TRIBHUVAN UNIVERSITY

INSTITUTE OF ENGINEERING

PULCHOWK CAMPUS

THESIS NO: 075/MSICE/005

DEEP LEARNING IN SPATIOTEMPORAL FETAL CARDIAC IMAGING

BY

DIPAK KUMAR NIDHI

075/MSICE/005

A THESIS

SUBMITTED TO THE DEPARTMENT OF ELECTRONICS AND COMPUTER ENGINEERING IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN INFORMATION AND COMMUNICATION ENGINEERING

DEPARTMENT OF ELECTRONICS AND COMPUTER ENGINEERING

LALITPUR, NEPAL

AUGUST, 2021

DEEP LEARNING IN SPATIOTEMPORAL FETAL CARDIAC IMAGING

BY

DIPAK KUMAR NIDHI

075/MSICE/005

Thesis Supervisor

Prof. Dr. Shashidhar Ram Joshi

A thesis submitted in partial fulfillment of the requirements for the degree of Masters of Science in Information and Communication Engineering

Department of Electronics and Computer Engineering

Institute of Engineering, Pulchowk Campus

Tribhuvan University

Lalitpur, Nepal

AUGUST, 2021

COPYRIGHT©

DECLARATION

I declare that the work hereby submitted for Master of Science in Information and Communication Engineering (MSICE) at IOE, Pulchowk Campus entitled **"DEEP LEARNING IN SPATIOTEMPORAL FETAL CARDIAC IMAGING"** is my own work and has not been previously submitted by me at any university for any academic award.

I authorize IOE, Pulchowk Campus to lend this thesis to other institution or individuals for the purpose of scholarly research.

DIPAK KUMAR NIDHI
075/MSICE/005
Date: AUGUST, 2021

RECOMMENDATION

The undersigned certify that they have read and recommended to the Department of Electronics and Computer Engineering for acceptance, a thesis entitled **"DEEP LEARNING IN SPATIOTEMPORAL FETAL CARDIAC IMAGING"**, submitted by **Dipak Kumar Nidhi** in partial fulfillment of the requirement for the award of the degree of **"Master of Science in Information and Communication Engineering"**.

.....
Supervisor: Prof. Dr. Shashidhar Ram Joshi
Dean of Engineering
Institute of Engineering, Tribhuvan University

.....
External Examiner: Subhash Dhakal
IT Director
Department of National Id and Civil Registration
Government of Nepal

.....
Committee Chairperson: Basanta Joshi, PhD
Program Coordinator, Msc in Information and Communication Engineering
Department of Electronics and Computer Engineering
Institute of Engineering, Tribhuvan University

Date: August, 2021

DEPARTMENTAL ACCEPTANCE

The thesis entitled **"DEEP LEARNING IN SPATIOTEMPORAL FETAL CARDIAC IMAGING"**, submitted by **Dipak Kumar Nidhi** in partial fulfillment of the requirement for the award of the degree of **"Master of Science in Information and Communication Engineering"** has been accepted as a bonafide record of work independently carried out by him in the department.

.....

Prof. Dr. Ram Krishna Maharjan
Head of the Department
Department of Electronics and Computer Engineering,
Pulchowk Campus,
Institute of Engineering,
Tribhuvan University,
Nepal.

ACKNOWLEDGEMENT

It has been a real joy for me to prepare my dissertation, and I would like to take this opportunity to thank everyone who has made it possible.

I would like to express my heartfelt gratitude and heartfelt thanks to Prof. Dr. Shashidhar Ram Joshi, Dean, Institute of Engineering, Tribhuvan University, who has provided guidance and constant encouragement throughout the course of my study and in the preparation of this dissertation with his extensive knowledge and professional expertise. Working under him during my entire post graduation has been an honour and a truly inspirational experience to cherish.

I am very grateful to Prof. Dr. Ram Krishna Maharjan, Head of the Department of Electronics and Computer Engineering at the Pulchowk Campus, for his precious assistance.

I am very much indebted to Dr. Rama Manjari Naik, Fetal Medicine Consultant , Indira IVF and my beloved wife Dr. Khushboo Srivastav, Fetal Medicine and OBYN Consultant, People's General Hospital, Kathmandu. Without their expert input, constant help and critical analysis, the completion of this work would not have been possible.

