifier. The random forest classifiers were generated using the spectral features computed for three different bandwidth ranges of the power spectrum. The classifiers were trained with seventy-five percent of the available 180 ROI samples and tested on the remaining ROIs. The data was divided into a training set and a test set randomly.

Out-of-bag (OOB) error, a statistical measurement that predicts the error rate of each sample not used in the generation of the classification trees, was computed for each random forest that was generated. Another statistical measure known as the relative predictive importance was also computed for each bandwidth. This measure evaluates the importance of each predictor or spectral feature in distinctly separating the classes of the tissues [126]. the mentioned two measurements help in determining the appropriate parameters required for the classifier and improves the classifier's performance. OOB error was used in this study to determine the optimal number of trees required for classification.

In some investigations, mostly a portion of the pipeline classification of view in ultrasound heart images was used. Using a graph estimation method prediction obtained from a 3D model trained on video classification and tailored to this issue, [127] a technique of predicting the pose of the ultrasound probe in [128]. A comprehensive pipeline is put up [126] for using ultrasound heart pictures to analyze the anatomy and operation of the heart as well as spot disease. They use their model of convolutional neural networks to detect the echocardiographic view in the ultrasound picture, which is the first stage in their workflow.

Support vector machines [128] combined with various image processing techniques are the common solution to this heart disease. To build classification models

based on they used an accumulation of local characteristics in a spatial pyramid called the "spatial pyramid histogram of words" employed SVM. By employing viewpoints. Histograms of gradients are used as features in an SVM model [129] to categorize ultrasound pictures into one of two views. use a different method that combines GIST feature extraction with SVM for classification. They are successful in categorizing the photos into eight groups.

To detect the left ventricle ultrasound image is used boosting instead of the SVM model. As they construct global templates based on the location of the left ventricle, they can categorize the anatomy in a variety of ways. In a subsequent boost pass, these values are used to identify possible four echocardiographic views that the image represents [130]. As a consequence of this method, any view without the left ventricle cannot be classified. By using a Kalman-filter framework, [131] classified standard echocardiographic views by matching splines to heart structures visible in the image. After that, the object's echocardiographic viewpoint is identified using a scored system.

Expectation-Maximization (EM) texture classification algorithm, employs boundary search deformable shape models, and atlas propagation with image registration. EM classification techniques have been introduced successfully to the statistical classification of pictures of heart disorders. [132, 133] where its images are of higher quality (higher signal-to-noise ratio). Myocardium incompressibility was subsequently integrated into the 2D segmentation [134]. These cardiac data consisted of multi-slice, short-axis-oriented pictures that covered the heart from the apex to the basal planes. Based on how strong they were, the myocardium and ventricular blood cavities could be easily separated.

RFRS - ReliefF and Rough Set methods were combined in a hybrid classification system utilized in her work [135]. The same structure is split into two parts feature-selection system and a classification model (ensemble classifier). An algorithm known as relief F extracts feature from a set of discretized data. A group of classifiers is suggested for the second system. The statistic log dataset from the UCI repository was used. The accuracy was 92.59 percent.

3D for cardiac disease detection, the Naive Bayes algorithm was utilized by the model in ultrasound images [136]. The clinical datasets were obtained from one of the major research institutes in Chennai for diabetes. In all, there are around 500 cases in the database, each with 11 different characteristics. The studies were carried out with a 70% split of percentages utilizing the weka data mining software. 86.4% of the time,

the proposed model was correct in its classification of the datasets it tested. On average, a recall rate of 74 percent, a precision rate of 71 percent, and an F1 measure of 71.2 percent.

In the realm of classification techniques, Ada Boost stands out. It generates many relatively inept learners and then combines them into a single robust one. This process exemplifies how a single learner's mistake can have repercussions for the rest of the class. As long as the weak learners' requirements are not met, the process of creating them and adjusting sample weights will continue. The function is determined by the classifier's voting strength and the classifier's attributes. To determine the final prediction of the test sample, the label with the maximum number of votes is used. As a result, Ada Boost favors classifiers that correctly predict over those that misclassify.

They proposed a hybrid method in which two machine learning algorithms supporting Swarm optimization and Ant colony optimization are effectually combined with the wrapper approach [137, 124]. They are using the UCI machine learning dataset repository. The various data are mixed with the cleveland clinic foundation for the experiment. After applying the ML model, the classification accuracy was 89.65% in the result of heart diseases [138]. The proposed localization algorithm uses form predicting in the ultrasound image in the b-mode approach to detect the heart artery wall to determine unknowns the about the region and wall tracking method in the ROI [139]. The proposed approach presents various ways to achieve the favorite result optimization with a different machine learning algorithm.

In some cases, most of the papillary muscle was part of the myocardium. In addition, the boundaries of the identified regions were not as smooth as the surface delineation supplied by the other two categories of approaches, which will be discussed in the following two sections. Lastly, the classification mostly uses how the intensity is spread out to separate areas that are close together. This makes me wonder if the classification can be used to divide up the whole heart using the cardiac US, where the intensity distributions of the four chambers and the large blood vessels can be almost the same. As far as anyone knows, this attempt hasn't led to any work.

2.4 Segmentation

Anatomical structures can be differentiated from each other in medical images. Unhealthy tissues, such as lesions in the heart, can also be detected with the help of segmentation. Furthermore, semantic segmentation has proved improved accuracy in

the healthcare industry. Irrespective of various challenges Hence. It is necessary to modify classical pipelines to accommodate the imbalance in data labels. CT scans, ultrasounds, X-rays, and magnetic resonance imaging (MRIs) are some of the technologies used to produce medical images. Integrated biomedical images provide 2D and 3D volumes that may be handled in a variety of ways. The simplest method is to use 2D FCNs, such as Darknet53 [139], and U-net [140]. among each slice of the region considered independently of the others. In terms of myocardial and vascular segmentation, the majority of strategies in recent years are on statistical shape models in conjunction with sensitivity distribution methods [141-143]. In reality, several of the heart segmentation challenges are based on statistical approaches, such as [144], which offers an unsupervised methodology based on phase separation between healthy hearts and lesions.

Segmentation is a process used in medical image processing to divide an image or a group of images into different areas or to define the edges of these areas. Most of the time, the segmented regions have important information for clinical uses. Using segmentation, the desired information may be extracted and separated into new images, each of which has some hidden information. After introducing threshold, edge detection, and region-growing methods in this section, we will move on to discuss their applications.

a. Threshold

The intensity values can be used to find a region of interest by setting a threshold value. Thresholding successfully necessitates that no sections in the image histogram have overlapping their intensity. This criterion is excessive for the segmentation of cardiac ultrasound pictures [145]. Separating the blood pool into different parts of the body is another example. The intensity ranges of the atrial and ventricular blood chambers can overlap. so, it is needed to know the anatomy.

b. Edge detection

Edge detection of the start and end of a region by using the intensity values of neighboring points. A review and comparison of several algorithms can be found in the manuscript [146]. In general, edge spots are deemed to have a high magnitude of intensity gradient. Local maximum gradient sites with zero second directional derivatives were thought to be edge candidates in [147]. To safeguard the real edge's low gradient value, a technique called hysteresis can be used to make detection easier

and more resilient in the face of noise. Hysteresis also helps to make smoother, more continuous curves so that edges don't look like dashes. It is possible to employ edge detection algorithms as a boundary profile when searching for boundaries [148].

c. Region-based segment

Region growing uses the fact that pixels in the same area are connected and have similar intensity values. This is also thought to be happening as the area grows. Based on mean, variance, shape smoothness, or some combination of these, the algorithm increases the volume by assimilation of nearby pixels that meet a predetermined similarity requirement [149]. The most important step in this method is to define the similarity criterion, which usually needs to be parameterized to avoid the leaking problem. As a result, various observations may yield different outcomes from the same seed point.

d. Other segmentation.

A proposed multi-objective mathematical programming (MOMP) model for predicting the cardiac medical image segmentation to evaluate the robustness of the suggested model by including the synthetic image space of the object with various concavities and Gaussian noise [151]. The article used a hybrid approach that involves combining different techniques in the exploiter that quickly diagnose and predict heart diseases [152].

The diseases were diagnosed in ultrasound heart images diseases using the RCNN mask mode algorithm used in the Ultrasonic image [153]. Therefore R-CNN is a powerful principle in the machine learning model to find ROI, using this RCNN algorithm provides the classification and segmentation of 95.5% of segmentation with ROI weight edge boundary [154]. With 132 ultrasonic images and a CNN model, Deep Learning methods are used to divide images of heart disease. However, they used the vessel segmentation liver image for the implementation in that the result is 99.14%-pixel accuracy with a mean accuracy value of 79.16% and an IOU value is 69.62% achieved according to the model.

An individual with heart disease risk gets predicted [155] by applying machine learning algorithms to training data. But, the accuracy during prediction was not improved by the CNN algorithm. An efficient intelligent medical system was introduced in [156] for identifying the patient's heart condition to perform an accurate cardiovascular disease diagnosis. But the time consumption was not reduced by the intelligent medical system.

e. Summary

The summary of this study is to find the existing technologies which are used to identify heart diseases and to handle various deep learning models in the last few years. The state of the existing program is how they used the DL model to diagnose CVD diseases using ultrasound images and predict heart issues in the medical field. Based on the survey intention find the existing drawbacks to overcome and predict the problem in images to diagnose heart issues in the earlier stage.

This research is intended to provide a summary of DL's medical uses. The analyses and papers that were published used a wide variety of data sources. When it comes to the literature on artificial intelligence in healthcare, medical imaging is one of the most well-studied fields. Because the deep learning convolutional neural network represents the evolution and improvement of many computer vision algorithms. Researchers in other studies used NLP algorithms to make sense of patient records and ultrasound images. Subsequently, a review was presented on the uses of AI systems in cardiovascular diseases, including a variety of studies that made use of the CAMUS [156, 157] dataset, which is generally recognized as one of the most noticeable resources for DL applications in cardiovascular medicine. Following that, the dataset was introduced, cleaned, and sorted so that it could be used with various classification methods. This includes the likes of the naive bayes, logistic regression, decision tree, support vector machine, and artificial neural network. Metrics were later developed to explain the best model.

Chapter 3

Kushner Stratonovich Dice Segmented Haar Wavelet Deep Convolutional Neural Learning Method

3.1 Introduction

At present Machine learning techniques implemented in most of the models are developed for classification and segmentation, even if they will not satisfy the aim of the healthcare department. So various versions of the deep learning model will be used for object detection in the medical diagnosis field as it becomes mandatory in the medical field. Filters and classifiers were created from the spectral features calculated in the first step to categorize the available tissue types. These filters contain classification of the images that are unsupervised machine learning techniques. Several such network models are included in yolo and sorensen, and the model's final prediction is determined by a simple majority of their delegates. Several research studies have incorporated random yolo & sørensen dice classifiers for tissue characterization and had shown successful classification rates. The number of trees was selected to minimize the out-of-bag error of the classifier. The proposed model classifiers were generated using the spectral features computed for three different bandwidth ranges of the power spectrum. The classifiers were trained with seventy-five percent of the available 180 ROI samples and tested on the remaining ROIs. The information was arbitrarily split together into a training set as well as an evaluation set.

The main contribution of the KSDSC model is given as follows:

- The KSDSC model aims to perform efficient heart disease prediction with higher accuracy and lesser time consumption. This method has five layers namely an input layer, three hidden layers, and an output layer.
- Ultrasound images are considered as input. After that, image pre-processing is carried out in KSDSC model using kushner stratonovich filter to reduce the noise pixels from the ultrasound image.

- Sørensen dice image segmentation process in KSDSC model divides the preprocessed image into several segments. Then, multiple features are extracted using haar wavelet transformation in KSDSC model from the segmented image.
- The output layer employs the softmax yolov4 darknet53 activation function in KSDSC model to match extracted features with disease features for heart disease prediction detection.

3.2 KSDSC model implementation

Heart disease has received large attention in medical research. Heart disease diagnosis is a difficult task at present the automated prediction system helps the heart condition of the patient for further treatment. Many researchers carried out their research on different heart disease prediction methods, with minimal accuracy and time consumption methods. To address these problems, kushner-stratonovich dice segmented haar wavelet (KSDSC) deep convolutional neural learning model is introduced [158]. The main objective of the KSDSC model is to perform an efficient heart disease prediction with higher accuracy and lesser time consumption. The structural diagram of the KSDSC model is illustrated in figure 3.1.

Deep convolutional neural network (DCNN) is a deep learning neural network that processes the images as structured data arrays. Convolutional deep neural networks are used in computer vision for visual applications like image classifications [159]. The designed DCNN comprised six layers, namely an input layer, three hidden layers, and one output layer for performing efficient heat disease prediction. The input layer collects the number of (USI) ultrasound images '$USI_1, USI_2, USI_3, \ldots USI_n$ ' [160]. The input layer '$I(t)$' of the network is formulated as,

$$I(t) = bias + [\sum_{i=1}^{n} USI_i * \omega e_{in}] \quad (3.1)$$

From equation (3.1), 'USI_i' represents the ultrasound images. The bias value is considered as '1'. 'ωe_{in}' denotes the weight value at the input layer. Then, the input is transmitted to the hidden layer 1, where image pre-processing is carried out.

Kushner-stratonovich filtered image denoising is performed in the hidden layer_1 to reduce the noise of ultrasound images for image quality enhancement. kushner-stratonovich filtered image denoising '$KSFID(x)$' is used to eliminate the

noise drawn based on the conditional probability density of stochastic non-linear dynamic system [161]. $KSFID$is formulated as

$$KSFID(x) = \sum_{i=1}^{n} USI_i h(x,t)dt \quad (3.2)$$

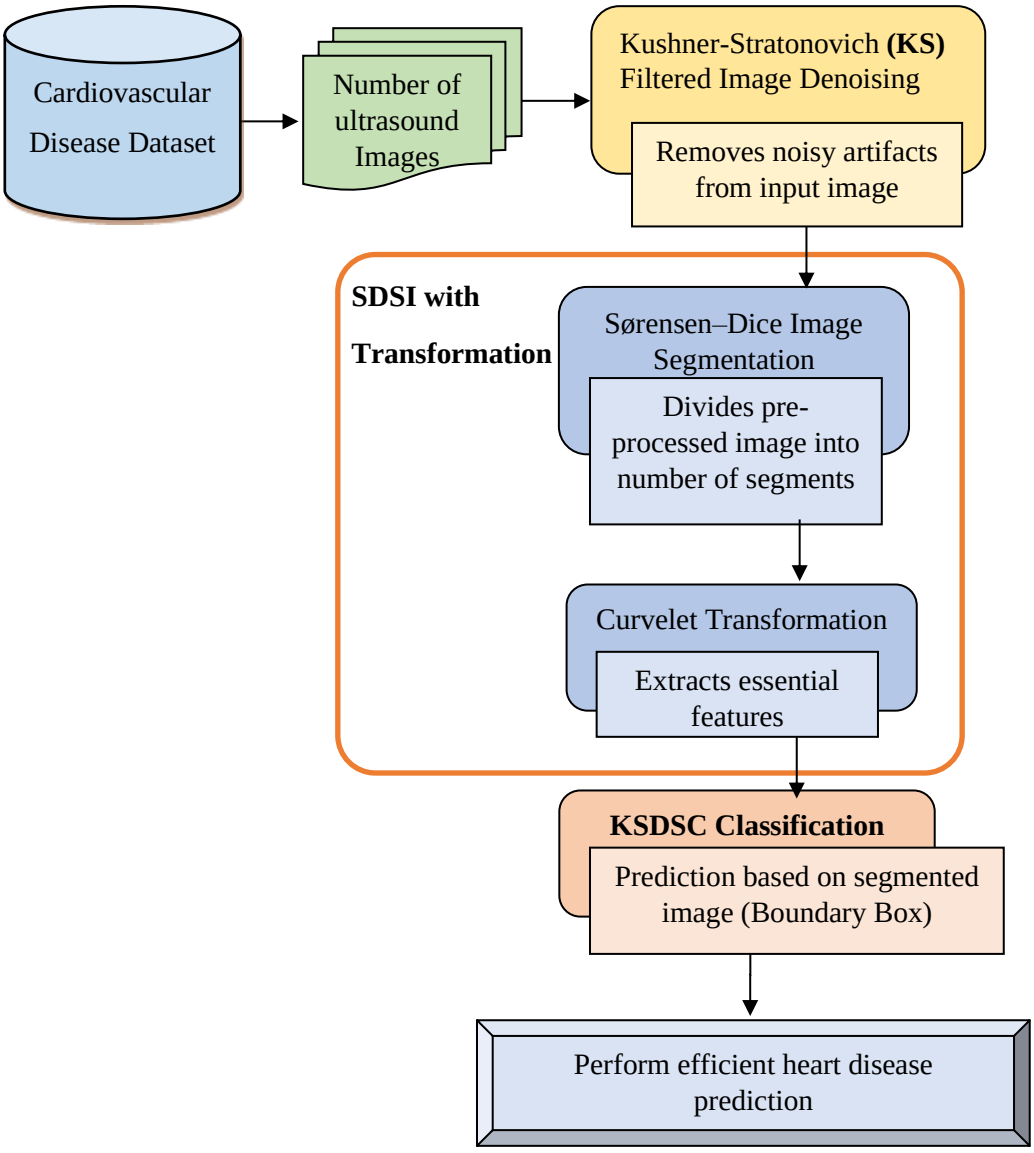

Figure 3.1 Workflow diagram of the proposed KSDSC model

From (3.1), '$KSFID(x)$' denotes the denoised image. where '$h(x,t)$' denotes the conditional probability density function. Figure 3.1 proposed workflow diagram represents the denoised image sent to the hidden layer 2. In that layer, the image processing technique known as sørensen dice image segmentation is used to segment the denoised image into sub-images based on characteristics like location, color, shape, texture, and even direction of movement.

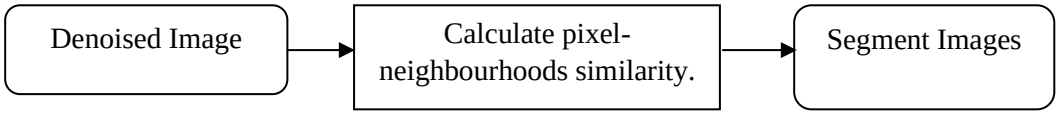

Figure 3.2 Sorensen–Dice image segmentation

Figure 3.2 describes the Sorensen dice similarity function taken with input denoised image. The sørensen-dice similarity index was used by the similarity function to identify each of the input image's pixels [162-164]. The linear relationship among both pixels in an input image is established by the similarity index. Specifically, it is calculated as

$$SDSI = \frac{2|Pixel_x \cap Pixel_y|}{|Pixel_x| + |Pixel_y|} \tag{3.3}$$

From equation 3.3, '$SDSI$' symbolizes the thesørensen–dice similarity index. '$Pixel_x$' denotes the image pixels, and '$Pixel_y$' denotes the neighboring pixel of the denoised image. The coefficient is the similarity value range between 0 and $1(0 \leq SDSI \leq 1)$.

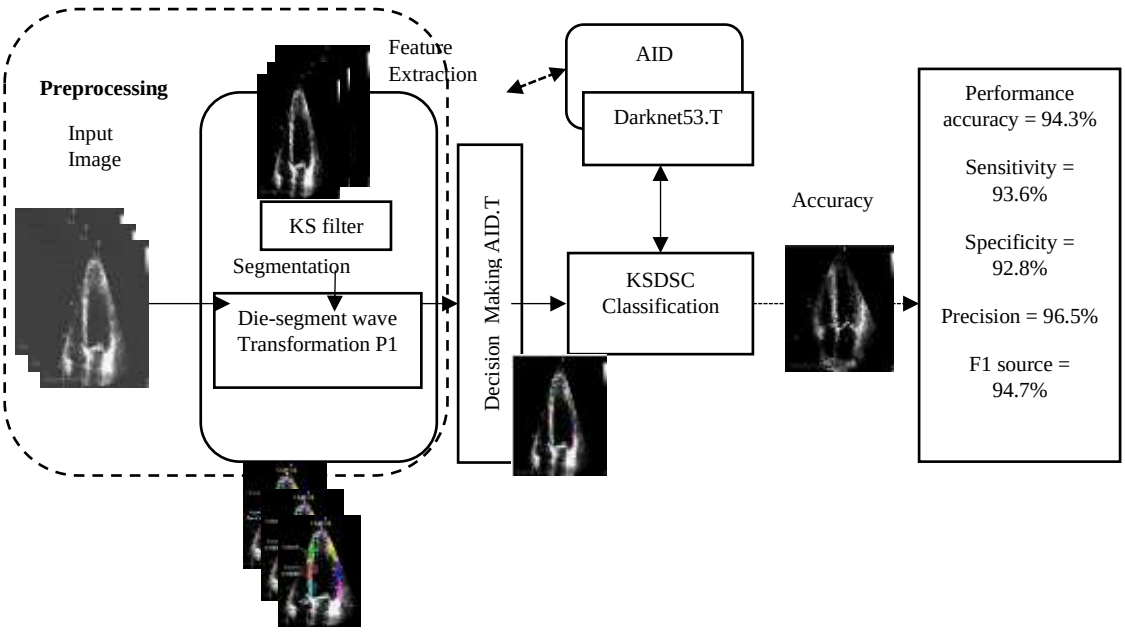

Figure 3.3 KSDSC proposed Model Architecture diagram

The KSDSC method performed a die-segment haar wavelet of an input image by convolution with the (KS) kushner stratonovich filter set illustrated in figure 3.3. For

wavelength as well as angle KS filters are used. KS filters split images into a variety of directions and scales. Assuming that the input ultrasound Image 'USI' is represented by 'M' pixels. 'F' stands for the fourier transform. The linear KS curvelet transform is denoted by 'W o'. 'S i' and's i' represents the filtering in the frequency & spatial domains, respectively [165].