I am thankful to our program coordinator Dr. Basanta Joshi for providing suitable platform to complete this task. I extend my thank form the bottom of my heart for his constant help, encouragement, and valuable suggestions..

I am also thankful to Prof. Dr. Subarna Shakya, Dr. Surendra Srestha, Dr. Sanjeeb Prasad Panday, and other faculties for their encouragement and precious guidance. Their wisdom and inside has helped me in completing this burden of this research smoothly.

I specifically thank my batch-mates, without whose help my thesis would not have been possible.

I have the utmost gratitude for all the mothers who have given consent to use their fetal scan for the purpose of the study. They are the backbone of my study and without their cooperation, this thesis work would never have been possible.

I am indebted to the Department of Electronics and Computer Engineering for incorporating this work as a part of our syllabus. It makes us the realization of our knowledge in the real-world applications and provides a benchmark to proceed for the thesis and research.

Last but not least, I want to express my sincere thanks to my late parents without whose hard

work and blessings, I wouldn't have been writing this dissertation. I also value the conviction of my father in law Mr. Hariom Srivastav and mother in law Mrs. Kumud Srivastav and thank them with my deepest gratitude for their unconditional love, inspiring wisdom and unwavering encouragement.

Sincerely,
Dipak Kumar Nidhi
(075MSICE005)

DEDICATION

To my late parents ,

The reason of what I become today.

Thank you for your great support and

continuous care.

Your memories and blessings are what keeps

me going.

ABSTRACT

Fetal echocardiography is a standard diagnostic tool used to evaluate and monitor fetuses with a compromised cardiovascular system associated with a number of fetal conditions. Deep learning is a computer technology which can perform specific tasks with specific goals. Deep learning techniques is used to evaluate fetal cardiac ultrasound cine loops and improve the evaluation of fetal abnormalities. In this study, I implemented convolutional neural network and recurrent neural network as CNN+LSTM, CNN+GRU and 3DCNN, deep learning models for the processing and classification of ultrasonographic cine loops into various classes. The CNN+LSTM, CNN+GRU, and 3D CNN algorithms were able to sort the fetal cardiac cine loops into 5 standard views with 92.63%, 94.99%, and 82.69% accuracy, respectively. Furthermore, the CNN+LSTM, CNN+GRU, and 3D CNN were able to accurately diagnose Tricuspid atresia (TA) and Hypoplastic left heart syndrome (HLHS) with 94.61%, 91.99%, and 86.54%, respectively. These deep learning-based algorithms found to be an effective tool for evaluating and monitoring normal and abnormal fetal heart cine loops.

Keywords:
Deep Learning, Fetal Cardiac Cine loops, CHD Lesions, CNN, RNN, 3D CNN

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

CNN	Convolutional Neural Network
RNN	Recurrenrt Neural Network
LSTM	Long Short-term memory
CHD	Congenital Heart Disease
4C	Four Chamber
3VV	Three Vessel View
LVOT	Left Ventricular Outflow Tract
RVOT	Right Ventricular Outflow Tract
ABDO	Abdominal
HLHS	Hypoplastic Left Heart Syndrome
FCN	Fully Convolutional Network
ReLu	Rectified Linear Unit
TP	True Positive
FP	False Positive
TN	True Negative
FN	False Negative
ROC	Receiver Operating Characteristics
AUC	Area Under The Curve
TA	Tricuspid Atresia

CHAPTER 1

INTRODUCTION

1.1 Background

Congenital heart disease(CHD) represents the wide range of anomaly present since birth which affects the normal structure of fetal heart and results in its defective functioning. Any condition which is present since birth is termed as congenital. CHD largely contributes to infant morbidity and mortality that results due to congenital anomaly. It is estimated that there are around 4-13 cases of CHD per thousand live births[1]. According to the World Health Organization's data (WHO), heart anomalies contributed to 42% of infant deaths from 1950 and 1994[2]. It has also been observed that heart defects affecting the fetal cardiac structure are often missed in routine antenatal fetal scanning [3]. Precise antenatal diagnosis offers potential clinical benefit with regards to later prognosis. This is true especially in cases where some form of intervention or surgery can be offered in utero itself (e.g. in utero aortic valvuloplasty for HLHS).These potential benefits all rely on accurate fetal diagnosis of CHD.