Depending on the similarity value, the pixels are divided into different regions in KSDSC model and the segmented image is sent to the hidden layer. In this layer, we apply HAAR wavelet transformation to both the test images and the background image to extract relevant features. The images are broken up into several horizontal and vertical layers. since ultrasound images are transformed through image decomposition, this method is frequently used. The letter 'X' represents the geographical coordinates. Curvelet transformation eliminates unnecessary representations due to its non-orthogonality. C_i represents the curvelet coefficient, and 'i' stands for the individual channels that result from the down-sampling convolution of the ultrasound image USI_i. This is how it is formally depicted.

$$C_i = \downarrow_i (S_i \otimes USI_i) \tag{3.4}$$

The representation of symbol '⊗' in equation 3.4 is the convolution operator. Down sampling by 2 and the result of every level provides the four sub-blocks for the following levels. Hidden layer 4 builds the feature representation after extracting the image's statistical characteristics such as form, colors, contrast, texture, and edges with different pixel attributes. The form is described as a physical structure and the ideal border is obtained. The picture-segmented image gives more information on the image's spatial illustration. Color characteristics are detected by obtaining RGB data from the source images. Each hidden layer's result is represented as,

$$Hidden(t) = \sum_{i=1}^{n} USI_i * \omega_{in} + [Hidden(t-1) * \omega_{hh}] \tag{3.5}$$

A hidden layer result is represented by 'Hidden(t)' in equation (3.5). 'ω_{in}' represents the weight of the hidden and input layers. 'Concealed(t-1)' represents the previously hidden layer's content. The weight between hidden layers is denoted by 'ω_{hh}'. The result of the hidden layer is then sent to the output layer. Within the layer that implements the KSDSC model and the softmax activation function, classification takes place. The KSDSC model employs the softmax yolov4 darknet53 activation function

better understand and accurately evaluate the training picture and illness image for cardiac disease [166]. It is given as,

$$Softmax_{fun} = \frac{e^{x_i}}{\sum_{j=1}^{K} 1+e^{x_i}} \tag{3.6}$$

where 'Softmax$_{fun}$' signifies the softmax yolov4 darknet53 activation function in equation (3.6). 'x_i ' represents the training segmented picture. An image is identified as having heart disease if it looks like both the training picture and the testing disease image. If not, it is simply called a standard picture. The KSDSC model feature extractor is designated as,

$$Output(t) = \omega_{oh} * Hidden(t) \tag{3.7}$$

where 'Output(t)' represents the output layer 'ω_{oh}' represents the weight between the output and hidden layers. As a result, it helps improve the predictability of cardiovascular disease, both in terms of accuracy and false positives. The algorithmic process of kushner-stratonovich dice segmented haar wavelet (KSDSC) Deep convolutional neural learning is described in algorithm 3.1

3.2.1 Curvelet Transformation

Curvelet transformation attains the sub-bands of ultrasound images with different resolution levels. Curvelet transform is a transformation technique with multiresolution analysis. Curvelet transform includes the directional parameters with needle-shaped elements and high directional specificity [167]. Curvelet coefficients are obtained from scaling and windowing functions. Equation 3.8 window frame width 'W' and length 'L' is selected by,

$$\text{Width} = (\text{Length})^2 \tag{3.8}$$

And it is termed as curvelet scaling law. The sub-band decomposition of the ultrasound image is carried out by,

$$\boldsymbol{p}b \rightarrow (\text{lp}f, \Delta 1p, \Delta 2p, \ldots) \tag{3.9}$$

'pb' represents the pixel bands. 'lp$f$' represents the low pass filter, and 'Δ1' and 'Δ2' denotes the high pass filter in equation 3.9. In this way, the image features are extracted. The output of the hidden layer is,

$$KSFID(x) = \sum_{i=1}^{n} USI_i h(x,t)dt \tag{3.10}$$

A. Investigational Analysis

Using python exploratory analysis of the proposed KSDSC layers is established as follows existing related approaches, such as the Jaya Algorithm with Red Deer Algorithm (J-RDA) [168-170] and the Combined Reinforcement Multitask Progressive Time-Series Networks (CRMPTN) [171-174] model, for predicting the presence or absence of heart disease. Most ultrasound Images are gathered from the CAMUS database [175]. The input image is pre-processed to eliminate the noisy artifacts and attain quality-improved ultrasound images. Then, the input pre-processed images are segmented. The texture, color, shape, and intensity features are extracted from segmented images [176-179]. At last, classification is performed with help of the extracted features for performing the heart disease prediction.

B. Impact of prediction accuracy

Prediction accuracy is defined as the ratio of the number of ultrasound images that are correctly identified as a diseased image or normal image through segmentation and classification process from given ultrasound input images.

The prediction accuracy is calculated as,

$$Pre_{Acc} = \left(\frac{Hit\ rate}{Total\ No\ of\ inputs}\right) * 100 \tag{3.11}$$

Here, the hit rate is calculated by dividing the number of cache hits by the total of cache hits as well as cache misses. 'Pre_{Acc}' represents the prediction accuracy in Equation (3.11). The prediction accuracy is expressed in percentage (%).

Algorithm 3.1 presents the step-by-step process of heart disease prediction in the KSDSC Models with improved accuracy. At the neural network, the raw ultrasound images are used as inputs.

Algorithm 3.1 Kushner-Stratonovich dice segmented curvelet (KSDSC) deep convolutional neural learning

Input: Database, Ultra Sound Image $USI_2, USI_3 \ldots . USI_n$, *Threshold, AID: {ai decision making, T-net}*

Output: Boundary Box for Image tested image.

Begin

Step 1: Initialize the number of ultrasound images$USI_2, USI_3 \ldots . USI_n$

Step 2: **for each** Ultrasound image

Step 3: Eliminates the noisy pixels through KS Filtering

Step 4: Partition the images through sørensen–dice similarity function

$$SDSI = \frac{2\left|Pixel_x \cap Pixel_y\right|}{|Pixel_x| + \left|Pixel_y\right|}$$

Step 5: Measure the similarity between pixels

Step 6: Extract the feature using curvelet transformation (AID.T)

Step 7: Apply (softmax, yolov4+, darknet53→T) [f(x) = x ⋅ tanh(ς(x))] activation function for extracted features

Step 8: **if** $(Softmax_{fun} > th)$**then**

Step 9: image is abnormal

Step 10: **Else**

Step 11: image is normal

Step 12: **End if**

Step 13: End for

End

The hidden layers then perform pre-processing, image segmentation feature extraction, and classification [180]. Based on the study, heart disease prediction is performed at the activation function, where the time complexity is lower.

i. Accuracy

Efficiency is a classification system measure that indicates the number of correct predictions. Accuracy is determined by comparing the actual number of hits to the total number of predictions. Accuracy can be represented as True/False when dealing with binary classification.

$$\text{Accuracy} = \frac{\text{TN+TN}}{\text{TP+TN+FP+FN}} \tag{3.12}$$

Equation 3.12 shows the Accuracy, TN represents the true negative, TP represents true positive, FP represents false positive, and FN represents false negative.

The performance efficiencies of the future model KSDSC with other existing models of the current segmentation and classification algorithms. The following terms are used to analyze the efficiency of the proposed work.

- Prediction time
- Prediction accuracy
- Sensitivity, Specificity, and F1-Score

ii. Precision

Precision is found by dividing the number of true positives by the number of things that were predicted to be positives.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{3.13}$$

iii. Recall (Sensitivity)

The recall rate, also known as the real positive rate or sensitivity, is determined by dividing the number of actual positive results by the total number of results that should have been positive.

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{3.14}$$

iv. F1- Measure

The F1 score is an alternative way to measure how good a machine learning model is at making predictions. It does this by looking at how well it does in each class rather than how well it does overall. F1 score is used a lot in recent research because it combines two metrics that are often at odds with each other: a model's precision and recall scores.

$$\text{F1} - \text{measure} = 2 * \frac{(\text{Recall} * \text{Precision})}{(\text{Recall+Precision})} \tag{3.15}$$

v. Convolution matrix

Convolution matrix for the proposed model KSDSC. Confusion matrices are tables used to measure a classification algorithm's efficacy. The results of a classification algorithm may be seen and reported using a confusion matrix.

Table 3.1 Confusion matrix based on the patient's records and actual class

	Number of Records	True Outcome Patients having CVD disease	
		P (Patients having CVD)	**N (Patients who do not have CVD)**
Actual Class	**P (Patients having CVD)**	TP (Patients who are correctly diagnosed with CVD	FP (Patients who are having CVD but are wrongly diagnosed as normal)
	N (Patients do not have CVD)	FN (Patients who are normal, but diagnosed with CVD)	TN (Patients who are normal and diagnosed as normal)

Table 3.1 shows a confusion matrix with healthy tissue as benign and diseased tissue as malignant. This evaluation process is used to categorize the patient's possible heart disease cases so that accurate predictions can be made.

3.3 Results and Discussion

In this section, the performance of the KSDSC Model and the existing related approaches namely the J-RDA and CRMPTN model are discussed with three metrics namely prediction accuracy, false positive rate, and prediction time [181].

Table 3.2 Tabulation of prediction accuracy

Number of US Images (Number)	Prediction Accuracy (%)		
	J-RDA	**CRMPTN model**	**KSDSC Model**
50	40	60	80
100	70	80	90
150	73	87	93
200	70	85	90
250	76	84	92
300	80	87	93
350	83	89	94
400	80	85	95
450	84	89	98

Table 3.2 describes the experimental results of prediction accuracy along with the number of ultrasound images in the range of 50 to 450 from the input dataset. The attained results of prediction accuracy using the KSDSC model are compared to the two conventional methods namely the J-RDA and CRMPTN models. The proposed KSDSC model efficiently achieves higher prediction accuracy than the existing methods. 50 ultrasound images are considered for conducting the experiments [182]. By using KSDSC model, the observed prediction accuracy is 90% while the prediction accuracy of the conventional first J-RDA method and second of CRMPTN is 70% and 80% respectively. The various prediction accuracy results are achieved for every technique. The graphical representation of prediction accuracy is illustrated in the 3.4 figure.

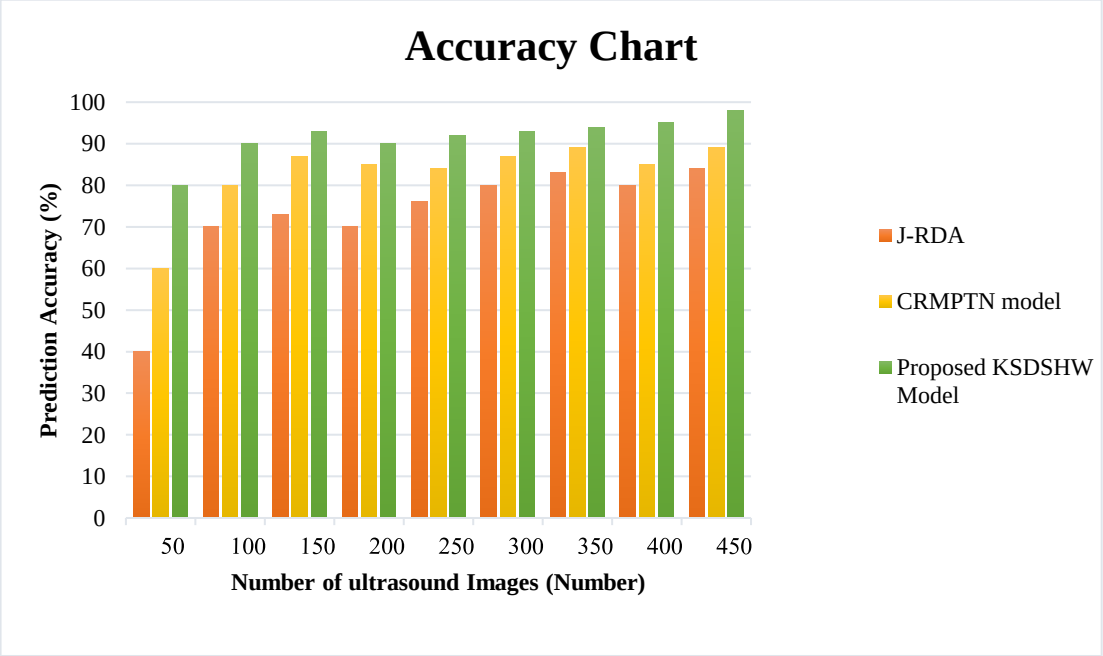

Figure 3.4 Measurement of prediction accuracy

Figure 3.3 illustrates the performance analysis of prediction accuracy along with the number of ultrasound images ranging from 50 to 450. The number of images is considered in the horizontal direction and the prediction accuracy is attained at the vertical axis.[183] The green color column denotes the prediction accuracy of the KSDSC model whereas the orange color and yellow color represent the prediction accuracy of the J-RDA and CRMPTN model respectively. The prediction accuracy of the proposed KSDSC model is higher when compared to the conventional methods. This is due to the application of a deep convolutional neural classifier in the KSDSC model.

The Kushner-Stratonovich filtered pre-processing techniques improve image quality by eliminating noise pixels. The Sørensen–dice image segmentation divides the pre-processed image into several segments. After that, the features are extracted to perform accurate classification for heart disease prediction. As a result, the proposed KSDSC model increases the prediction accuracy by 30% when compared to J-RDA and 11% when compared to the CRMPTN model respectively.

A. Impact of false positive rate

The false positive rate is computed as the ratio of many ultrasound images that are incorrectly predicted as diseased ultrasound images or normal images through performing segmentation and classification processes from given input ultrasound images [184]. The false positive rate is calculated as,

$$FP_{Rate} = \left(\frac{Number\ of\ images\ incorrectly\ predicted}{Number\ of\ US\ images}\right) * 100 \quad (3.16)$$

From (3.16), 'FP_{Rate}'symbolizes the false positive rate. The false positive rate is measured in terms of percentage (%).

Table 3.3 Tabulation of False positive rate

Number of US Images (Number)	False positive rate (%)		
	J-RDA	CRMPTN model	KSDSC Model
50	60	40	20
100	30	20	10
150	27	13	7
200	30	15	10
250	24	16	8
300	20	13	7
350	17	11	6
400	20	15	5
450	16	11	2

Table 3.3 explains the experimental results of the false positive rate along with the number of ultrasound images in the range varying from 50 to 450 from the input dataset. The achieved results of the false positive rate using the KSDSC model are compared to two existing techniques namely the J-RDA and CRMPTN models. The proposed KSDSC model efficiently attains a lesser false positive rate than the existing methods. 200 ultrasound images are considered for experimental consideration for conducting the experiments. By using KSDSC model, the observed false positive rate is 30% while the false positive rate of the conventional existing methods is J-RDA & CRMPTN value of 15% and 10% respectively. The different false positive rate results are attained for every method. The graphical illustration of the false positive rate is illustrated in figure 3.5.

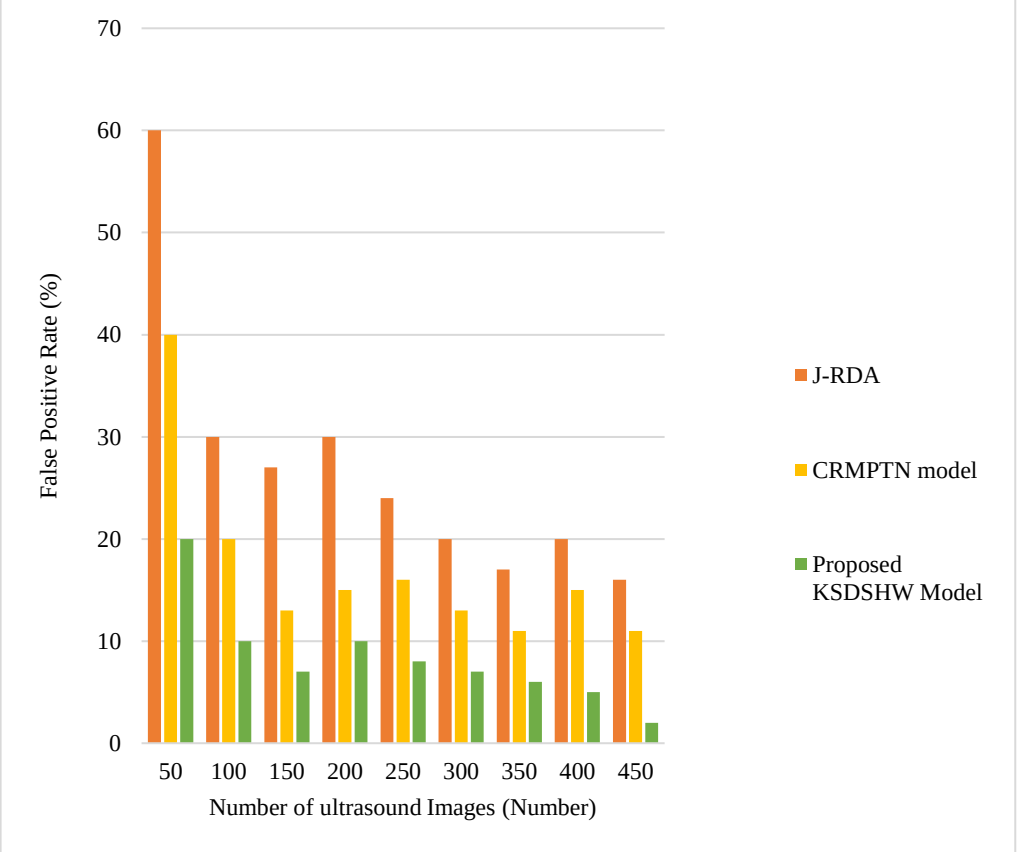

Figure 3.5 Measurement of false positive rate

Figure 3.5 describes the performance analysis of the false positive rate along with the number of ultrasound images varying from 50 to 450. The number of ultrasound images is taken in the horizontal direction and the false positive rate is considered on the vertical axis. The green color column denotes the false positive rate of the KSDSC model whereas the orange color and yellow color symbolize the false

positive rate of the J-RDA and CRMPTN models. The false positive rate of the proposed KSDSC model is lesser when compared to the conventional methods. This is because of using the deep convolutional neural classifier in the proposed KSDSC model. The Kushner-Stratonovich filtered pre-processing eliminates the noisy pixels to enhance the image quality [185]. Then, Sørensen–dice image segmentation partitions the pre-processed image into several segments. After that, the features are extracted from the segmented image for heart disease prediction with a lesser false positive rate. As a result, the proposed KSDSC model reduces the false positive rate by 70% when compared to J-RDA and 52% when compared to the CRMPTN model.

B. Analysis of Prediction Time

Prediction time is calculated as the amount of time consumed to predict heart disease from given input ultrasound images through pre-processing, segmentation, and classification task. The prediction time is calculated as,

$$Pre_{Time} = Number\ of\ MRI\ images * t(PSI) \tag{3.12}$$

From (3.12), 'Pre_{Time}' symbolizes the prediction time. '$t(PSI)$' denotes the consumed for predicting a single ultrasound image. The prediction time is computed in terms of milliseconds (ms).

Table 3.4 Tabulation of prediction time

Number of US Images (Number)	Prediction Time (ms)		
	J-RDA	CRMPTN	Proposed KSDSC
50	27	20	15
100	31	22	18
150	34	26	20
200	37	30	23
250	40	33	25
300	43	36	27
350	45	39	29
400	48	42	32
450	51	46	35

Table 3.4 illustrates the experimental results of prediction time along with the number of ultrasound images in the range varying from 50 to 450 from the input dataset. The results of prediction time using the KSDSC model are compared to two conventional techniques namely the J-RDA and CRMPTN models. The proposed KSDSC model efficiently attained lesser prediction time than existing methods. [187] 400 ultrasound images are considered for experimental consideration for conducting the experiments. By using KSDSC model, the observed prediction time is $32ms$ while the prediction time of the traditional methods first and second are $48ms$ and $42ms$ respectively. The different prediction time results are achieved for each technique. The graphical representation of prediction time is illustrated in the 3.6 figure.

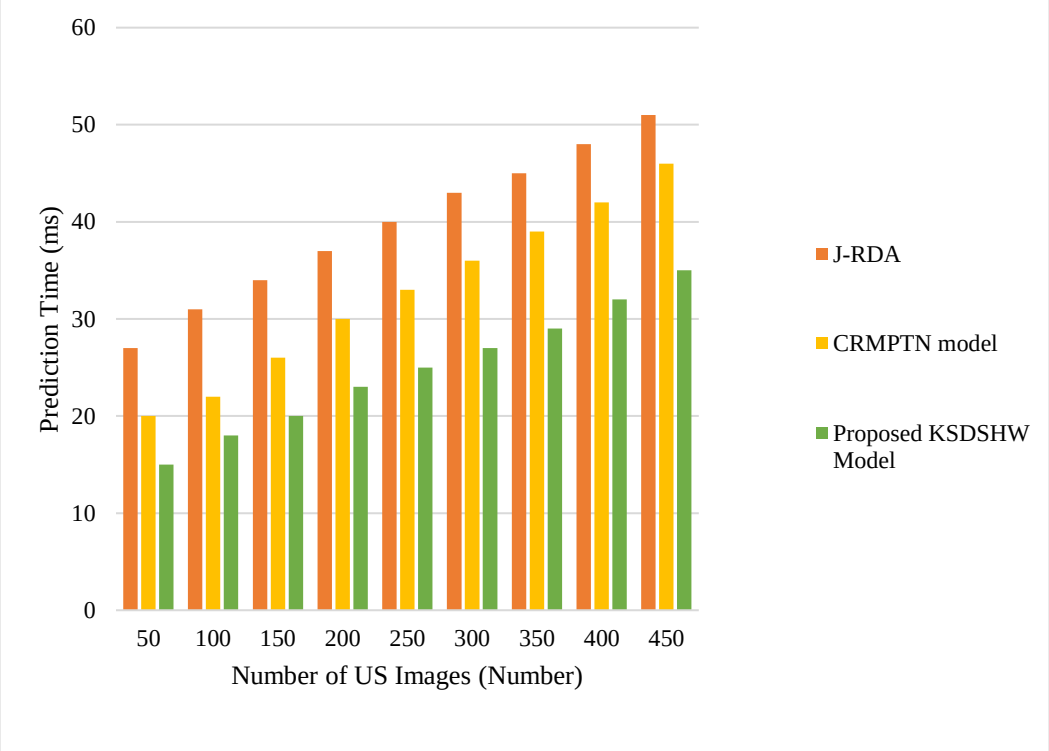

Figure 3.6 Measurement of prediction time

Figure 3.6 illustrates the result analysis of prediction time along with the number of ultrasound images ranging from 50 to 450. The number of ultrasound Images is considered in the horizontal direction and the prediction time is taken on the vertical axis. The green color column symbolizes the prediction time of the KSDSC model whereas the orange color and yellow color denotes the prediction time of the J-RDA and CRMPTN model. The prediction time of the proposed KSDSC model is lesser when

compared to traditional techniques. This is because of the deep convolutional neural classifier application in the proposed KSDSC model. In the designed model, Kushner-Stratonovich filtered pre-processing eliminates the noisy pixels to improve image quality. Then, Sorensen–dice image segmentation divides the pre-processed image into different segments. The features are extracted from segmented images for performing the heart disease prediction with lesser time consumption. Therefore, the proposed KSDSC model reduces the prediction time by 38% when compared to J-RDA [1] and 24% when compared to the CRMPTN model [2] respectively.

Conclusion

A heart disease prediction first proposed model is KSDSC model is introduced with higher accuracy and lesser time consumption. Kushner-Stratonovich filter eliminates the noisy pixel from ultrasound images to improve image quality. Sorensen–dice image segmentation process segments the preprocessed image and multiple features are extracted using haar wavelet transformation from the segmented image. The softmax yolov4+ with darknet53 activation function matches the extracted features with disease features for performing the heart disease prediction with higher accuracy and lesser time consumption. The experimental evaluation is carried out with help of a heart disease prediction dataset. The qualitative and quantitative results discussion shows that the KSDSC model has received better performance in terms of achieving higher accuracy and lesser time consumption and a false positive rate when compared to conventional methods.

Chapter 4

Adaptive Weighted Mean Filter-YOLOv4+ Deep Learning Model

4.1 Introduction

In this Chapter, the proposed model is AWMF (Adaptive Weighted Mean Filter) for pre-processing and YoloV4+ (You only look once version 4 plus) is implemented for segmentation and classification with the ultrasonic image used to identify cardio issues in that image. There are two phases to this model's execution. Pre-processing is the first stage to process the input, after which the input image can move into the second layer. A proposed filter has two stages namely noise detection and noise removal. The secondary stage is considered the preliminary part as it processes the input layer, backbone, and neck there are processes to detect and classify an object that has been identified in the affected area. Splitting the process is to increase the speed to make predictions quicker in real-time.

The main contribution of the AWMYOLOv4+ model is given as,

- The AWMYOLOv4+ model aims to perform efficient heart disease prediction with higher accuracy and lesser time consumption with an input layer, three hidden layers, and one output layer.
- In the beginning phase of filtering, an adaptive weighted mean (AWM) filter approach is used to assess whether the pixel within the input picture is corrupted or uncorrupted; if the noisy pixels are found, they are substituted by the pixel's mean value mostly in the filtering layer.
- Following the filtering operation image pre-processing is performed in the AWMYOLOv4+ model with the adaptive weighted mean filter to reduce noisy pixels in the ultrasound image.
- The second phase is going to train pictures, which are fed into the structure and analyzed on the GPU system. The Backbone and Neck then execute feature extraction and aggregation. The object detector is a combination of the recognition neck and observation head.

- Inside the AWMFYOLOv4+ models, the output layer uses the softmax yolov4+ darknet53 with Mish activation function to link wavelet transforms with clinical features for heart disease prediction detection.
- Finally, the head does the detection/prediction. As well as the produced border will have a high degree of precision.

4.2 AWMYolov4+ methodology

At present ML implemented in most of the models is developed for classification and segmentation purposes even if it will not satisfy the aim of the health care department. so various versions ML model will be used for object detection in medical diagnosis's some of the stands are provided that using and detecting is necessary use in the medical field.

YOLOv4+ is a duel as fast as EfficientDet, with the same efficiency. In comparison to YOLOv3 [188], the average precision (AP) and frame per second have increased by 10% and 12%, accordingly. The outstanding speed and accuracy of YOLOv4+'s fully documented publication are valuable contributions to the scientific community.

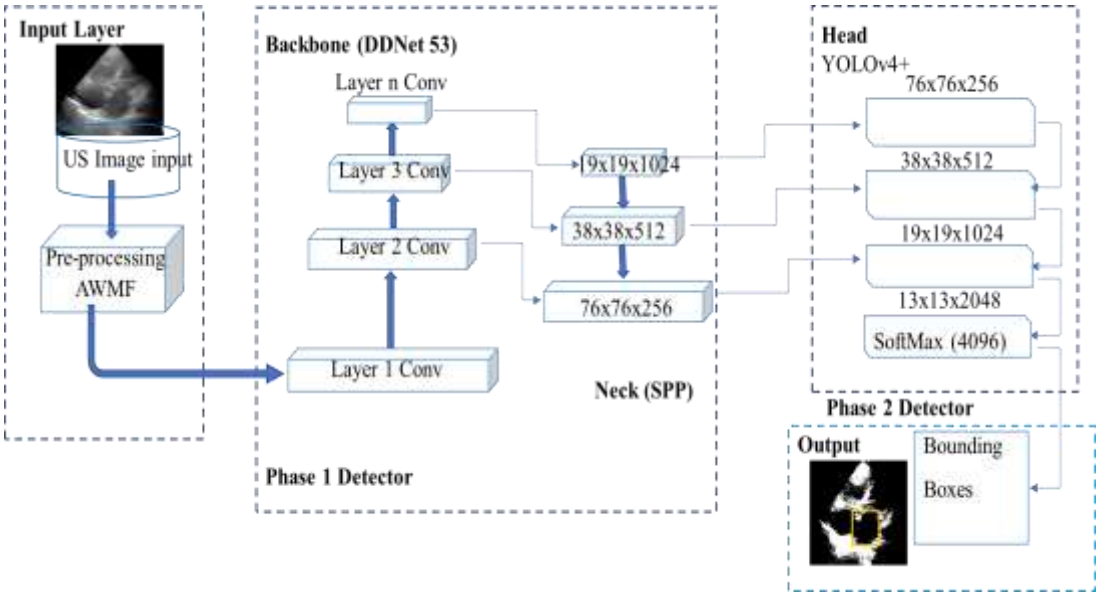

Figure 4.1 Proposed workflow architecture diagram for the AWMyolov4+

Figure 4.1 The proposed architecture shows in begins with a data pre-processing phase in which noises are removed from the input images. The second phase network is taken from the first phase's output images and it can apply the dense darknet53

convolution neural network. After that, the spatial pyramid pooling, Bof, and BoS techniques are used in the network to improve accuracy and help the head predict the object based on its category so that the detector can find its boundary box.

4.2.1 Input Layer

The input layer retrieves the image from the data source and trains it for classification and segmentation in the first layer[189]. A cardiovascular ultrasonic image is the main input source for the first layer with a good resolution even in this image have to noise.

Algorithm:4.1 Adaptive weighted mean filter yolov4+ with deep learning algorithm.(AWMYolov4+)

Input: Event Attributes, Position, and noise US1,US2, …, USn Images, transformer
Output: Mean values, Segment value, boundary co-ordinates

Step 1: Get the noisy image as X.
Step 2: Calculate the weight matrix from input image X (i,j) as W (i,j) = w1,w2, …, wn by using the method Perceptron Delta Rule in BPNN.
Step 3: Using the above weights to calculate the weighted median value as follows,
1. If the selected window, ranges from X = [x1, x2...xn], W = [w1, w2...wn].
2. WM = Med [w1 * x1 ,w2 * x2,w3 * x3…….wn * xn]
Step 4: Calculate the no of detected noisy pixels per N × N window by detecting noisy pixels using 3 × 3, 5 × 5, and 7 × 7 windows on input image X.
Step 5: Sort all the pixels NxN sub-window ‘sw’ from the original image X (i, j). and find the med value of ‘sw’.
Step 6: Replaced by weighted mean value 'WM’ with corrupted and go to step 5
Step 7: Check the corrupted and uncorrupted pixels.
Step 8: Calculate the value RGB, Intensity, and KNN values.
Step 9: Calculate the value position, x&y coordinated (or) horizontal & vertical coordinate, motion rate(0/1) for boundary box.
Step 10: The process is done iteratively.

The input image is pre-processed using the AWM filter and forwarded to the module. The preprocessing steps include noise detection and noise removal [190]. Effective noise monitoring yields a binary representation in which non-noisy pixels

have values of 0 and noisy pixels are 1. Using 3 x 3 and 5 x 5 frames, the number of noise removal is calculated across all identified pixels, and then an adaptive weighted mean filter (AWM) with a given sliding window is implemented flexibly based on the existence of the number of noise removal in an area. AWM has done effective noise detection by utilizing the count of the noisy pixels, by applying a weightage of 0 to discover noisy places in 3 x 3, 5 x 5, and 7 x 7. The proposed thresholds are in the range of 1-3, 4-12, and 13 and higher depending on the number of noisy pixels.

The steps that must be followed by the initial phase of the AWMF

1) Drag 3 x 3 windows across each pixel x(m, n) at location (m, n).
2) R(m, n)=[r1(m, n), r2(m, n), ..., r9(m, n)] is the result of sorting the pixels in ascending order for every frame W3 (m, n).
3) The pixel x(m, n) is a noisy one unless only x(m, n) = r1(m, n) or x(m, n) = r9(m, n); otherwise, it is noise-free.

In the second phase, you'll select noisy pixels by processing noise candidates, with specific steps to take depending on which ones are chosen in the final stage. Reposition the 11x11 window. W11(m1,n1) across all noisy candidates at locations (m1, n1), and sort the window to get (A) as

$$R\ (m1, n1) = [r1(m1, ,n1),r2(m1, ,n1), \ldots , r121(m1, ,n1)] \quad (4.1)$$

- Compute the distance vector D(m1, ,n1)=[d1(m1, n1), d2(m1, n1), …, d120(m1, n1)] where di(m1, n1) = r(i + 1)(m1, n1) − ri(m1, n1).
- Find first the four largest distances dmax1 = di(m1, n1) , dmax2 = dj(m1, n1), dmax3 = dk(m1, n1), and dmax4 = dl(m1, n1) where dmax1 > dmax2 > dmax3 > dmax4.
- If (x(m1, n1) < wmin or x(m1, n1) > wmax), A noisy cell, x(m1, n1), can be generated with the parameters wmin = r(p + 1) and wmax = rt. p is the minimum of(i, j, k, l), and (i, j, k, l) is the maximum of (p, t). Ultrasound images are considered as input [191]. From the first step image pre-processing is carried out by AWM (Adaptive Weighted with Mean Filter) to minimize the noise pixels in US images.

$$X = [X1, X2, X3,..... Xn] \quad (4.2)$$

Then W is an array [W1, W2, W3,.....Wn].

$$WM = MED[W1 * X1,W2 * X2,W3 * X3,.....Wn * Xn] \quad (4.3)$$

Equation 4.2&4.3 represents the value 'X' stands for the source image data input, 'W' stands for the BPNN weighted correlations, and 'WM' stands for the BPNN weighted mean value. It's a Neural Network that uses Back Propagation to learn.

Algorithm 4.2: Adaptive weighted mean filter algorithm

Input: Database, Ultrasound image $USI_2, USI_3, \dots, USI_n$

Begin

Step 1: Initialize the number of ultrasound images$USI_2, USI_3, \dots, USI_n$

Step 2: **for each** ultrasound image

Step 3: Level A

Step 4: IF (Zmin < Zmed < Zmax) then

Step 5: Zmed does not make any noise. level B to

Step 6: see if Zxy is noise

Step 7: Else Zmed is a noise.

- Enlarge the frame
- Repeat Level A until...

Step 8: Zmed isn't noise.

Step 9: Level B:

Step 10: If (Zmin < Zxy < Zmax) then

Step 11: Zxy is not noise.

Step 12: The output is Zxy.

Step 13: OTHERWISE, Zxy is a noise problem.

Step 14: The output is WM.

Step 15: End for

End

This research led to the development of the YoloV4+ model. It is implemented for classification and segmentation with an ultrasonic image used to identify issues in the image. This model executes in two phases. The input is first transferred through pre-processing, and then the image can advance to the second layer. The secondary stage is the first step in this stage's transfer of the input layer, backbone layer, and neck layer, which is necessary to determine whether an object has been located in the area and to detect and categorize it. The primary reason for splitting the process is to speed up the ability to make predictions quickly and in real-time. YOLOv4+ is a duel and fast as efficientdet, with the same efficiency[192, 193]. In comparison to YOLOv3, the AP and

FPS have increased by 10% and 12%, accordingly. The outstanding speed and accuracy of YOLOv4+'s fully documented publication are valuable contributions to the scientific community.

Table 4.1 Filtration used for a certain region of segmentation technique

	Pixel Range		
No. of noisy pixels	1 - 3	4 – 12	13 - onward
Best filter	DWMF 3 × 3	DWMF 5 × 5	DWMF 7 × 7

Table 4.1 shows the region segment model's filtering operation is divided into noise detection and noise reduction phases. The result of various noise detection techniques is a binary image with non-noisy pixel values represented as 0 (zero) and noisy pixel values represented as 1. The number of noisy pixels in a region determines whether a dynamic weighted mean filter (DWM) with window sizes of 3 x 3 and 5 x 5 is used. In addition to impulse noise detection, DWM counts the number of noisy pixels by applying a weightage 0 to find noisy spots in 3 x 3, 5 x 5, and 7 x 7 windows.

4.2.2 Backbone and neck

CSPDarkNet53 is built on the DenseNet architecture. The Dense interconnectivity pattern is created by concatenating the previous inputs with the current input before going into the fully connected layers.

CSPDarkNet53 is divided into two parts

- Convolutional Base Layer
- Cross Stage Partial (CSP) Block

The image data were separated into two pieces by CSPDarknet53. The gradient adjustments from beginning to conclusion are recorded into the feature map in the first phase, which decreases the amount of calculation & memory costs while ensuring good accuracy.

This approach reduces computation costs. It has higher accuracy when compared to other ResNet models and so performs well [194]. The backbone is the deep learning architecture, it is the collection of the three-stage namely bag of freebies (BoF), bag of specials (BoS), and CSPDarknet53 Increased training costs and changing the training strategy to reduce the cost of leaving are included in the bag of freebies. Deep

learning models with increased variability and image orders are commonly used for developing computer vision and data argumentation to improve the training model. using weighted linear interpolation of two existing images, all backbone models can mix up images and labels for the augmented image. The combination of two images, noise areas, and memory corruption of labels increases the robustness of both training and generating adversarial networks [195]. When training, there is an extreme imbalance between foreground and background classes, which is what is addressed by focal loss. Most of the time, cross-entropy is used for a loss function as part of a classification problem. With this function, highly probable errors will be penalized more severely.

$$FL(p_t) = -a_t(1 - p_t)^{\gamma} \log\log(p_t) \tag{4.4}$$

By introducing a (1-pt) ^ γ gamma factor in the focal loss function, a new cross-entropy is one of the based loss functions is proposed in equation 4.4. Based on this coefficient, the correction of misclassified objects is given priority [196]. A focusing parameter γ allows the down-weighting rate for easy examples to be controlled smoothly. When γ = 0, FL is equal to CE, and as γ increases, the modulating factor's effect also increases. All too often, people think they are right when in fact they are wrong. The model may not learn from the data if it has 100% confidence in its prediction.

YoloV4+ model sort out the problem of overfitting is a major loss in the calculation in lowers or raises the upper bound to prediction and the lower value is 0.9 and its loss is 0.1 percentage. calculation of the real bounding box the position and size difference evaluate the prediction [197]. Comparing the IoU loss and l_2 loss, we see that the IoU loss optimizes the bounding box with an equal weight instead of optimizing four coordinates separately. The IoU loss might therefore be able to provide a more accurate bounding box prediction than the L_2 loss. In addition, the IoUs are naturally skewed to [0, 1] when the bounding boxes differ in size. G represents Ground truth [$\tilde{x} = \tilde{x}(\tilde{x}_t, \tilde{x}_b, \tilde{x}_l, \tilde{x}_r)$] and P as a Prediction [$x = (x_t, x_b, x_l, x_r)$].

$$l_2 \ \text{loss} = \| P - G \|_2^2 \tag{4.5}$$

$$\text{IOU loss} = -\ln\left(\frac{\text{Intersection(P,G)}}{\text{Union(P,G)}}\right) \tag{4.6}$$

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Ares of union}} \qquad (4.6(a))$$

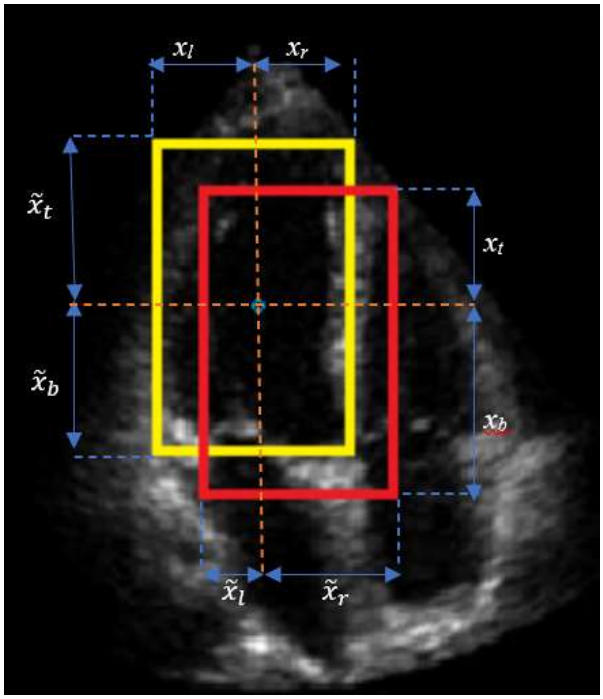

Figure 4.2 Bound box representation for prediction and ground truth region

4.2.3 Mish activation

Mish Activation can be defined by *f(x) = x*tanh(softplus(x))* as a novel self-regularized non-monotonic activation function. The soft-plus activation formula as design as $\varsigma(x)=\ln(1+e^x)$. it is necessary prerequisites for Dying ReLU and another activation phenomenon are eliminating the hazards and training to make as easier to use MISH's activation function. During the backpropagation phase, neurons that have a large negative bias will not be up-to-date, causing them to be ineffective [198]. Mish has a bound below and an unbounded range above of [≈ -0.31, ∞]. Mish structures support both activities and programs access has been provided because they preserve some negative information. Indefinite mish tries to avoid intensity, which typically causes learning to slow down due to almost zero gradients. A strong regularization effect results from being bounded below as well.

C. CSPDarknet53 and DenseNet

A cross stage partial architecture can be derived from either a DenseNet or a standard dense layer architecture by appending the previous input to the input data before

proceeding to the dense layer [199]. Each dense block in DenseNet consists of k dense layers, and the network as a whole consists of two stages namely the dense block and the transition layer. The input of the (I + 1) dense layer is obtained by concatenating the product of the i^{th} dense layer's source and the same dense layer's output. Here are some formulas that demonstrate this process as X1 = W1 * X0, X2 = W2 * [X0, X1], …. Xk = Wk * [X0, X1, ..., Xk-1].

Dense net forward, in which the kernel operator [*] represents a convolution operator, numbers [x0, x1, ...] represent the weights and output of a dense layer, and w_i and x_i are the weights and output of the i^{th} layer, respectively. The input is divided into two parts x0' and x0" in the CSP, instead of concatenating it with the i^{th} output. A dense layer x0" will pass through the first part, and a concatenated part x0' will pass through the second part. The result will be concatenated at the output of the dense layer x0"**.** Mathematically, this equals the following $X_k = W_k * [X_{0''}, X_1, \ldots, X_{k-1}]$, $X_T = W_T * [X_{0''}, X_1, \ldots, X_k]]$, $X_U = W_U * [X_{0'}, X_T]$.

The equation for Partially forwarded in the cross-stage. Consequently, different dense layers will be repeatedly exposed to copies of gradient information.

4.2.4 Neck

Feature maps from different stages of the backbone are gathered in the neck, which is a subset of the bag of specials. In a nutshell, this feature aggregator combines versions of features. In the subsequent sections the neck of the object detection in the pipeline [200]. In the second stage, the elements that were previously created in the neck are aggregated and supplied to the head for detection.

4.2.5 Spatial pyramid pooling

Spatial Pyramid Pooling (SPP) is a method that can be used to collect both fine and gross details in a given area. Using a sliding kernel with sizes of 1x1, 5x5, 9x9, and 13x13 cells, this algorithm is implemented. The output is obtained by merging maps with varying kernel sizes. SPP is widely accepted as it contains an enhanced receptive field. No matter how big the convolution layer seems to be, fixed-size components can be generated with the help of spatial pyramid pooling.

Inside the neck region, YOLOv4+ utilizes both the FPN as well as the PANet, whereas YOLOv3 utilizes only the FPN. The FPN accomplishes the second iteration

from a lower resolution to a better resolution and then concatenates with the large-size Zero crossing rate (ZCR). The PANet model utilizes bottom-up path enhancement with preceding local convolution layers via wavelet transform to reduce the overall information path across high-resolution and low-resolution elements.

A. Bag of specials

Object detection is improved by using methods from the bag of special methods, which increases computing costs by a small margin but improves accuracy. Obtaining feature maps from various stages is the dominant function of the neck. An average neck is typically made up of a couple of top-down and a couple of bottom-up paths.

In a fully interconnected network, it is necessary to have a fixed-size image; when detecting objects, it is not necessary to have fixed-size images. Due to this issue, scaling the images, which may cause a part of the object to detect to be removed and lead to a less accurate model. CNN also causes the sliding window to be fixed in size. Our convolution neural networks generate features as part of the output of their computation; these features correspond to the output of our various filters. N*K + base layer filters capable of detecting circular geometric shapes, N is the dense layer contained and K is the activator output growth rate. The filter would produce a feature map highlighting the shapes and keeping the coordinates of the shapes in the image.

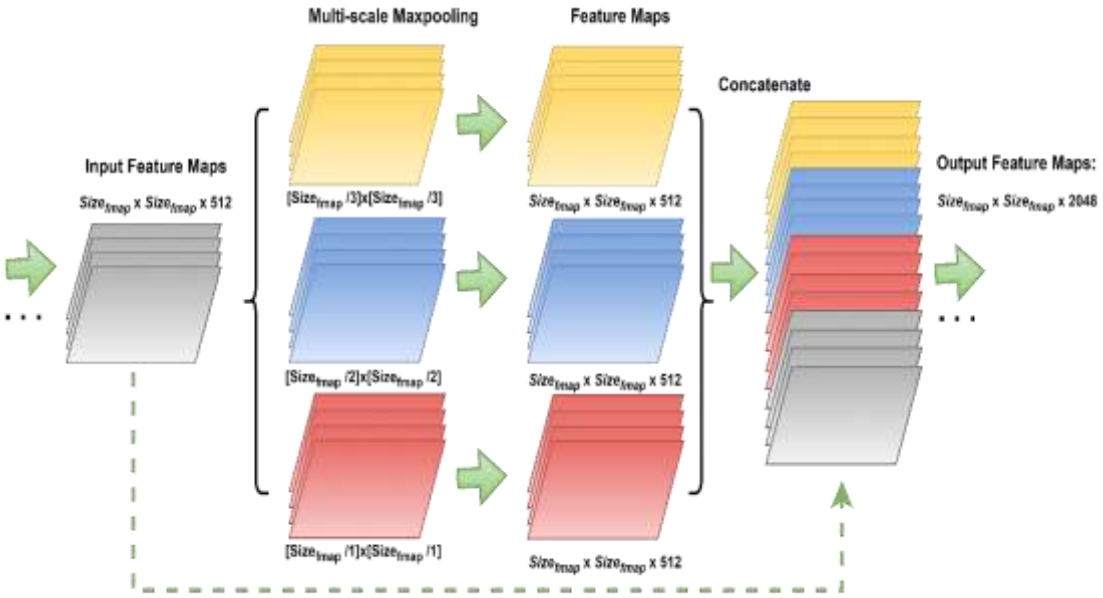

Figure 4.3 Modified SPP- spatial pyramid pooling layers

B. Spatial pyramid pooling

These Layers generate fixed-sized features regardless of the size of our feature maps. Our feature maps will be generated in different representations based on pooling layers such as Max Pooling.

The SPP consists of three levels. Consider the conv5 layer, which contains 256 features. The source from the first step is to pool the feature maps into one value (grey part in figure 4.3). There are 1, 256 fields in this vector.

- Several feature maps are then pooled to have four values (green in figure 4.3). The resulting vector is (4,256).
- The same is true for the pooling of each feature (blue part in figure 4.3). In this case, the vector has a size (of 16, 256).
- In step 3, the three vectors are created. This fixed-size vector is then concatenated to create the input for the fully connected network.

4.2.6 Head (detector)

In YOLOv4+, the overall purpose of the head is to do predictions, which involves bounded-box classification as well as regression. It gives information concerning control points of bounding boxes (x, y, h, w). The label is accompanied by the dimensions (height, width, center), as well as a prediction score (how well it predicted the label). So, each anchor box can use the YOLOv4+ head.

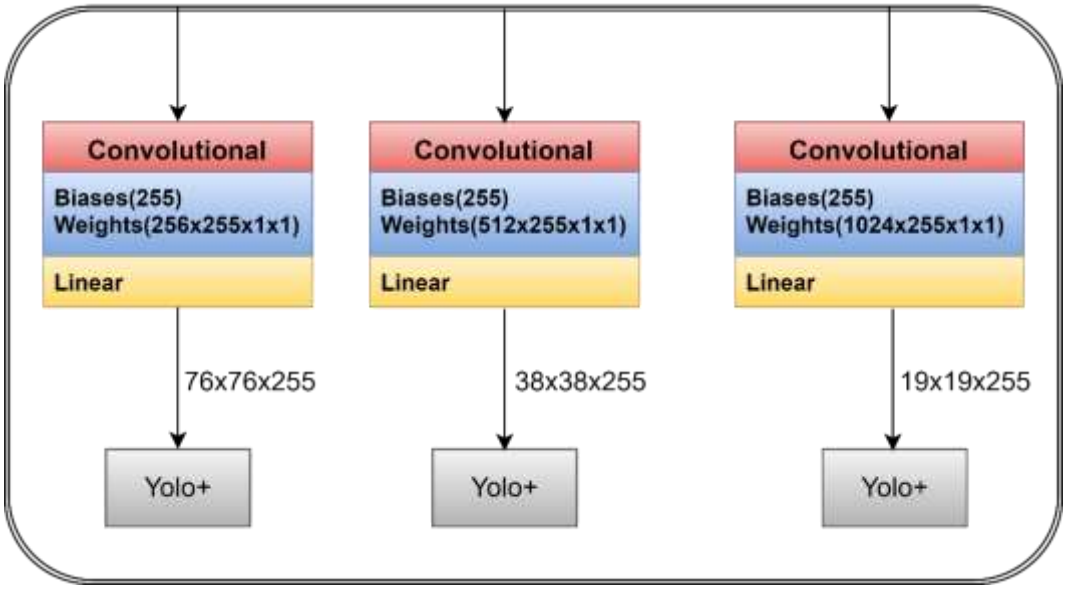

Figure 4.4 Different scales applied at the Head declaration for the proposed model

Finding bounding boxes and carrying out classification tasks are the primary functions of this section. Scores and bounding box coordinates (x, y, height, width) are identified. The x and y coordinates here represent the location of the midpoint of the b-box with respect to the edge of the grid cell. Dimensions are estimated in terms of the entire picture.

$b_w = p_w e^{t_w}$, $b_h = p_h e^{t_h}$ $b_x = \sigma(t_x) + c_x$, $b_y = \sigma(t_y) + c_y$

Figure 4.4 The output resolution of YOLOv4+ in the head area with the number of feature maps was 76 x 76/256, 38 x 38/512, and 19 x 19/1024. The loss function is the only thing that has changed. The CIoU is utilized as YOLOv4+ is loss function to determine the gap between both the ground truth and the predicted box.

This model head can be called an object detector as it finds a region where the object might be inside or outside of the boundary box but does not tell what kind of object might be present there. Our detection systems are divided into two-stage detectors and one-stage detectors; both types of detectors may be either anchor-based or anchor-free. This section will focus more specifically on the head.

4.2.7 Anchor box

Anchor boxes are used to contain many items of varying sizes in a single frame, each center of which is located within a cell. However, the grid was used to identify a specific object within a frame [204]. When anchor boxes grow in size, the ground truth and the prediction arrays shift to accommodate the new data. Here is an illustration of a scale-independent anchor box, primarily on zero.

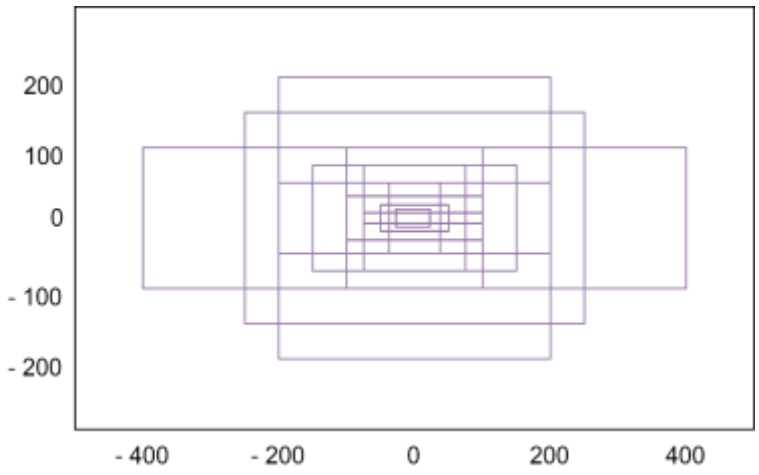

Figure 4.5 Boundary box selection based on the anchor tag

Figure 4.5 represents the anchor box designed to find many items of varying sizes whose centers are all within the same cell. If you alter the number of anchor boxes, you'll also need to adjust the length of the ground truth and prediction arrays to account for the new configuration. For assumption, a set of 85 boxes in the area, are divided into 80 different categories for making predictions (Pc, P1, P2..., P80, X1, Y1, X2, Y2). There will be a need to make any predictions for some of the 9-bounding boxes in an array of size 85x9 = 765.

4.2.8 Bag of Specials

Those methods that only change the training strategy or raise training costs (but do not affect inference) are considered a 'bag of freebies'. Based on the above diagram 4.6, there are a plethora of opportunities one can pursue, but the proposed techniques discuss only the most important ones. During one-stage detection the head executes dense prediction. In dense prediction, the bounding box would be represented by a vector with the center, height, and width of its predicted area, a confidence score, and a label.

A. Loss of CIoUs

Two new concepts are introduced by the CIoU loss, in contrast to the IoU loss. A bounding box center point distance represents the distance between the actual center point and the predicted center point of the bounding box [206]. Second, we consider the aspect ratio of the true and predicted bounding boxes, are compared the quality of a bounding box can be measured with these 3 indicators. equation 4.7 CIoU loss.

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(b,b^{gt})}{c^2} + \alpha v \tag{4.7}$$

A trade-off parameter is a positive(α) trade-off parameter, while v measures the consistency of aspect ratio, and their central points are expressed by b and b_{gt}, respectively. The Euclidean distance in ρ(·) is the diagonal length of the smallest enclosing box, while * measures the positive trade-off parameter. Equation 4.8 consistency of aspect ratio with *h* being the height and *w* being the width of the bounding box.

$$v = \frac{4}{\pi^2}\left(arctan\frac{w^{gt}}{h^{gt}} - arctan\frac{w}{h}\right)^2 \tag{4.8}$$

B. Cross mini batch normalization procedure

If the batch size is too small, the batch normalization procedure will not take place. In summary, the sample size biases both the mean and standard deviation estimates. If the sample size is too small, the distribution will probably not be completely representative

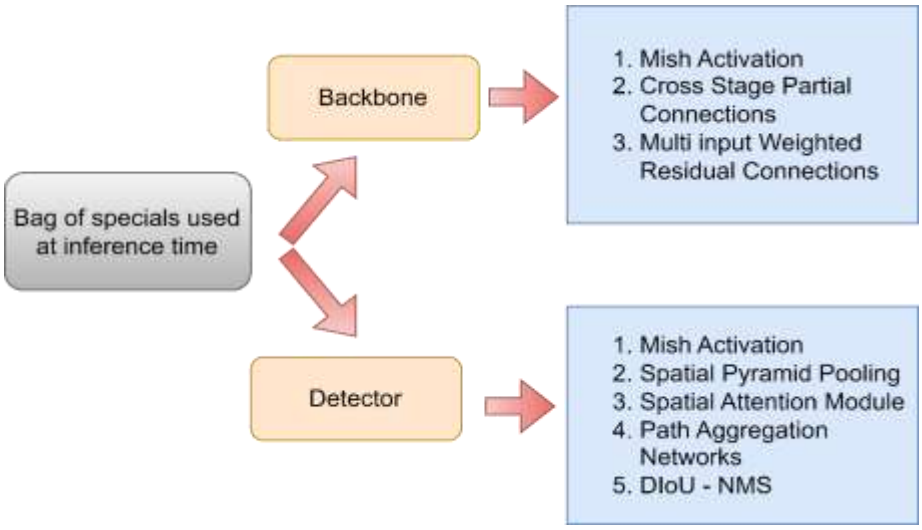

Figure 4.6 Detector and Backbone network (CSPDDarknet53) selection for the bag of specials network

4.2.9 Self-adversarial training

Self-Adversarial Training (SAT) is a novel approach that consists of two stages of forward progress and one of reversal. During the first stage of neural network processing, the original image is altered rather than the network weights [207]. In this way, the neural network poses an adversarial attack on itself, altering the original image to give the impression that it does not contain the desired object. After the second stage, the neural network is trained to detect an object based on the original label on this modified image before adding noise to the image.

A. Using multiple anchors for a single ground truth

Since convolutional networks cannot predict directly boxes associated with different ratios of images, the anchors that divide the image space based on different strategies are used. Many anchor boxes are created by convolution of the features map

to represent objects of various sizes. Using the IOU, it is decided whether to assign certain anchor boxes to objects or backgrounds based on thresholds below the surface.

B. Random training shapes

Object detectors trained on a fixed input image shape are common among single-stage detectors. Different image sizes can be used to improve generalization. (YOLO Training on Multiple Scales)

C. Bag of specials

Bag of Specials can be thought of as an extra feature that can be added to any existing object detector to make it more accurate on benchmark datasets. There are many different modules, but this research is on YOLOv4+. The selection of the methods that are used to refine the detector results could be varied depending on the architecture of the object detectors, but it is still possible to accomplish the ultimate goal.

In the latest state-of-the-art models of deep learning, attention layers are very common, in language processing. YOLOv4+ models are used to enable the recognition of the most significant features created by convolution layers and their removal. Max pooling and Average pooling are two transforms that must be applied to the output feature map of a convolutional layer to create the spatial attention module (SAM). After concatenating and passing each feature in a convoluted layer, a sigmoid function is applied to highlight the locations of the most important features.

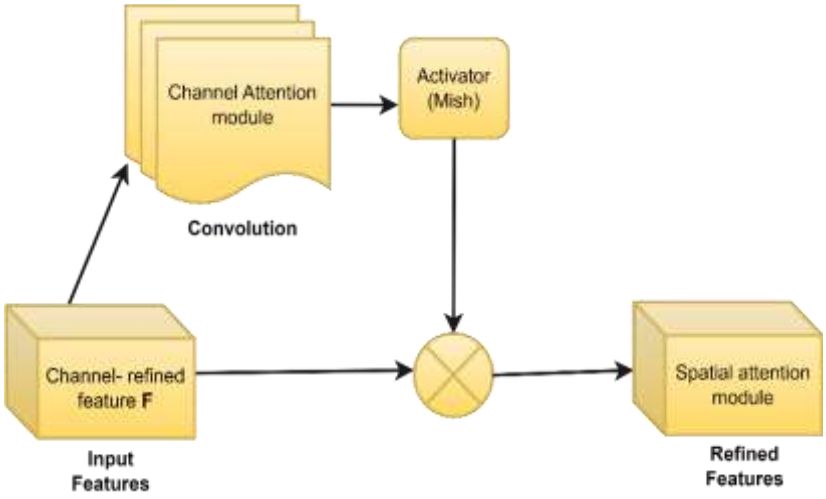

Figure 4.7 YOLOv4+ in modified SAM

A modified version of SAM figure 4.7 for the YoloV4+ that removes the layers of max pooling and average pooling is applied, random training shapes: A fixed input picture form is used to train several single-stage item detectors. The model is trained with varied picture sizes to increase generalization. (YOLO Multi-Scale Training).

D. DIoU-NMS

Non-Maximum Suppression (NMS) removes overlapping boxes representing the same object but keeps the box that is most precise compared to the true bounding box.

$$R_{\mathrm{DIoU}} = \frac{\rho^2(\mathrm{b},\mathrm{b}^{gt})}{c^2} \tag{4.9}$$

The equation 4.9 RDIoU is shown in figure 4.7 as assume b and b^{gt} are the central points of B and B^{gt}, where c is the diagonal distance separating the two boxes, and ρ(·) is the Euclidean distance between them.

$$s_i = \begin{cases} s_i, & IoU - R_{DIoU}(M, B_i) < \varepsilon, \\ 0, & IoU - R_{DIoU}(M, B_i) \geq \varepsilon, \end{cases} \tag{4.11}$$

From the equation 4.11 DIoU-threshold ε, if so, we retain the bounding box is retained otherwise, delete it.

4.3 Performance evaluation

In this section, the performance of the AWMYolov4+ Model and the existing related approaches namely the DBN and RCNN model are discussed with three metrics, NMS is specifically for testing whether the overlap rate minus the distance between the two centers is lower than the prediction accuracy, classification error, model loss, Precision, Sensitivity, and F1 Score [208]. Table 4.2 shows the hyperparameters tuned model configuration with AWMF with YOLOv4+.

In this section, the performance of the AWMYolov4+ model and the existing related approaches namely the DBN and RCNN model are discussed with some metrics, prediction accuracy, classification error, false positive rate, precision, Sensitivity, F1 score, and specificity. true positive is TP, false positive is FP, true negative is TN, and false negative is FN.

Table 4.2 Hyperparameters of Yolov4+

Parameters	Values
Learning Rate	0.001
Batch size	20
Epochs	100
Optimizer	Adam
Activation Function	Mish
Filter size	$3 \times 3, 5 \times 5, 7 \times 7$

Table 4.2 shows the assessment criteria like mean absolute error MAE, RMSE, RAE, and RRSE are used to evaluate the classification of error rate approaches, and precision; F-measure, recall, and accuracy are utilized to compute the performance accuracy of the techniques used.

A. Mean absolute error

Mean absolute error is the average of absolute errors. Absolute error is the deviation from the projected value. MAE assesses continuous variable precision. Mean absolute error measures the average mistake in a set of predictions without considering their direction. MAE is the average of the absolute discrepancies between the prediction and the observation over the test sample, weighted equally. The average error in model prediction, expressed in units of the dependent variable, can be described using either the mean absolute error or the root mean square error.

$$MAE = \frac{1}{2}\sum_{j=1}^{n}|y_i - y| \tag{4.12}$$

Equation 4.12 is shown here, $'n'$ is the number of errors, $|y_i - y|$ is an absolute error.

B. Root mean squared error

Calculating the root mean square error, also known as the root mean square deviation, is a common approach to assessing the reliability of forecasts. Equation 4.13 use the euclidean distance, it demonstrates how far apart forecasts are from the actual values that have been measured.

$$\mathrm{RMSE} = \sqrt{\frac{1}{\mathrm{n}}}\sum_{\mathrm{j}=1}^{\mathrm{n}}(\mathrm{y_i} - 1)^2 \tag{4.13}$$

C. Relative absolute error

Error in proportion to the mean of the quantity being measured is referred to as the relative error, and it is calculated as the ratio of mean absolute error to the mean of the item being measured.

$$\mathrm{RAE} = \frac{\sum_{\mathrm{j}=1}^{\mathrm{n}}|\mathrm{p_{ij}} - \mathrm{T_j}|}{\sum_{\mathrm{j}=1}^{\mathrm{n}}|\mathrm{T_j} - \mathrm{T}|} \tag{4.14}$$

Equation 4.14 represents p_{ij} is the value predicted by a certain model $'i'$ for record $'j'$(out of records), $T_{ij}$is a target value of records $'ij'$and TN is considered a perfect fit with the numerator which is equal to 0.

D. Root relative squared error (RRSE)

A simple predictor would have reduced the root relative squared error. The relative squared error is calculated by dividing the total squared error by the total squared error of the simple predictor. Equation 4.15 By calculating the square root of the relative squared error, the mistake is reduced to the same dimensions as the projected quantity. Another significance is the distinction between high-relative and root-relative mistakes.

$$\mathrm{RRSE} = \sqrt{\frac{\sum_{\mathrm{j}=1}^{\mathrm{n}}(\mathrm{P_{ij}} - \mathrm{T_j})^2}{\sum_{\mathrm{j}=1}^{\mathrm{n}}(\mathrm{T_j} - \mathrm{T})^2}} \tag{4.15}$$

E. Accuracy

Accuracy is determined by comparing the actual number of hits to the total number of predictions. Accuracy can be represented as True/False when dealing with binary classification.

$$\mathrm{Accuracy} = \frac{\mathrm{TN+TP}}{\mathrm{TP+TN+FP+FN}} \tag{4.16}$$

F. Precision

Precision is found by dividing the number of true positives by the number of things that were predicted to be positives.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{4.17}$$

G. Recall / Sensitivity/True Positive rate

The recall rate, also known as the real positive rate or sensitivity, is determined by dividing the number of actual positive results by the total number of results that should have been positive.

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{4.18}$$

$$\text{Specificity} = \frac{\text{TN}}{(\text{TN+FP})} \tag{4.19}$$

H. F1- Measure

The F1 score is an alternative way to measure how good a machine learning model is at making predictions. It does this by looking at how well it does in each class rather than how well it does overall. Equation 4.20 representing the F1 score is used a lot in recent research because it combines two metrics that are often at odds with each other: a model's precision and recall scores.

$$\text{F1} - \text{measure} = 2 * \frac{(\text{Recall} * \text{Precision})}{(\text{Recall+Precision})} \tag{4.20}$$

People who are misclassified face a wide range of prices, thus it is set cost1 at 10 and cost2 at 1. The following index is used to evaluate each classifier

$$\text{E}_\text{i} = \frac{\text{Accuracyy}_{\text{i+1}} - \frac{(\text{MC}_\text{i})}{\text{cost}_1 + \text{cost}_2}}{2} \tag{4.21}$$

Equations 4.21 & 4.22 describe an independent classifier's overall performance as measured by its mean classification accuracy (MC_i) across the training period, while attitudes toward the effectiveness of the i^{th} classifier are measured by E_i attitudes.

$$\text{MC} = \left(\frac{(\text{FP} * \text{cost}_2 + \text{FN} * \text{cost}_1)}{(\text{TP+TN+FP+FN})}\right) * 100\% \tag{4.22}$$

Table 4.3 represents the confusion matrix to understand and compare the state of the maximum possible prediction result with reality. This evaluation process to classify the patient's heart disease possible cases to predict accuracy.

Table 4.3 Confusion matrix based on the patient's records and actual class

	Number of Records	Prediction class CVD disease	
		P (Patients having CVD)	**N (Patients who do not have CVD)**
Actual Class	**P (Patients having CVD)**	TP (Patients who are correctly diagnosed with CVD	FP (Patients who are having CVD but are wrongly diagnosed as normal)
	N (Patients do not have CVD)	FN (Patients who are normal, but diagnosed with CVD)	TN (Patients who are normal and diagnosed as normal)

4.3.1 Prediction time

Prediction Time is defined as the ratio of the number of ultrasound images that are correctly identified as a diseased image or normal image through segmentation and classification process from given input ultrasound images [209]. The prediction accuracy is calculated as the Window for the prediction value of T (t + window), the timestamp where the prediction will be done. this is how we give the information of irregular time series to our models is given.

$$y(w+r) = f(t_0, \ldots, t_w, X_{i,0}, \ldots, X_{i,w}, Y_0, \ldots, Y_w, t_{w+r}, Y_{w+r}) \quad (4.23)$$

In this sense, we may have two distinct equations to model our system, depending on whether or not we use earlier response values as new predictors.

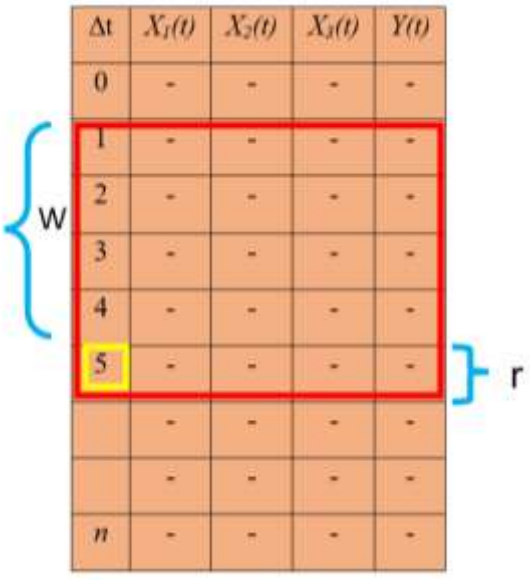

Figure 4.8 Prediction of the window for selection-based time series

Figure 4.8 Machine learning approaches for time series, red box prediction of T(t + window), the yellow box has represented the timestamp where the prediction will be done also represents the equation 4.23.

This is how to give the information of irregular time series is given to models. The function will be to flatten all the information contained in the window, which is all the values inside the W window, and the timestamp(s) to make the prediction.

Table 4.4 Comparison of prediction time with the existing model

Dataset	Prediction Time (ns)		
	DBN	**RCNN**	**AWMYolov4+**
100	120.225	130.113	95.30
200	200.350	190.88	170.450
250	256.20	277.503	218.58
350	281.631	300.336	251.205
450	312.55	349.881	298.350

Table 4.4 lists the experimentally determined prediction times as well as the range of 100 - 450 Ultrasound images from the input dataset. The findings of the AWMYoloV4+ Model's prediction of time are compared to those of the DBN and RCNN models. In comparison to current techniques, the suggested AWMYoloV4+ Model effectively reduces prediction time [210]. Consider that ultrasound images ranging from 100 to 450 are required for conducting experiments. In comparison to the standard approaches other two existing methods, which have prediction times of 70% and 80%, respectively the AWMYoloV4+ Model's observed prediction time is 40%. For each technique, varied prediction time results are obtained. Figure 4.9 shows a graphical representation of forecast time.

Figure 4.9 depicts the performance evaluation of prediction time and the number of ultrasound photos ranging from 100 to 450. The number of ultrasound Images is considered in the horizontal direction, and the prediction time is attained on the vertical axis.

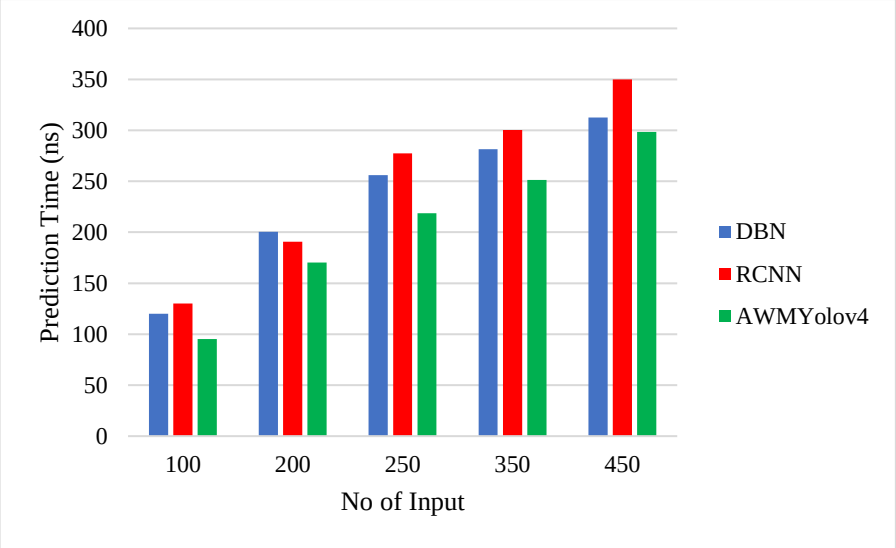

Figure 4.9 Comparisons bar chart for the prediction time

The AWMYolov4+ Model's prediction time is indicated by the greenish color column, while the DBN and RCNN models' prediction times are indicated by the blue and red coloring, respectively.

When compared to conventional methods, the proposed AWMYolov4+ Model has a greater prediction accuracy. This is because the AWMYolov4+ Model uses a deep convolutional neural classifier. By removing the noisy pixels, the Adaptive Weighted Mean pre-processing enhances the image quality. The pre-processed image was separated into several segments by the Yolov4+ backbone module, which uses image segmentation. Following that, the features are retrieved to quickly and accurately classify heart disease prediction. Therefore, compared to DBN and RCNN models, the suggested AWMYolov4+ Model reduces prediction time by 30% and 15%, respectively.

4.3.2 Classification error

A classification error occurs when a respondent fails to accurately answer a survey question, which is called a measurement error [211]. Both a false negative reaction and a false positive response are possible for nominal categorical data.

Table 4.5 lists the experimentally determined Classification Error using the image ranges between 100 - 450 ultrasound images from the input dataset. The findings of the AWMYoloV4+ Model's prediction of time Classification Error are compared to those of the DBN and RCNN models, two standard approaches.

Table 4.5 Evaluations of classification error

Dataset	Classification Error (%)		
	DBN	RCNN	AWMYolov4+
100	9.51	7.21	5.3
200	9.52	7.45	5.1
250	8.76	7.55	4.33
350	8.5	7.1	4.1
450	7.5	7.0	4.0

In comparison to current techniques, the suggested AWMYolov4+ model effectively reduces classification error. Consider that 100 ultrasound images are required for conducting experiments. In comparison to the standard approaches first and second, which have classification errors of 9.5% and 7.21%, respectively, the AWMYolov4+ Model's observed classification error is 5.3%. For each technique, varied Classification error results are obtained. Figure 4.10 shows a graphical representation of forecast time.

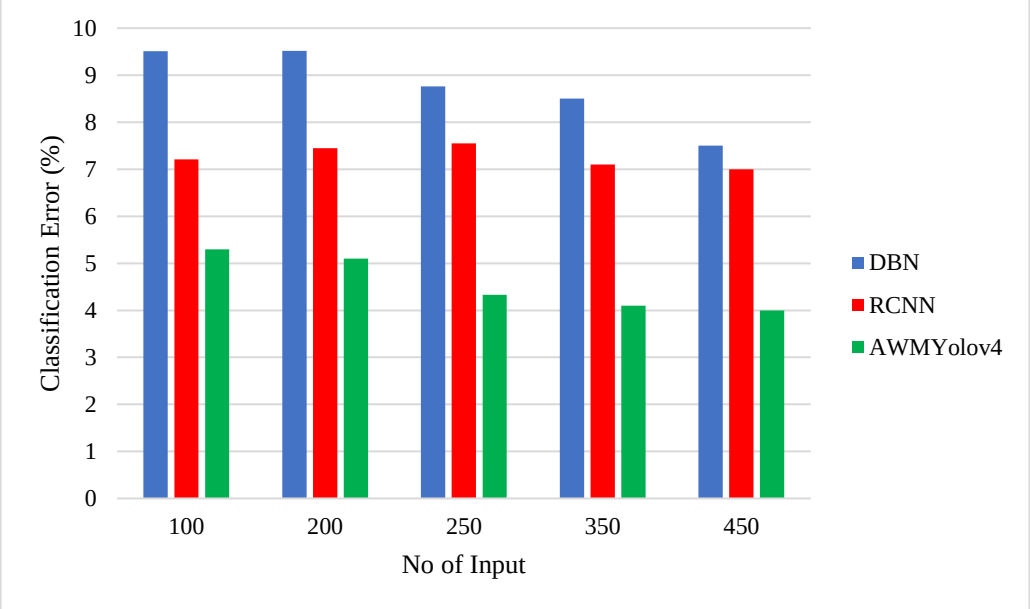

Figure 4.10 Bar chart for classification error compared with the existing methods

The number of ultrasound images is considered in the horizontal direction, and the Classification Error is attained on the vertical axis. The AWMYolov4+ Model's prediction time is indicated by the greenish color column, while the DBN and RCNN

model's Classification errors are indicated by the blue and red coloring, respectively. When compared to conventional methods, the proposed AWMYolov4+ Model has a greater Classification Error prediction.

4.3.3 Impact of false positive rate

The false positive rate is computed as the ratio of the number of ultrasound images that are incorrectly predicted as diseased ultrasound images or normal images through performing segmentation and classification processes from given input ultrasound images. The false positive rate is calculated as,

$$\mathrm{FP_{Rate}} = \left(\frac{\text{The number of images correctly predicted}}{\text{Total no of Inputs}}\right) * 100 \tag{4.25}$$

From equation 4.25 'FP_{Rate}symbolizes the false positive rate. The false positive rate is measured in terms of percentage (%).

Table 4.6 Comparison of false positive rate in percentage

Dataset	False positive rate (%)		
	DBN	RCNN	AWMYolov4+
100	71	65	33
200	35	33	21
250	30	27	15
350	24	21	12
450	20	16	10

Table 4.6 represents the achieved results of the false positive rate using the AWMYolov4+ Model are compared to two existing techniques namely the DBN and RCNN model. The proposed AWMYolov4+ model efficiently attains a lesser false positive rate than existing methods [212]. The number of ultrasound images is 250 for experimental consideration. By using the AWMYolov4+ model, the observed false positive rate is 30% while the false positive rate of the conventional method's first and second models are 27% and 15% respectively. The different false positive rate results are attained for every method. The graphical illustration of the false positive rate is illustrated in figure 4.11.

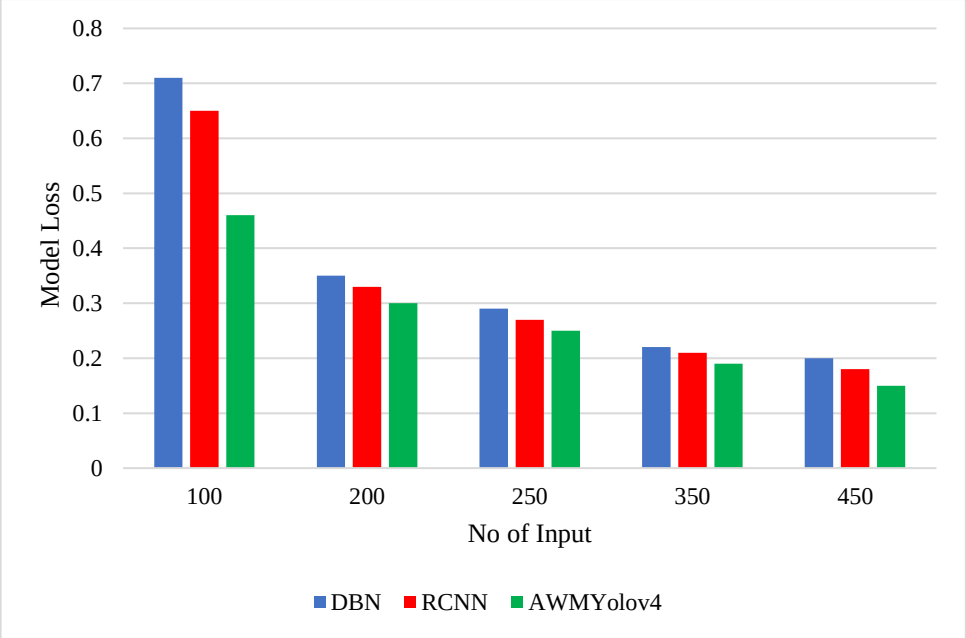

Figure 4.11 Comparison of model loss

4.3.4 Precision, Sensitivity, and F1-Score

After taking into account the classifier's procedure accuracy and misclassification costs, a final prediction result is generated. For convenience, the cost of accurate classification is set to zero so that it can compare the costs of incorrect categorization for different classifiers [213].

Table 4.7 Comparison of the proposed model with existing Precision, Sensitivity, and F1-score value

Models	**Parameters**		
	Precision	**Sensitivity**	**F1 Score**
DBN	90.03	91.67	90.02
RCNN	91.70	92.54	91.04
AWMYolov4+	**95.60**	**94.67**	**95.15**

Each classifier has a section of its mean classification. When a healthy person is mistakenly labeled as having heart disease, an expensive and pointless course of therapy is initiated.

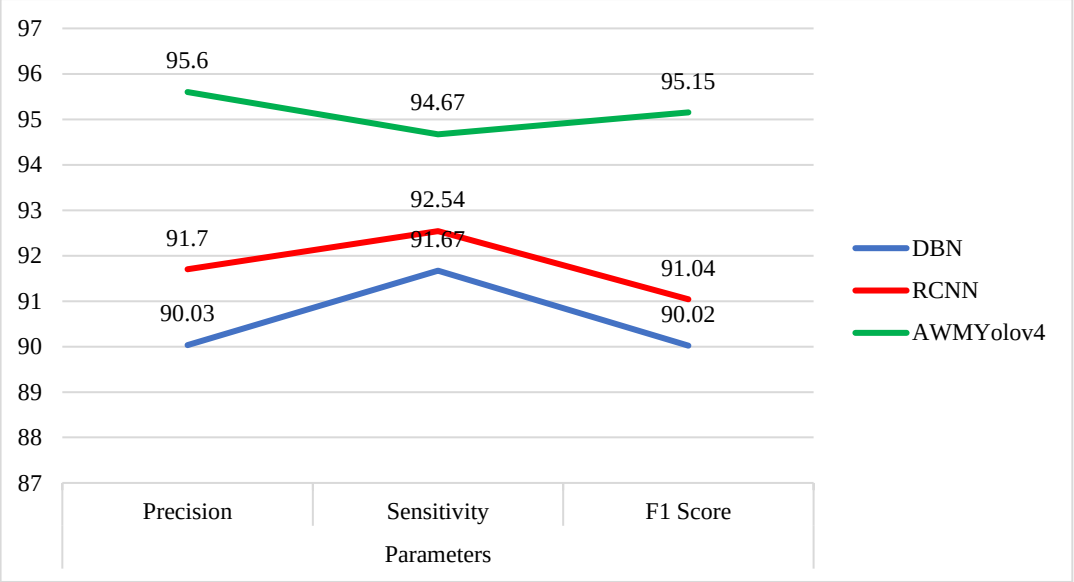

Figure 4.12 Graph representation for precision, sensitivity, F1 Score

In the second scenario, people with heart disease may be assured they are healthy and thus miss out on the best opportunity for treatment, worsening or possibly causing their patient to die.

Summary

In this chapter, the second method of the deep learning model to predict heart diseases from a dataset of cardiovascular diseases based on the provided parameters for the patient is discussed. It is our goal to classify the ultrasonic image to identify the exact defect area within the box and compare the best model with others. The above table shows that different machine-learning models should be preferred based on this model. To evaluate a model, a stranded process of selecting features, tuning parameters, and calibrating the model based on the performance of the data is used. Different algorithms are compared in the final stage, with the proposed optimized model by yolov4+ having the highest accuracy of 95.6%.

Chapter 5

Result Analysis and Discussion

Accuracy, sensitivity, and specificity are the principal component benchmarks used to evaluate the performance of deep learning models used in medical diagnosis studies. As many as nine features were used to create KSDSC classifiers for various image sizes. Accuracy, sensitivity, and specificity are the three primary statistical metrics used to evaluate the efficacy of deep learning techniques in medical decision support analyses [214]. The effectiveness of the algorithm was assessed in my research using AWMYOLOv4+ a statistical tool. True positives, false negatives, false positives, and false negatives are used in all calculations. The classifiers accuracy in separating and identifying the tissue types is quantified by its accuracy. A classifiers sensitivity is measured by how well it can identify the samples of a known tissue type while it is specificity is measured by how well it can confidently rule out other tissue types or the sample in question.

A highly sensitive test is required for a positive diagnosis, while a highly specific test aids in making a negative diagnosis with greater certainty. To calculate precision, sensitivity, and specificity based on the equation (4.16,4.17,4.18,4.19) the value is compared with two proposed models. In this research, the three primary darknet53 building blocks the dense connection block, the residual inception block, and the SPP pooling block on this benchmark dataset are assessed. Yolov4+ is used as the reference architecture and added one of the proposed blocks independently to the FCN to test its efficacy. Finally, they evaluate how well they perform in comparison to different deep learning models.

Convolution matrix for the proposed model KSDSC and AWMYoloV4+. Confusion matrices are tables used to measure a classification algorithm's efficacy. The results of a classification algorithm may be seen and reported using a confusion matrix. Table 5.1 shows a confusion matrix with healthy tissue as benign and diseased tissue as malignant.

In a research experiment, AWMYolov4+ was chosen. It is an optimization based on the dense darknet53 network. First, both the initial and second moments of the yolov4+ are used to estimate the average weighted mean and individual adaptive learning

rate of the different parameters. This method has advantages over other optimization methods [25]. In this experiment, the number of iterations is 100, and the learning rate is 0.001. All of the statistics we talked about in this research are the average results of 10 experiments. This is to make sure that the results are accurate. It tells us that our average accuracy is 95.06%, that our recall is 94.67%, that our specificity is 95.60%, and that our classification error is 4.33%. Table 5.1 shows the experiment's results in detail, and Figure 5.1 shows the confusion matrix of the predicted results.

Table 5.1 Confusion matrix based on the patient's records and actual class

	Number of Records	**True Outcome Patients having CVD disease**	
		P (Patients having CVD)	**N (Patients who do not have CVD)**
Actual Class	**P (Patients having CVD)**	TP (Patients who are correctly diagnosed with CVD	FP (Patients who are having CVD but are wrongly diagnosed as normal)
	N (Patients do not have CVD)	FN (Patients who are normal, but diagnosed with CVD)	TN (Patients who are normal and diagnosed as normal)

Table 5.1 shows the confusion matrix, which helps you understand the state of the best possible prediction result and compare it to the real world. This evaluation process is used to categorize the patient's possible heart disease cases so that accurate predictions can be made.

The confusion matrix between the KSDSC Model and the AWMYolov4+ Model is displayed in Figure 5.1. We find that the KSDSC model properly labels 408 images as having CVD-positive effects and 81 images as having CVD-negative effects. For this reason, our initial model accurately classified 489 images while incorrectly classifying 11. 3 of these 11 misclassified images are predicted as CVD-affected even though they are not. Lastly, 8 images had a Type II error, where the images that were actually affected by CVD were predicted to be unaffected by CVD. In our analysis shows that the AWMYolov4+ second model accurately classifies 410 photos as having CVD-positive effects and 87 images as having CVD-negative effects. Thus, 497 photos were correctly identified by our first model, whereas 3 were misclassified. Even though only 1 of these 3 photos really shows CVD-related damage, they were incorrectly labelled as such.

Finally, a Type II mistake was present in 2 photos, meaning that the images that were genuinely afflicted by CVD were misclassified as being unaffected.

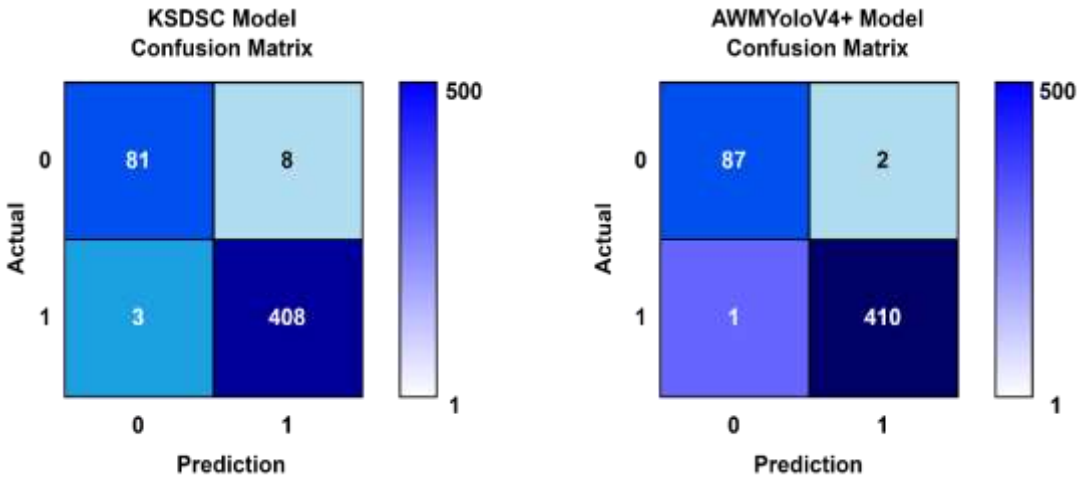

Figure 5.1 Proposed model KSDSC and AWMYolov4+ for heart diseases on CAMUS dataset

This report's classification model achieved an overall average accuracy of 95.06%, indicating that ultrasound doppler data may be used to distinguish between different types of tissue. The following tables detail the accuracy, sensitivity, and specificity of each model that is used to classify the previously discussed 500 ultrasound patient images of varying cardiac tissue types which are referred from Chapter 3 and 4.

Table 5.2 Comparison of the proposed model's

Models	Model Loss (%)	Prediction Time(ms)	Classification Error (%)	Specificity (%)	Sensitivity (%)	Accuracy (%)
KSDSC	13	**270.080**	2.5	92.04	92.34	93.05
AWMYolov4	**10**	298.350	**1.5**	**95.60**	**94.67**	**95.06**

Table 5.2 shows the performance report of the proposed model's output to the networks and the evaluation parameter is shown below in the table. Results that are the most similar to the best are highlighted in bold.

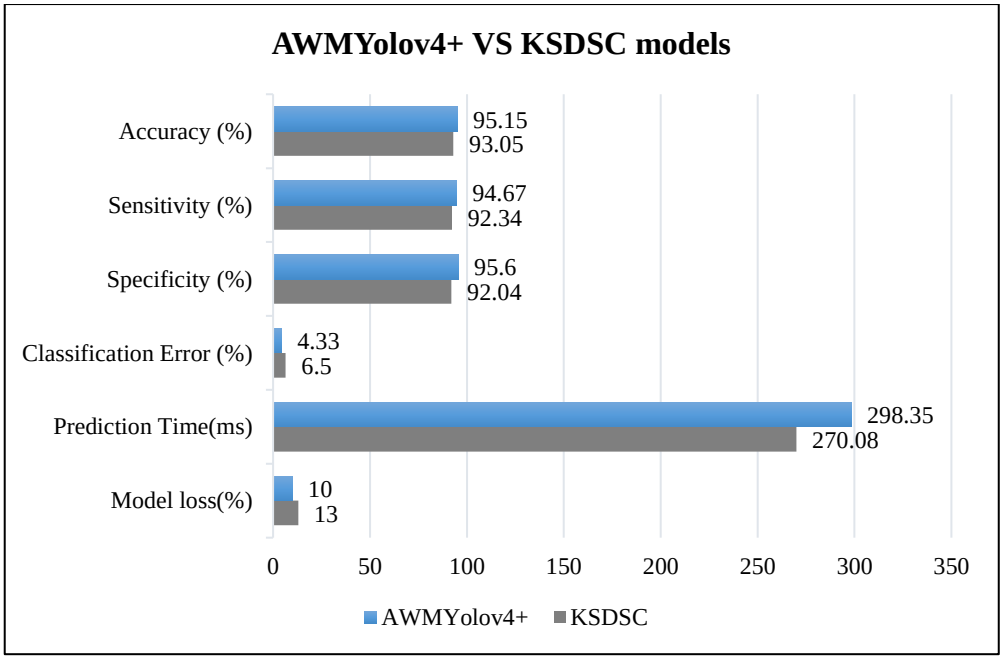

Figure 5.2 The proposed model compared with the proposed network

Figure 5.2 depicts the performance evaluation of the proposed AWMYolov4+ Model's and KSDSC Model prediction time, accuracy, sensitivity specificity, and model loss. The AWMYOLOv4+ is indicated by the blue color and the grey color is indicated by the KSDSC models prediction parameter Proposed model first and second. When compared to conventional methods, the proposed AWMYolov4+ Model is having greater prediction accuracy. This is because the AWMYolov4+ Model uses a deep convolutional neural classifier. By removing the noisy pixels, the Adaptive Weighted Mean pre-processing enhances the image quality. The pre-processed image was separated into several segments by the YoloV4+ backbone module, which uses image segmentation. Following that, the features are retrieved to make a quick and accurate classification of heart disease prediction.

Chapter 6

Conclusion and Future work

The obstacles faced in the analysis of medical images have been addressed in this thesis. The deep learning (DL)-based KSDSC model, AWMyolov4+, deep neural networks, and multi-instance learning suggested in this thesis are some of the examples that make use of machine learning concepts.

In the field of medical imaging, manually labeled images are often in short supply. If models are only developed using the small set of annotated images, then the unlabelled images will not be useful. Machine learning models' effectiveness can be enhanced by using images without annotations. To boost the effectiveness of the model, a novel approach that makes use of manually annotated images is proposed in this research. The proposed method makes use of unsupervised learning by training an additional deep neural network to understand the meaning of images with the help of artificial labels. Then, the trained weights were applied to the main neural network to kick off the training process, which took place using the annotated images as inputs. Existing applicable self-supervised methods, such as patch relative positions and local context prediction, are likely to learn trivial features because medical images are typically static, grayscale images.

Here, the images are randomly changed, and applied deep learning concepts to retain the original image. Because of this, the DL has no choice but to pick up on the meanings conveyed by the images. In Chapter 5, the performance of the proposed method on classification, localization, and segmentation issues in the field of medical image processing was analyzed. These applications used medical imaging data, such as cardiac ultrasounds, which were classified in various ways. The outcomes proved that 1) the proposed context restoration strategy enhances KSDSC performance in every scenario, and 2) self-supervised learning can yield results on par with those of fully supervised learning in some circumstances.

One original method for using images without annotations was provided by self-supervised learning, which restores image context. Based on our findings, the self-supervised training-based kushner stratonovich filter is used to solve a variety of problems and boosts the performance of ultrasound images.

The quantity of labeled training data can be reduced to improve performance. In addition, when compared to other self-supervised learning methods, context restoration performed exceptionally well when it comes to understanding the semantics of images. However, these conclusions are based entirely on evidence from the real world. No one has yet determined whether or not there is a best self-supervised learning strategy overall, or whether or not there is a best self-supervised learning strategy for specific tasks.

Furthermore, an entity for the segmentation of low-quality medical images using multi-scale patches is considered in this research. The use of this framework for AWMF on medical images is discussed in Chapter 4. Image patches at varying scales provide invaluable contextual information. Having this knowledge allows for precise segmentation of lesions, even in low-quality images. As the segmentation was guided by an already-existing AWMF graph, the efficiency of the framework was increased. With data from extensive clinical trials, an in-depth analysis of the framework was carried out. The similarity in segmentation predicted lesion volume, and disease ratings with two scoring systems were used to evaluate the proposed framework's performance more than that of human experts. It is found that the proposed method performed well and it is more reliable and meaningful than human experts.

Edge detection is a common difficulty in medical image analysis used in a wide variety of clinical applications. It would be helpful to have a segmentation technique that could be applied to a variety of contexts. The (AWMYOLOv4+) Adaptive weighted mean Filtered Image Denoising (AWMFID(x)) method uses the variable probabilities of a non-linear dynamic system to remove noise from an input image. Also, the AWMYOLOv4+, based on the DarkNet53, uses a similarity index to calculate the correlation between pixels.

The DarkNet's three essential building blocks the dense connection block, the head and neck inception block, and the SPP pooling block expand and deepen the network beyond that of a standard FCN activated with mish while allowing for fine-grained regulation over the network's basis function.

The course of the most recent but not least, multiple experts' annotations were incorporated into the analysis presented in Chapter 4. There are disagreements among specialists in the annotations, both in terms of pixel-level delineations and overall subject ratings. However, all of the chapters' model training relies on the specifications of a single expert. The likelihood of bias in the evaluation increases when these

distinctions are used as ground truth labels during training. Therefore, the models perform poorly in some circumstances.

The standard feature selection, parameter tuning, and model calibration to assess the model's efficacy in light of the data are used. In the final phase, various algorithms have been evaluated and contrasted and found that the proposed optimized model by AWMYolov4+ achieves the highest accuracy rate (94.67 percent).

Future work

Medical image analysis holds a lot of promise for the future, as many issues, including image segmentation and classification can be effectively tackled with DNNs. Large quantities of notated images produce the most promising results. Nevertheless, it is challenging to construct datasets as large as the CAMUS [33] in the field of medical imaging. As a result, one must exercise caution with both under and over-fitting. A theoretical analysis is required to determine the performance upper bound and the means to achieve the bound to solve underfitting issues.

The key to solving overfitting issues in training models is to generalize from images with sparse, sparsely annotated, or no annotations at all. A strong theoretical analysis is needed, but it is currently unavailable.

Most humans can identify objects of interest (e.g., vehicles, animals) in an image dataset comprised entirely of images captured in their natural environments. Despite this, when presented with a medical image, various specialists may arrive at different conclusions about what should be annotated. In a perfect world, multiple experts would annotate each image, and the resulting consensus would be used to inform model creation. However, the time and effort required are prohibitively high. In reality, a dataset is annotated by a select team of specialists. Each expert annotates a small sample of images, with a maximum of four experts contributing to each image's annotation. It is important to evaluate the accuracy of the various experts' annotations before moving forward with model development.

To better quantify patients, diagnose conditions, and plan treatments, medical image analysis has proven to be an invaluable resource. The result of a classification task is a predicted label for a specific category. When performing a detection task, the result is typically a highlighted bounding box. The result of most segmentation operations is a soft probability map that can be further processed into binary maps. The amount of segmentation work will then be known. Both geometric reconstruction and visualization

can benefit from the binary segmentation map. The relationship between the variables of interest and health outcomes can also be investigated, by providing additional clinical value.

For instance, the significant contributions to the analysis of stroke patient medical images, and our CVD segmentation work has yielded promising results. After that, it is clinically relevant to look into whether or not the ICH measurements agreed with these segmentations. The development of a lesion can also be tracked with subsequent imaging studies. Clinicians can do a better job of diagnosing and treating lesions if they have a better grasp on how they develop over time.

References

[1] Roth, G.A., Mensah, G.A., Johnson, C.O., Addolorato, G., Ammirati, E., Baddour, L.M., Barengo, N.C., Beaton, A.Z., Benjamin, E.J., Benziger, C.P. and Bonny, A., 2020. Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the GBD 2019 study. Journal of the American College of Cardiology, 76(25), pp.2982-3021.

[2] Han, X., 2022. Survival Analysis and Robustness for Chronic Diseases (Doctoral dissertation, New York University).

[3] Lakhanpal, A., Sagar, K.M., Murugan, K. and Singh, A.S., 2021, March. Discover Pretend Disease News Misleading Data in Social Media Networks Using Machine Learning Techniques. In 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE) (pp. 784-788). IEEE.

[4] Salte, I.M., Østvik, A., Smistad, E., Melichova, D., Nguyen, T.M., Karlsen, S., Brunvand, H., Haugaa, K.H., Edvardsen, T., Lovstakken, L. and Grenne, B., 2021. Artificial intelligence for automatic measurement of left ventricular strain in echocardiography. Cardiovascular Imaging, 14(10), pp.1918-1928.

[5] Rahman, S.U. and Hassan, M., 2013. HEART S ROLE IN THE HUMAN BODY: A. HEART, 2(2), pp.1-6.

[6] Boon, N., Norell, M., Hall, J., Jennings, K., Penny, L., Wilson, C., Chambers, J. and Weston, R., 2006. National variations in the provision of cardiac services in the United Kingdom: second report of the British Cardiac Society Working Group, 2005. Heart, 92(7), pp.873-878.

[7] Benso, A., Di Carlo, S. and Politano, G., 2010. A cDNA microarray gene expression data classifier for clinical diagnostics based on graph theory. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 8(3), pp.577-591.

[8] Shahjehan, R.D. and Bhutta, B.S., 2022. Coronary artery disease. In StatPearls [Internet]. StatPearls Publishing.

[9] Hosseini, K., Mortazavi, S.H., Sadeghian, S., Ayati, A., Nalini, M., Aminorroaya, A., Tavolinejad, H., Salarifar, M., Pourhosseini, H., Aein, A. and Jalali, A., 2021. Prevalence and trends of coronary artery disease risk factors and their effect on age of diagnosis in patients with established coronary artery disease: Tehran Heart Center (2005–2015). BMC cardiovascular disorders, 21, pp.1-11.

[10] BE, L.G.K.T.B., 2017. Setio AAA Ciompi F Ghafoorian M van der Laak JA van Ginneken B Sánchez CI A survey on deep learning in medical image analysis. Med Image Anal, 42(1995), p.60.

[11] Narula, S., Shameer, K., Salem Omar, A.M., Dudley, J.T. and Sengupta, P.P., 2016. Machine-learning algorithms to automate morphological and functional assessments in 2D echocardiography. Journal of the American College of Cardiology, 68(21), pp.2287-2295.

[12] Wang, C. and Smedby, Ö., 2014. Model-based left ventricle segmentation in 3D ultrasound using phase image. In MICCAI Workshop" Grand Challenge on Endocardial Three-dimensional Ultrasound Segmentation", Boston, US, 2014..

[13] Liu, S., Wang, Y., Yang, X., Lei, B., Liu, L., Li, S.X., Ni, D. and Wang, T., Deep learning in medical ultrasound analysis: a review. Engineering. 2019; 5 (2): 261–75.

[14] Balaji, K. and Lavanya, K., 2019. Medical image analysis with deep neural networks. In Deep learning and parallel computing environment for bioengineering systems (pp. 75-97). Academic Press.

[15] Gonçalves, L.F., Espinoza, J., Romero, R., Kusanovic, J.P., Swope, B., Nien, J.K., Erez, O., Soto, E. and Treadwell, M.C., 2006. Four-dimensional ultrasonography of the fetal heart using a novel Tomographic Ultrasound Imaging display. Journal of perinatal medicine, 34(1), pp.39-55.

[16] Martin-Isla, C., Campello, V.M., Izquierdo, C., Raisi-Estabragh, Z., Baeßler, B., Petersen, S.E. and Lekadir, K., 2020. Image-based cardiac diagnosis with machine learning: a review. Frontiers in cardiovascular medicine, p.1.

[17] Dorazil, J., Říha, K. and Dutta, M.K., 2019, July. Common carotid artery wall localization in B-mode ultrasound images for initialization of artery wall tracking methods. In 2019 42nd International Conference on Telecommunications and Signal Processing (TSP) (pp. 605-608). IEEE.

[18] Litjens, G., Ciompi, F., Wolterink, J.M., de Vos, B.D., Leiner, T., Teuwen, J. and Išgum, I., 2019. State-of-the-art deep learning in cardiovascular image analysis. JACC: Cardiovascular imaging, 12(8 Part 1), pp.1549-1565.

[19] Dommaraju, V.S., Nathani, K., Tariq, U., Al-Turjman, F., Kallam, S. and Patan, R., 2020. ECMCRR-MPDNL for cellular network traffic prediction with big data. IEEE Access, 8, pp.113419-113428.

[20] Huang, X., Zhu, H. and Wang, J., 2021. Adoption of snake variable model-based method in segmentation and quantitative calculation of cardiac ultrasound medical images. Journal of Healthcare Engineering, 2021.

[21] Molloy, A., Beaumont, K., Alyami, A., Kirimi, M., Hoare, D., Mirzai, N., Heidari, H., Mitra, S., Neale, S.L. and Mercer, J.R., 2021. Challenges to the development of the next generation of self-reporting cardiovascular implantable medical devices. IEEE Reviews in Biomedical Engineering, 15, pp.260-272.

[22] Andreopoulos, A. and Tsotsos, J.K., 2008. Efficient and generalizable statistical models of shape and appearance for analysis of cardiac MRI. Medical image analysis, 12(3), pp.335-357.

[23] Khalil, A., Ng, S.C., Liew, Y.M. and Lai, K.W., 2018. An overview on image registration techniques for cardiac diagnosis and treatment. Cardiology research and practice, 2018.

[24] Pencina, M.J., Navar, A.M., Wojdyla, D., Sanchez, R.J., Khan, I., Elassal, J., D'Agostino Sr, R.B., Peterson, E.D. and Sniderman, A.D., 2019. Quantifying importance of major risk factors for coronary heart disease. Circulation, 139(13), pp.1603-1611.

[25] Damodharan, D., Goel, A.K., 2022 , Cardio Vascular Diseases Detection Using Ultrasonic Image by Retaining Deep Learning Model. IJEER 10(3), pp.639-643.

[26] Gahungu, N., Trueick, R., Bhat, S., Sengupta, P.P. and Dwivedi, G., 2020. Current challenges and recent updates in artificial intelligence and echocardiography. Current Cardiovascular Imaging Reports, 13, pp.1-12.

[27] Kusunose, K., Abe, T., Haga, A., Fukuda, D., Yamada, H., Harada, M. and Sata, M., 2020. A deep learning approach for assessment of regional wall motion abnormality from echocardiographic images. Cardiovascular Imaging, 13(2_Part_1), pp.374-381.

[28] Louridi, N., Douzi, S. and El Ouahidi, B., 2021. Machine learning-based identification of patients with a cardiovascular defect. Journal of Big Data, 8, pp.1-15.

[29] Kawasaki, M., Hattori, A., Ishihara, Y., Okubo, M., Nishigaki, K., Takemura, G., Saio, M., Takami, T. and Minatoguchi, S., 2010. Tissue Characterization of Coronary Plaques and Assessment of Thickness of Fibrous Cap Using Integrated Backscatter Intravascular Ultrasound–Comparison With Histology and Optical Coherence Tomography–. Circulation Journal, 74(12), pp.2641-2648.

[30] Sonoura, T., Kodera, S., Shakya, S. and Kanda, J., 2019. Efficacy of cilostazol for sick sinus syndrome to avoid permanent pacemaker implantation: A retrospective case–control study. Journal of Cardiology, 74(4), pp.328-332.

[31] Acharya, U.R., Faust, O., Molinari, F., Sree, S.V., Junnarkar, S.P. and Sudarshan, V., 2015. Ultrasound-based tissue characterization and classification of fatty liver disease: A screening and diagnostic paradigm. Knowledge-Based Systems, 75, pp.66-77.

[32] Klingensmith, J.D., Haggard, A., Fedewa, R.J., Qiang, B., Cummings III, K., DeGrande, S., Vince, D.G. and Elsharkawy, H., 2018. Spectral analysis of ultrasound radiofrequency backscatter for the detection of intercostal blood vessels. Ultrasound in medicine & biology, 44(7), pp.1411-1422.

[33] Lang, R.M., Badano, L.P., Tsang, W., Adams, D.H., Agricola, E., Buck, T., Faletra, F.F., Franke, A., Hung, J., de Isla, L.P. and Kamp, O., 2012. EAE/ASE recommendations for image acquisition and display using three-dimensional echocardiography. European Heart Journal–Cardiovascular Imaging, 13(1), pp.1-46.

[34] Yodwut, C., Weinert, L., Klas, B., Lang, R.M. and Mor-Avi, V., 2012. Effects of frame rate on three-dimensional speckle-tracking–based measurements of myocardial deformation. Journal of the American Society of Echocardiography, 25(9), pp.978-985.

[35] Leclerc, S., Smistad, E., Pedrosa, J., Østvik, A., Cervenansky, F., Espinosa, F., Espeland, T., Berg, E.A.R., Jodoin, P.M., Grenier, T. and Lartizien, C., 2019. Deep learning for segmentation using an open large-scale dataset in 2D echocardiography. IEEE transactions on medical imaging, 38(9), pp.2198-2210.

[36] Krishnan Rajan, N., Song, Z., Hoffmann, K.R., Belohlavek, M., McMahon, E.M. and Borazjani, I., 2016. Automated three-dimensional reconstruction of the left

ventricle from multiple-axis echocardiography. Journal of biomechanical engineering, 138(1), p.011003.

[37] Zhou, J., Du, M., Chang, S. and Chen, Z., 2021. Artificial intelligence in echocardiography: detection, functional evaluation, and disease diagnosis. Cardiovascular ultrasound, 19(1), pp.1-11.

[38] Wei, L.I.U., Jinming, C.H.E.N., Bo, L.I.U., Wei, H.U., Xingjin, W.U. and Hui, Z.H.O.U., 2022. Tongue image segmentation and tongue color classification based on deep learning. Digital Chinese Medicine, 5(3), pp.253-263.

[39] Jindal, H., Agrawal, S., Khera, R., Jain, R. and Nagrath, P., 2021. Heart disease prediction using machine learning algorithms. In IOP conference series: materials science and engineering (Vol. 1022, No. 1, p. 012072). IOP Publishing.

[40] Madani, A., Arnaout, R., Mofrad, M. and Arnaout, R., Fast and accurate view classification of echocardiograms using deep learning. npj Digit. Med. 1, 6 (2018).

[41] Mohanasundaram, R., Malhotra, A.S., Arun, R. and Periasamy, P.S., 2019. Deep learning and semi-supervised and transfer learning algorithms for medical imaging. In Deep learning and parallel computing environment for bioengineering systems (pp. 139-151). Academic Press.

[42] Liu, N., Shen, J., Xu, M., Gan, D., Qi, E.S. and Gao, B., 2018. Improved cost-sensitive support vector machine classifier for breast cancer diagnosis. Mathematical Problems in Engineering, 2018, pp.1-13.

[43] Louridi, N., Douzi, S. and El Ouahidi, B., 2021. Machine learning-based identification of patients with a cardiovascular defect. Journal of Big Data, 8, pp.1-15.

[44] Damodharan, D., Kumar, N.S., Anandhan, K. and Kumar, S., 2019. A Review On Spotting and Segmenting The Tumor Using Mammography. Journal of Innovation in Computer Science and Engineering, 9(1), pp.29-33.

[45] Deepika, P. and Sasikala, S., 2020, November. Enhanced model for prediction and classification of cardiovascular disease using decision tree with particle swarm optimization. In 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 1068-1072). IEEE.

[46] Sengupta, P.P. and Chandrashekhar, Y., 2022. Imaging With Deep Learning: Sharpening the Cutting Edge. Cardiovascular Imaging, 15(3), pp.547-549.

[47] Das, R., Turkoglu, I. and Sengur, A., 2009. Effective diagnosis of heart disease through neural networks ensembles. Expert systems with applications, 36(4), pp.7675-7680.

[48] Ravindhar, N.V., Anand, H.S. and Ragavendran, G.W., 2019. Intelligent diagnosis of cardiac disease prediction using machine learning. International Journal of Innovative Technology and Exploring Engineering, 8(11), pp.1417-1421.

[49] Amelio, L. and Amelio, A., 2019. Classification methods in image analysis with a special focus on medical analytics. Machine Learning Paradigms: Advances in Data Analytics, pp.31-69.

[50] Ghorbani, A., Ouyang, D., Abid, A., He, B., Chen, J.H., Harrington, R.A., Liang, D.H., Ashley, E.A. and Zou, J.Y., 2020. Deep learning interpretation of echocardiograms. NPJ digital medicine, 3(1), p.10.

[51] Hajjam, E.H., Andr, E. and AK, G., 2020. Informatics in Medicine Unlocked Classification models for heart disease prediction using feature selection and PCA.

[52] Damodharan, D. and Goel, A.K., 2022, July. Comprehensive Study—A Deep Learning and Machine Learning Classification Methods for Cardiogram Images. In Artificial Intelligence on Medical Data: Proceedings of International Symposium, ISCMM 2021 (pp. 3-13). Singapore: Springer Nature Singapore.

[53] Andreadis, I.I. and Nikita, K.S., 2019. Tele-, Mobile-and Web-Based Technologies in Cardiovascular Medicine. Cardiovascular Computing—Methodologies and Clinical Applications, pp.261-277.

[54] Slomka, P.J., Miller, R.J., Isgum, I. and Dey, D., 2020, July. Application and translation of artificial intelligence to cardiovascular imaging in nuclear medicine and noncontrast CT. In Seminars in Nuclear Medicine (Vol. 50, No. 4, pp. 357-366). WB Saunders.

[55] Sonawane, R. and Patil, H., 2022. Automated heart disease prediction model by hybrid heuristic-based feature optimization and enhanced clustering. Biomedical Signal Processing and Control, 72, p.103260.

[56] Chanchal, A., Singh, A.S. and Anandhan, K., 2021, September. A Modern Comparison of ML Algorithms for Cardiovascular Disease Prediction. In 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO) (pp. 1-5). IEEE.

[57] Zhuang, Z., Jin, P., Joseph Raj, A.N., Yuan, Y. and Zhuang, S., 2021. Automatic segmentation of left ventricle in echocardiography based on YOLOv3 model to achieve constraint and positioning. Computational and Mathematical Methods in Medicine, 2021, pp.1-11.

[58] Beunza, J.J., Puertas, E., García-Ovejero, E., Villalba, G., Condes, E., Koleva, G., Hurtado, C. and Landecho, M.F., 2019. Comparison of machine learning algorithms for clinical event prediction (risk of coronary heart disease). Journal of biomedical informatics, 97, p.103257.

[59] Shankar, V., Kumar, V., Devagade, U., Karanth, V. and Rohitaksha, K., 2020. Heart disease prediction using CNN algorithm. SN Computer Science, 1(3), p.170.

[60] Li, W., Zuo, M., Zhao, H., Xu, Q. and Chen, D., 2022. Prediction of coronary heart disease based on combined reinforcement multitask progressive time-series networks. Methods, 198, pp.96-106.

[61] Lalmuanawma, S., Hussain, J. and Chhakchhuak, L., 2020. Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review. Chaos, Solitons & Fractals, 139, p.110059.

[62] Aljanabi, M., Qutqut, M.H. and Hijjawi, M., 2018. Machine learning classification techniques for heart disease prediction: a review. International Journal of Engineering & Technology, 7(4), pp.5373-5379.

[63] Henglin, M., Stein, G., Hushcha, P.V., Snoek, J., Wiltschko, A.B. and Cheng, S., 2017. Machine learning approaches in cardiovascular imaging. Circulation: Cardiovascular Imaging, 10(10), p.e005614.

[64] Rahim, A., Rasheed, Y., Azam, F., Anwar, M.W., Rahim, M.A. and Muzaffar, A.W., 2021. An integrated machine learning framework for effective prediction of cardiovascular diseases. IEEE Access, 9, pp.106575-106588.

[65] Patro, S.P., Nayak, G.S. and Padhy, N., 2021. Heart disease prediction by using novel optimization algorithm: a supervised learning prospective. Informatics in Medicine Unlocked, 26, p.100696.

[66] Waris, S.F. and Koteeswaran, S., 2021. Heart disease early prediction using a novel machine learning method called improved K-means neighbor classifier in python. Materials Today: Proceedings.

[67] Samundeeswari, E.S., Saranya, P.K. and Manavalan, R., 2016, March. Segmentation of breast ultrasound image using regularized K-means (ReKM) clustering. In 2016 international conference on wireless communications, signal processing and networking (WiSPNET) (pp. 1379-1383). IEEE.

[68] Son, Y.J., Kim, H.G., Kim, E.H., Choi, S. and Lee, S.K., 2010. Application of support vector machine for prediction of medication adherence in heart failure patients. Healthcare informatics research, 16(4), pp.253-259.

[69] Dey, D., Slomka, P.J., Leeson, P., Comaniciu, D., Shrestha, S., Sengupta, P.P. and Marwick, T.H., 2019. Artificial intelligence in cardiovascular imaging: JACC state-of-the-art review. Journal of the American College of Cardiology, 73(11), pp.1317-1335.

[70] Sengupta, P.P., Huang, Y.M., Bansal, M., Ashrafi, A., Fisher, M., Shameer, K., Gall, W. and Dudley, J.T., 2016. Cognitive machine-learning algorithm for cardiac imaging: a pilot study for differentiating constrictive pericarditis from restrictive cardiomyopathy. Circulation: Cardiovascular Imaging, 9(6), p.e004330.

[71] Lai, Z. and Deng, H., 2018. Medical image classification based on deep features extracted by deep model and statistic feature fusion with multilayer perceptron. Computational intelligence and neuroscience, 2018.

[72] Murphy A., Moore C., 2023, Decision tree (machine learning). Reference article, Radiopaedia.org Accessed on 12 Mar 2023.

[73] Padmavathi, S. and Ramanujam, E., 2015. Naïve Bayes classifier for ECG abnormalities using multivariate maximal time series motif. Procedia Computer Science, 47, pp.222-228.

[74] Xu, Z., Tao, B., Liu, C., Han, D., Zhang, J., Liu, J., Li, S., Li, W., Wang, J., Liang, J. and Cao, F., 2021. Three-dimensional quantitative assessment of myocardial infarction via multimodality fusion imaging: methodology, validation, and preliminary clinical application. Quantitative Imaging in Medicine and Surgery, 11(7), p.3175.

[75] Rani, M., Bakshi, A. and Gupta, A., 2020, March. Prediction of Heart Disease Using Naïve bayes and Image Processing. In 2020 International Conference on Emerging Smart Computing and Informatics (ESCI) (pp. 215-219). IEEE.

[76] Nagavelli, U., Samanta, D. and Chakraborty, P., 2022. Machine learning technology-based heart disease detection models. Journal of Healthcare Engineering, 2022.

[77] Li, R., Zhang, W., Shen, S., Yao, J., Li, B., Zhou, B., Chen, G. and Wang, Z., 2021. An intelligent heartbeat classification system based on attributable features with AdaBoost+ Random forest algorithm. Journal of Healthcare Engineering, 2021, pp.1-19.

[78] Ren, X., Guo, H., Li, S., Wang, S. and Li, J., 2017. A novel image classification method with CNN-XGBoost model. In Digital Forensics and Watermarking: 16th International Workshop, IWDW 2017, Magdeburg, Germany, August 23-25, 2017, Proceedings 16 (pp. 378-390). Springer International Publishing.

[79] Yin, Z.X. and Xu, H.M., 2022. An unsupervised image segmentation algorithm for coronary angiography. BioData Mining, 15(1), p.27.

[80] Li, R., Zhang, W., Shen, S., Yao, J., Li, B., Zhou, B., Chen, G. and Wang, Z., 2021. An intelligent heartbeat classification system based on attributable features with AdaBoost+ Random forest algorithm. Journal of Healthcare Engineering, 2021, pp.1-19.

[81] Wang, C., Yang, G. and Papanastasiou, G., 2022. Unsupervised image registration towards enhancing performance and explainability in cardiac and brain image analysis. Sensors, 22(6), p.2125.

[82] Hu, M., Zhang, J., Matkovic, L., Liu, T. and Yang, X., 2023. Reinforcement learning in medical image analysis: Concepts, applications, challenges, and future directions. Journal of Applied Clinical Medical Physics, p.e13898.

[83] Quer, G., Arnaout, R., Henne, M. and Arnaout, R., 2021. Machine learning and the future of cardiovascular care: JACC state-of-the-art review. Journal of the American College of Cardiology, 77(3), pp.300-313.

[84] Yin, Sixing; Han, Yameng; Pan, Judong; Wang, YIning; Li, Shufang; Yu, F. Richard (2021): Left Ventricle Contouring in Yin, S., Han, Y., Pan, J., Wang, Y. and Li, S., 2022, December. Left Ventricle Contouring in Cardiac Images Based on Deep Reinforcement Learning. In International Conference on Medical Imaging with Deep Learning (pp. 1470-1481). PMLR.

[85] Yin, S., Han, Y., Pan, J., Wang, Y. and Li, S., 2022, December. Left Ventricle Contouring in Cardiac Images Based on Deep Reinforcement Learning. In International Conference on Medical Imaging with Deep Learning (pp. 1470-1481). PMLR.

[86] Ye, Z., Kumar, Y.J., Sing, G.O., Zhang, J. and Ni, X., 2020. Deep echocardiography: a first step toward automatic cardiac disease diagnosis using machine learning. Journal of Internet Technology, 21(6), pp.1589-1600.

[87] Waris, S.F. and Koteeswaran, S., 2021. WITHDRAWN: Heart disease early prediction using a novel machine learning method called improved K-means neighbor classifier in python.

[88] Zhang, D., Chen, Y., Chen, Y., Ye, S., Cai, W., Jiang, J., Xu, Y., Zheng, G. and Chen, M., 2021. Heart disease prediction based on the embedded feature selection

method and deep neural network. Journal of healthcare engineering, 2021, pp.1-9.

[89] Aversano, L., Bernardi, M.L., Cimitile, M., Iammarino, M., Montano, D. and Verdone, C., 2022, May. Using Machine Learning for early prediction of Heart Disease. In 2022 IEEE International Conference on Evolving and Adaptive Intelligent Systems (EAIS) (pp. 1-8). IEEE.

[90] Ali, F., El-Sappagh, S., Islam, S.R., Kwak, D., Ali, A., Imran, M. and Kwak, K.S., 2020. A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. Information Fusion, 63, pp.208-222.

[91] Anwar, S.M., Majid, M., Qayyum, A., Awais, M., Alnowami, M. and Khan, M.K., 2018. Medical image analysis using convolutional neural networks: a review. Journal of medical systems, 42, pp.1-13.

[92] Samir, A.A., Rashwan, A.R., Sallam, K.M., Chakrabortty, R.K., Ryan, M.J. and Abohany, A.A., 2021. Evolutionary algorithm-based convolutional neural network for predicting heart diseases. Computers & Industrial Engineering, 161, p.107651.

[93] Dangare, C. and Apte, S., 2012. A data mining approach for prediction of heart disease using neural networks. International Journal of Computer Engineering and Technology (IJCET), 3(3).

[94] Rai, A., Damodharan, D., Singh, A.S. and Anandhan, K., 2022, July. Enhanced Machine Learning Model for Design and Evaluation of Various Algorithms Applied in Clinical Cardiovascular Disease. In 2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT) (pp. 304-310). IEEE.

[95] Mir, N.F., Basha, A.S. and Tiwari, A.A., 2021, November. Application of Deep Learning in Dynamic Link-Level Virtualization of Cloud Networks Through the Learning Process. In 2021 International Symposium on Computer Science and Intelligent Controls (ISCSIC) (pp. 169-173). IEEE.

[96] Kimura, N., Yoshinaga, I., Sekijima, K., Azechi, I. and Baba, D., 2019. Convolutional neural network coupled with a transfer-learning approach for time-series flood predictions. Water, 12(1), p.96.

[97] Dutta, A., Batabyal, T., Basu, M. and Acton, S.T., 2020. An efficient convolutional neural network for coronary heart disease prediction. Expert Systems with Applications, 159, p.113408.

[98] Gao, H., Lin, S., Yang, Y., Li, C. and Yang, M., 2018. Convolution neural network based on two-dimensional spectrum for hyperspectral image classification. Journal of Sensors, 2018, pp.1-13.

[99] Sarvamangala, D.R. and Kulkarni, R.V., 2022. Convolutional neural networks in medical image understanding: a survey. Evolutionary intelligence, 15(1), pp.1-22.

[100] Ho, N. and Kim, Y.C., 2022. Estimation of Cardiac Short Axis Slice Levels with a Cascaded Deep Convolutional and Recurrent Neural Network Model. Tomography, 8(6), pp.2749-2760.

[101] Zhu, L., Xu, Z. and Fang, T., 2021. Analysis of Cardiac Ultrasound Images of Critically Ill Patients Using Deep Learning. Journal of Healthcare Engineering, 2021.

[102] Dezaki, F.T., Liao, Z., Luong, C., Girgis, H., Dhungel, N., Abdi, A.H., Behnami, D., Gin, K., Rohling, R., Abolmaesumi, P. and Tsang, T., 2018. Cardiac phase detection in echocardiograms with densely gated recurrent neural networks and global extrema loss. IEEE transactions on medical imaging, 38(8), pp.1821-1832.

[103] Choi, E., Schuetz, A., Stewart, W.F. and Sun, J., 2017. Using recurrent neural network models for early detection of heart failure onset. Journal of the American Medical Informatics Association, 24(2), pp.361-370.

[104] Li, X., Jiang, Y., Rodriguez-Andina, J.J., Luo, H., Yin, S. and Kaynak, O., 2021. When medical images meet generative adversarial network: recent development and research opportunities. Discover Artificial Intelligence, 1, pp.1-20.

[105] Priyadarshini, G.I., Sujatha, K., Lakshmi, B.D. and Kamatchi, C., 2021. Detection of Cardiac Stenosis Using Radial Basis Function Network. In Congress on Intelligent Systems: Proceedings of CIS 2020, Volume 1 (pp. 487-492). Springer Singapore.

[106] Xu, M., Qian, P., Zheng, J., Ge, H. and Muzic, R.F., 2020. A novel radial basis neural network-leveraged fast training method for identifying organs in MR images. Computational and Mathematical Methods in Medicine, 2020.

[107] Al Bataineh, A. and Manacek, S., 2022. MLP-PSO hybrid algorithm for heart disease prediction. Journal of Personalized Medicine, 12(8), p.1208.

[108] Isa, A., Iliyas, I.I. and Zarma, M.L., 2022. Computational Intelligence Approaches for Enhancing Biomedical Image Processing Applications Based on Breast Cancer. In Biomedical Signal and Image Processing-Advanced Imaging Technology and Application. IntechOpen.

[109] Jarur, M.C. and Mora, M., 2006. Heart cavity detection in ultrasound images with SOM. In MICAI 2006: Advances in Artificial Intelligence: 5th Mexican International Conference on Artificial Intelligence, Apizaco, Mexico, November 13-17, 2006. Proceedings 5 (pp. 1211-1219).

[110] Smistad, E., Østvik, A., Salte, I.M., Leclerc, S., Bernard, O. and Lovstakken, L., 2018, October. Fully automatic real-time ejection fraction and MAPSE measurements in 2D echocardiography using deep neural networks. In 2018 IEEE International Ultrasonics Symposium (IUS) (pp. 1-4). IEEE.

[111] Cao, Y., Liu, Z., Zhang, P., Zheng, Y., Song, Y. and Cui, L., 2019. Deep learning methods for cardiovascular image. Journal of Artificial Intelligence and Systems, 1(1), pp.96-109.

[112] Rajasekaran, S. and Kousalya, G., 2022. Virtual Nursing Using Deep Belief Networks for Elderly People (DBN-EP). Comput. Syst. Sci. Eng., 42(3), pp.985-1000.

[113] Wu, J., Mazur, T.R., Ruan, S., Lian, C., Daniel, N., Lashmett, H., Ochoa, L., Zoberi, I., Anastasio, M.A., Gach, H.M. and Mutic, S., 2018. A deep Boltzmann machine-driven level set method for heart motion tracking using cine MRI images. Medical image analysis, 47, pp.68-80.

[114] Wahid, F., Azhar, S., Ali, S., Zia, M.S., Almisned, F.A. and Gumaei, A., 2022. Pneumonia Detection in Chest X-Ray Images Using Enhanced Restricted Boltzmann Machine. Journal of Healthcare Engineering, 2022.

[115] Sarra, R.R., Dinar, A.M., Mohammed, M.A. and Abdulkareem, K.H., 2022. Enhanced heart disease prediction based on machine learning and $\chi 2$ statistical optimal feature selection model. Designs, 6(5), p.87.

[116] Javeed, A., Rizvi, S.S., Zhou, S., Riaz, R., Khan, S.U. and Kwon, S.J., 2020. Heart risk failure prediction using a novel feature selection method for feature refinement and neural network for classification. Mobile Information Systems, 2020, pp.1-11.

[117] Farhad, M., Masud, M.M. and Beg, A., 2022, January. Deep Learning Based Cardiac Phase Detection Using Echocardiography Imaging. In Advanced Data Mining and Applications: 17th International Conference, ADMA 2021, Sydney, NSW, Australia, February 2–4, 2022, Proceedings, Part I (pp. 3-17). Cham: Springer International Publishing.

[118] Martin-Isla, C., Campello, V.M., Izquierdo, C., Raisi-Estabragh, Z., Baeßler, B., Petersen, S.E. and Lekadir, K., 2020. Image-based cardiac diagnosis with machine learning: a review. Frontiers in cardiovascular medicine, p.1.

[119] Cherian, R.P., Thomas, N. and Venkitachalam, S., 2020. Weight optimized neural network for heart disease prediction using hybrid lion plus particle swarm algorithm. Journal of Biomedical Informatics, 110, p.103543.

[120] Loizou, C.P., Kyriacou, E., Griffin, M.B., Nicolaides, A.N. and Pattichis, C.S., 2021. Association of intima-media texture with prevalence of clinical cardiovascular disease. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 68(9), pp.3017-3026.

[121] Azarmehr, N., Ye, X., Howard, J.P., Lane, E.S., Labs, R., Shun-Shin, M.J., Cole, G.D., Bidaut, L., Francis, D.P. and Zolgharni, M., 2021. Neural architecture search of echocardiography view classifiers. Journal of Medical Imaging, 8(3), pp.034002-034002.

[122] Ming, C.T., Omar, Z., Mahmood, N.H. and Kadiman, S., 2015. A literature survey of ultrasound and computed tomography-based cardiac image registration. Jurnal Teknologi, 74(6).

[123] Gárate-Escamila, A.K., El Hassani, A.H. and Andrès, E., 2020. Classification models for heart disease prediction using feature selection and PCA. Informatics in Medicine Unlocked, 19, p.100330.

[124] Siddiqi, M.H., Salamah Alhwaiti, Y., Alrashdi, I., Ali, A. and Faisal, M., 2021. Segmentation and classification of heart angiographic images using machine learning techniques. Journal of Healthcare Engineering, 2021.

[125] Litjens, G., Ciompi, F., Wolterink, J.M., de Vos, B.D., Leiner, T., Teuwen, J. and Išgum, I., 2019. State-of-the-art deep learning in cardiovascular image analysis. JACC: Cardiovascular imaging, 12(8 Part 1), pp.1549-1565.

[126] Cao, Y., Liu, Z., Zhang, P., Zheng, Y., Song, Y. and Cui, L., 2019. Deep learning methods for cardiovascular image. Journal of Artificial Intelligence and Systems, 1(1), pp.96-109.

[127] Liu, X., Gao, K., Liu, B., Pan, C., Liang, K., Yan, L., Ma, J., He, F., Zhang, S., Pan, S. and Yu, Y., 2021. Advances in deep learning-based medical image analysis. Health Data Science, 2021.

[128] Khened, M., Kollerathu, V.A. and Krishnamurthi, G., 2019. Fully convolutional multi-scale residual DenseNets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers. Medical image analysis, 51, pp.21-45.

[129] Bernard, O., Lalande, A., Zotti, C., Cervenansky, F., Yang, X., Heng, P.A., Cetin, I., Lekadir, K., Camara, O., Ballester, M.A.G. and Sanroma, G., 2018. Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: is the problem solved?. IEEE transactions on medical imaging, 37(11), pp.2514-2525.

[130] Kim, M., Yan, C., Yang, D., Wang, Q., Ma, J. and Wu, G., 2020. Deep learning in biomedical image analysis. In Biomedical information technology (pp. 239-263). Academic Press.

[131] Zhou, Y., Peng, Z. and Zhou, X.S., 2012, May. Automatic view classification for cardiac MRI. In 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI) (pp. 1771-1774). IEEE.

[132] Ma, P., Li, Q. and Li, J., 2022. Application of artificial intelligence in cardiovascular imaging. Journal of Healthcare Engineering, 2022.

[133] Xing, X., Yu, D. and Zhang, W., 2015. Data calibration based on multisensor using classification analysis: a random forests approach. Mathematical Problems in Engineering, 2015.

[134] Piovoso, M. and Laplante, P.A., 2003. Kalman filter recipes for real-time image processing. Real-Time Imaging, 9(6), pp.433-439.

[135] Müller, D. and Kramer, F., 2021. MIScnn: a framework for medical image segmentation with convolutional neural networks and deep learning. BMC medical imaging, 21(1), pp.1-11.

[136] Liu, X., Wang, X., Su, Q., Zhang, M., Zhu, Y., Wang, Q. and Wang, Q., 2017. A hybrid classification system for heart disease diagnosis based on the RFRS method. Computational and mathematical methods in medicine, 2017.

[137] Anand, M.V., KiranBala, B., Srividhya, S.R., Younus, M. and Rahman, H., 2022. Gaussian Naïve Bayes Algorithm: A Reliable Technique Involved in the Assortment of the Segregation in Cancer. Mobile Information Systems, 2022.

[138] Fahad, L.G., Tahir, S.F., Shahzad, W., Hassan, M., Alquhayz, H. and Hassan, R., 2020. Ant colony optimization-based streaming feature selection: an application to the medical image diagnosis. Scientific Programming, 2020, pp.1-10.

[139] Cai, L., Gao, J. and Zhao, D., 2020. A review of the application of deep learning in medical image classification and segmentation. Annals of translational medicine, 8(11).

[140] Hoeser, T. and Kuenzer, C., 2020. Object detection and image segmentation with deep learning on earth observation data: A review-part i: Evolution and recent trends. Remote Sensing, 12(10), p.1667.

[141] Yin, X.X., Sun, L., Fu, Y., Lu, R. and Zhang, Y., 2022. U-Net-Based medical image segmentation. Journal of Healthcare Engineering, 2022.

[142] Chen, C., 2020. Deep learning for cardiac image segmentation: a review. Front. Cardiovasc. Med. 7, 25 (2020).
[143] Lu, H., She, Y., Tie, J. and Xu, S., 2022. Half-UNet: A simplified U-Net architecture for medical image segmentation. Frontiers in Neuroinformatics, 16.
[144] Lei, T., Wang, R., Wan, Y., Du, X., Meng, H. and Nandi, A.K., 2020. Medical Image Segmentation Using Deep Learning: A Survey.
[145] Wang, A., Zhang, Q., Han, Y., Megason, S., Hormoz, S., Mosaliganti, K.R., Lam, J.C. and Li, V.O., 2022. A novel deep learning-based 3D cell segmentation framework for future image-based disease detection. Scientific Reports, 12(1), pp.1-15.
[146] Pare, S., Kumar, A., Singh, G.K. and Bajaj, V., 2020. Image segmentation using multilevel thresholding: a research review. Iranian Journal of Science and Technology, Transactions of Electrical Engineering, 44, pp.1-29.
[147] Niu, Z. and Li, H., 2019, June. Research and analysis of threshold segmentation algorithms in image processing. In Journal of Physics: Conference Series (Vol. 1237, No. 2, p. 022122). IOP Publishing.
[148] Al-Amri, S.S., Kalyankar, N.V. and Khamitkar, S.D., 2010. Image segmentation by using edge detection. International journal on computer science and engineering, 2(3), pp.804-807.
[149] Muthukrishnan, R. and Radha, M., 2011. Edge detection techniques for image segmentation. International Journal of Computer Science & Information Technology, 3(6), p.259.
[150] Masubuchi, S., Watanabe, E., Seo, Y., Okazaki, S., Sasagawa, T., Watanabe, K., Taniguchi, T. and Machida, T., 2020. Deep-learning-based image segmentation integrated with optical microscopy for automatically searching for two-dimensional materials. npj 2D Materials and Applications, 4(1), p.3.
[151] Lalaoui, L. and Mohamadi, T., 2013. A comparative study of image region-based segmentation algorithms. International Journal of Advanced Computer Science and Applications, 4(6).
[152] Liu, H., 2022. Image Segmentation Techniques for Intelligent Monitoring of Putonghua Examinations. Advances in Mathematical Physics, 2022.
[153] Noor khalid (2022) “Review on Region-Based Segmentation Using Watershed and Region Growing Techniques and their Applications in Different Fields”, Journal La Multiapp, 3(5), pp. 241-249. doi: 10.37899/journallamultiapp.v3i5.714.
[154] Yu, H., Sun, P., He, F. and Hu, Z., 2021. A weighted region-based level set method for image segmentation with intensity inhomogeneity. Plos one, 16(8), p.e0255948.
[155] Mishra, D., Chaudhury, S., Sarkar, M., Manohar, S. and Soin, A.S., 2018, May. Segmentation of vascular regions in ultrasound images: A deep learning approach. In 2018 IEEE International Symposium on Circuits and Systems (ISCAS) (pp. 1-5). IEEE.

[156] Tang, J., 2010, April. A color image segmentation algorithm based on region growing. In 2010 2nd international conference on computer engineering and technology (Vol. 6, pp. V6-634). IEEE.

[157] Tang, J., 2010, April. A color image segmentation algorithm based on region growing. In 2010 2nd international conference on computer engineering and technology (Vol. 6, pp. V6-634). IEEE.

[158] Le, Q.T., Ladret, P., Nguyen, H.T. and Caplier, A., 2020. Image aesthetic assessment based on image classification and region segmentation. Journal of Imaging, 7(1), p.3.

[159] Priyadarshi, G. and Halim, A., 2022. A hybrid Haar wavelet collocation method for nonlocal hyperbolic partial differential equations. arXiv preprint arXiv:2211.07249.

[160] Parida, P. and Bhoi, N., 2017. Wavelet based transition region extraction for image segmentation. Future Computing and Informatics Journal, 2(2), pp.65-78.

[161] Jiao, L. and Zhao, J., 2019. A survey on the new generation of deep learning in image processing. IEEE Access, 7, pp.172231-172263.

[162] Cherian, A.K., Poovammal, E., Philip, N.S., Ramana, K., Singh, S. and Ra, I.H., 2021. Deep learning based filtering algorithm for noise removal in underwater images. Water, 13(19), p.2742.

[163] Ismail, R. and Nagy, S., 2022. Ways of improving of active contour methods in colonoscopy image segmentation. Image Analysis & Stereology, 41(1).

[164] Wang, C., Qin, Y., Jin, H., Kim, I., Vergara, J.D.G., Dong, C., Jiang, Y., Zhou, Q., Li, J., He, Z. and Zou, Z., 2019. A low power cardiovascular healthcare system with cross-layer optimization from sensing patch to cloud platform. IEEE transactions on biomedical circuits and systems, 13(2), pp.314-329.

[165] Kessler, D.A., MacKay, J.W., Crowe, V.A., Henson, F.M., Graves, M.J., Gilbert, F.J. and Kaggie, J.D., 2020. The optimisation of deep neural networks for segmenting multiple knee joint tissues from MRIs. Computerized Medical Imaging and Graphics, 86, p.101793.

[166] Tantary, A.Y. and Shah, F.A., 2020. An intertwining of curvelet and linear canonical transforms. Journal of Mathematics, 2020, pp.1-14.

[167] Ahmad, T., Ma, Y., Yahya, M., Ahmad, B., Nazir, S. and Haq, A.U., 2020. Object detection through modified YOLO neural network. Scientific Programming, 2020, pp.1-10.

[168] Damodharan, D. and Goel, A.K., 2022. Kushner-Stratonovich Dice Segmented Curvelet (KSDSC) deep convolutional neural learning for heart disease prediction.

[169] Nguyen-Van, S., Lieu, Q.X., Xuan-Mung, N. and Nguyen, T.T.N., 2022. A new study on optimization of four-bar mechanisms based on a hybrid-combined differential evolution and Jaya algorithm. Symmetry, 14(2), p.381.

[170] Hauser, K. and Matthes, J., 2017. Medical students' medication communication skills regarding drug prescription—a qualitative analysis of simulated physician-patient consultations. European journal of clinical pharmacology, 73, pp.429-435.

[171] Wadood, A., Farkoush, S.G., Khurshaid, T., Yu, J.T., Kim, C.H. and Rhee, S.B., 2019. Application of the jaya algorithm in solving the problem of the optimal coordination of overcurrent relays in single-and multi-loop distribution systems. Complexity, 2019, pp.1-13.

[172] Bosl, W.J., Loddenkemper, T. and Vieluf, S., 2022. Coarse-graining and the Haar wavelet transform for multiscale analysis. Bioelectronic Medicine, 8(1), pp.1-7.

[173] Babaoğlu, I., Fındık, O. and Bayrak, M., 2010. Effects of principle component analysis on assessment of coronary artery diseases using support vector machine. Expert Systems with Applications, 37(3), pp.2182-2185.

[174] Kelly, S., Voegerl, T., Banzhaf, W. and Gondro, C., 2021. Evolving hierarchical memory-prediction machines in multi-task reinforcement learning. Genetic Programming and Evolvable Machines, 22, pp.573-605.

[175] Islam, M.R., Liu, S., Wang, X. and Xu, G., 2020. Deep learning for misinformation detection on online social networks: a survey and new perspectives. Social Network Analysis and Mining, 10, pp.1-20.

[176] Zhao, D., Ferdian, E., Talou, G.D.M., Quill, G.M., Gilbert, K., Wang, V.Y., Gamage, T.P.B., Pedrosa, J., D'hooge, J., Sutton, T.M. and Lowe, B.S., 2022. MITEA: A dataset for machine learning segmentation of the left ventricle in 3D echocardiography using subject-specific labels from cardiac magnetic resonance imaging. Frontiers in Cardiovascular Medicine, 9.

[177] Ramesh, G., Madhavi, K., Reddy, P.D.K., Somasekar, J. and Tan, J., 2021. WITHDRAWN: Improving the accuracy of heart attack risk prediction based on information gain feature selection technique.

[178] Micucci, M. and Iula, A., 2022. Recent advances in machine learning applied to ultrasound imaging. Electronics, 11(11), p.1800.

[179] Lakshmanan, M., Karnan, H. and Natarajan, S., 2020. Smart diagnosis of cardiac arrhythmias using optimal feature rank score algorithm for solar based energy storage ECG acquisition system. In Smart Healthcare for Disease Diagnosis and Prevention (pp. 125-139). Academic Press.

[180] Yang, S., Chen, M., Pomerleau, D. and Sukthankar, R., 2010, June. Food recognition using statistics of pairwise local features. In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (pp. 2249-2256). IEEE.

[181] Vidhya, S., Singh, D.A.A.G. and Leavline, E.J., 2015. Feature extraction for document classification. Int. J. Innov. Res. Sci. Eng. Technol., 4(6), pp.50-56.

[182] Damodharan, D. and Goel, A.K., 2021, March. A Novel Approach on classification in Machine Learning model-based USG Cardiogram Images. In 2021 international conference on advance computing and innovative technologies in engineering (ICACITE) (pp. 947-950). IEEE.

[183] Singh, A., Sengupta, S. and Lakshminarayanan, V., 2020. Explainable deep learning models in medical image analysis. Journal of Imaging, 6(6), p.52.

[184] D, D. and Goel, A. K. . (2022) "An Implementation of Cardiovascular Disease Prediction in Ultrasonography Images using AWMYOLOv4 Deep Learning

Mode", International Journal on Recent and Innovation Trends in Computing and Communication, 10(9), pp. 40–52. doi: 10.17762/ijritcc.v10i9.5669.

[185] Kamila, N.K. ed., 2015. Handbook of research on emerging perspectives in intelligent pattern recognition, analysis, and image processing.

[186] S. Suganya, 2015, Analysis of Feature Extraction of Optical Character detection in Image Processing Systems, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) NCRTET – 2015 (Volume 3 – Issue 04),

[187] Chowdhary, C.L. and Acharjya, D.P., 2020. Segmentation and feature extraction in medical imaging: a systematic review. Procedia Computer Science, 167, pp.26-36.

[188] Song, T., Yu, X., Yu, S., Ren, Z. and Qu, Y., 2021. Feature extraction processing method of medical image fusion based on neural network algorithm. Complexity, 2021, pp.1-10.

[189] Ali, F., El-Sappagh, S., Islam, S.R., Kwak, D., Ali, A., Imran, M. and Kwak, K.S., 2020. A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. Information Fusion, 63, pp.208-222.

[190] Roy, A.M., Bose, R. and Bhaduri, J., A fast accurate fine-grain object detection model based on YOLOv4 deep neural network. arXiv 2021. arXiv preprint arXiv:2111.00298.

[191] Jeong, H.J., Park, K.S. and Ha, Y.G., 2018, January. Image preprocessing for efficient training of YOLO deep learning networks. In 2018 IEEE International Conference on Big Data and Smart Computing (BigComp) (pp. 635-637). IEEE.

[192] Shao, C., Kaur, P. and Kumar, R., 2021. An improved adaptive weighted mean filtering approach for metallographic image processing. Journal of Intelligent Systems, 30(1), pp.470-478.

[193] Saponara, S., Elhanashi, A. and Zheng, Q., 2022. Developing a real-time social distancing detection system based on YOLOv4-tiny and bird-eye view for COVID-19. Journal of Real-Time Image Processing, 19(3), pp.551-563.

[194] Q. Tomas, J.P., D. Celis, M.N., B. Chan, T.K. and A. Flores, J., 2022, March. Trash detection for computer vision using scaled-yolov4 on water surface. In The 11th International Conference on Informatics, Environment, Energy and Applications (pp. 1-8).

[195] Wang, G., Ding, H., Li, B., Nie, R. and Zhao, Y., 2022. Trident-YOLO: Improving the precision and speed of mobile device object detection. IET Image Processing, 16(1), pp.145-157.

[196] Wang, H., Wang, J., Bai, K. and Sun, Y., 2021. Centered multi-task generative adversarial network for small object detection. Sensors, 21(15), p.5194.

[197] LM, L.Y.W.H.D., 2020. Nguyen TN Han D. Lee A. Jang I. Moon H. A deep learning-based hybrid framework for object detection and recognition in autonomous driving IEEE Access, 8, pp.194228-194239.

[198] Tian, D., Han, Y., Wang, S., Chen, X. and Guan, T., 2022. Absolute size IoU loss for the bounding box regression of the object detection. Neurocomputing, 500, pp.1029-1040.

[199] Cheng, Z. and Zhang, F., 2020. Flower end-to-end detection based on YOLOv4 using a mobile device. Wireless Communications and Mobile Computing, 2020, pp.1-9.
[200] Sun, B., Wang, X., Oad, A., Pervez, A. and Dong, F., 2023. Automatic Ship Object Detection Model Based on YOLOv4 with Transformer Mechanism in Remote Sensing Images. Applied Sciences, 13(4), p.2488.
[201] Feng, W., Zhu, Y., Zheng, J. and Wang, H., 2021. Embedded YOLO: a real-time object detector for small intelligent trajectory cars. Mathematical Problems in Engineering, 2021, pp.1-11.
[202] Kumar, A., Kalia, A., Sharma, A. and Kaushal, M., 2021. A hybrid tiny YOLO v4-SPP module based improved face mask detection vision system. Journal of Ambient Intelligence and Humanized Computing, pp.1-14.
[203] Zhu, X., Bai, Y. and Pei, Y., 2021, October. Improved Pedestrian Detection Algorithm of Yolov4 Network Structure. In Proceedings of the 5th International Conference on Computer Science and Application Engineering (pp. 1-6).
[204] Zhou, X., Yi, J., Xie, G., Jia, Y., Xu, G. and Sun, M., 2022. Human Detection Algorithm Based on Improved YOLO v4. Information Technology and Control, 51(3), pp.485-498.
[205] Yang, Q., Rong, X., Guo, R., Zhao, H. and Zhao, Y., 2022. Real-Time Detection of Insulators and Drop Fuses Based on Improved YOLOv4. Scientific Programming, 2022.
[206] Wang, C.Y., Liao, H.Y.M., Wu, Y.H., Chen, P.Y., Hsieh, J.W. and Yeh, I.H., 2020. CSPNet: A new backbone that can enhance learning capability of CNN. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops (pp. 390-391).
[207] Huang, M.L. and Wu, Y.S., 2023. GCS-YOLOV4-Tiny: A lightweight group convolution network for multi-stage fruit detection. Mathematical Biosciences and Engineering, 20(1), pp.241-268.
[208] Zhuo, L., Tan, S., Li, B. and Huang, J., 2022. Self-adversarial training incorporating forgery attention for image forgery localization. IEEE Transactions on Information Forensics and Security, 17, pp.819-834.
[209] Nepal, U. and Eslamiat, H., 2022. Comparing YOLOv3, YOLOv4 and YOLOv5 for autonomous landing spot detection in faulty UAVs. Sensors, 22(2), p.464.
[210] Jiang, Y., Li, W., Zhang, J., Li, F. and Wu, Z., 2023. YOLOv4-dense: A smaller and faster YOLOv4 for real-time edge-device based object detection in traffic scene. IET Image Processing, 17(2), pp.570-580.
[211] Zhang, Z., Xia, S., Cai, Y., Yang, C. and Zeng, S., 2021. A Soft-YoloV4 for High-Performance Head Detection and Counting. Mathematics, 9(23), p.3096.
[212] Glučina, M., Baressi Šegota, S., Anđelić, N. and Car, Z., 2022. Automated Detection and Classification of Returnable Packaging Based on YOLOV4 Algorithm. Applied Sciences, 12(21), p.11131.
[213] Diwan, T., Anirudh, G. and Tembhurne, J.V., 2022. Object detection using YOLO: Challenges, architectural successors, datasets and applications. Multimedia Tools and Applications, pp.1-33.

[214] Bhatti, U.A., Tang, H., Wu, G., Marjan, S. and Hussain, A., 2023. Deep Learning with Graph Convolutional Networks: An Overview and Latest Applications in Computational Intelligence. International Journal of Intelligent Systems, 2023.

[215] Yugander, P., Tejaswini, C.H., Meenakshi, J., Varma, B.S. and Jagannath, M., 2020. MR image enhancement using adaptive weighted mean filtering and homomorphic filtering. Procedia Computer Science, 167, pp.677-685.

[216] Erkan, U., Thanh, D.N., Enginoğlu, S. and Memiş, S., 2020, June. Improved adaptive weighted mean filter for salt-and-pepper noise removal. In 2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE) (pp. 1-5). IEEE.

[217] Guo, C., Lv, X.L., Zhang, Y. and Zhang, M.L., 2021. Improved YOLOv4-tiny network for real-time electronic component detection. Scientific reports, 11(1), p.22744.

[218] Bochkovskiy, A., Wang, C.Y. and Liao, H.Y.M., 2020. Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934.

[219] Yu, J. and Zhang, W., 2021. Face mask wearing detection algorithm based on improved YOLO-v4. Sensors, 21(9), p.3263.

[220] Liu, X., Hu, J., Wang, H., Zhang, Z., Lu, X., Sheng, C., Song, S. and Nie, J., 2022. Gaussian-IoU loss: Better learning for bounding box regression on PCB component detection. Expert Systems with Applications, 190, p.116178.

[221] Wang, R., Wang, Z., Xu, Z., Wang, C., Li, Q., Zhang, Y. and Li, H., 2021. A real-time object detector for autonomous vehicles based on YOLOv4. Computational Intelligence and Neuroscience, 2021.

[222] Pang, H., Zhang, Y., Cai, W., Li, B. and Song, R., 2022. A real-time object detection model for orchard pests based on improved YOLOv4 algorithm. Scientific Reports, 12(1), p.13557.

[223] Prasetyo, E., Suciati, N. and Fatichah, C., 2021, June. Yolov4-Tiny and Spatial Pyramid Pooling for Detecting Head and Tail of Fish. In 2021 International Conference on Artificial Intelligence and Computer Science Technology (ICAICST) (pp. 157-161). IEEE.

[224] Pan, X., Jia, N., Mu, Y. and Bai, W., Msfe-Panet: Improved Yolov4-Based Small Object Detection Method in Complex Scenes. Available at SSRN 4186484.

[225] Li, Y., Wang, J., Huang, J. and Li, Y., 2022. Research on Deep Learning Automatic Vehicle Recognition Algorithm Based on RES-YOLO Model. Sensors, 22(10), p.3783.

[226] Hua, J., Hao, T., Zeng, L. and Yu, G., 2021. YOLO Mask, an Instance Segmentation Algorithm Based on Complementary Fusion Network. Mathematics, 9(15), p.1766.