A routine fetal anomaly scan is recommended for all pregnant mothers between 18-22 weeks. This includes basic examination of all the major organ system of the fetus to rule out any structural defect. The majority of the information gathered from basic cardiac ultrasonography is done by study of the four chamber(4C) view of the fetal cardia. The next step includes the survey of three vessel(3VV and 3VT) and pulmonary and aortic outflow tracts(RVOT and LVOT vies). This extended basic cardiac evaluation improvises the diagnosis rates for heart abnormalities as shown in Figure **??**. Systematic assessment of the fetal heart can be done using a sweep technique which includes a transverse sweep of the probe from the fetal abdomen towards upper mediastinum. In this sweep, fetal four chamber(4C) is visualised followed by 3V, 3VT and then the outflow tracts(LVOT, RVOT). If any structural defect is detected during basic cardiac sonography or if specific condition (e.g. Maternal diabetes, suspected chromosomal anomalies, teratogenic drug intake) warrants, a further detailed evaluation of fetal heart is done which is called Fetal echocardiography. Fetal echocardiography is described as comprehensive sonographic assessment that is used to recognize and distinguish fetal cardiac abnormalities in the antenatal period. This dedicated diagnostic modality is an augmentation to the "basic" and "extended basic" fetal heart screening framework[4].

Major anomalies considered in this study are

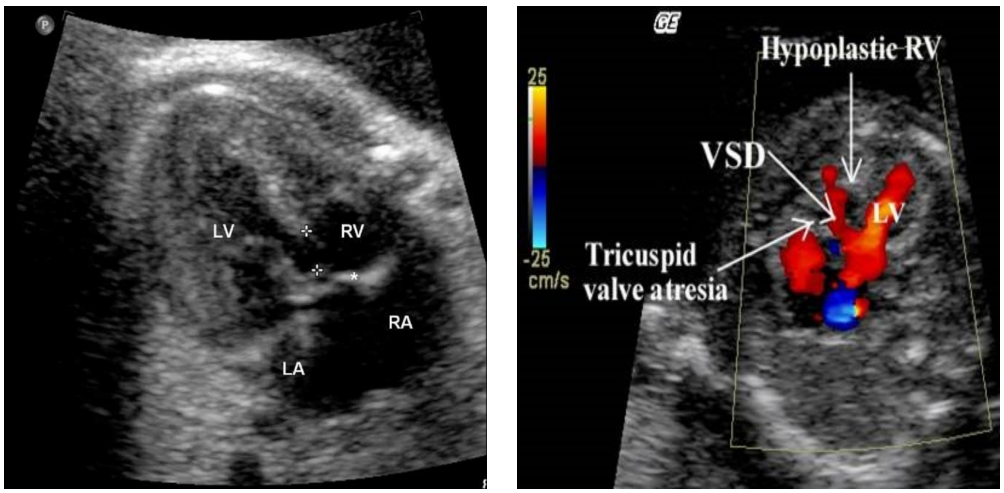
1. Tricuspid atresia (TA)

   Tricuspid atresia is a congenital defect affecting the tricuspid valve. Its incidence is

around 0.7 percent in the fetal life and 2 percent in the neaonatal period.The tricuspid valve, located on the right side of the heart, regulates blood flow from the right atrium to the right ventricle. When the tricuspid valve fails to develop, no blood can flow from the right atrium to the right ventricle and further to the pulmonary arteries. The valve can be entirely atretic, with a fibrous bar replacing the connection, or it can be imperforate. The little rudimentary right ventricle is shifted anteriorly and fills via inlet VSD which is usually associated with this condition. The ventriculo-arterial connection can be concordant or discordant (less frequently) [5].

Tricuspid atresia is a birth defect that is classified as a "critical congenital heart problem" because a newborn with tricuspid atresia will probably require surgery or other interventions soon after birth if it survives in utero.

The four-chamber ultrasound image is used to make the diagnosis: there is a dominant left ventricle, a rudimentary right ventricular chamber, and a minor VSD that allows the right ventricle to fill as shown in Figure 1.1(a). The between the right atrium and ventricle is absent and in its place, hyperechoic fibrous tissue is seen. To validate the absence of right atrioventricular filling and demonstrate flow across the VSD, color Doppler can be employed as shown in Figure 1.1(b). The ventriculoatrial connection can be assessed using four-dimensional echocardiography. Given the radically distinct prognoses of the two lesions, it's important to determine the side of the hypoplastic chamber to avoid mistaking a tricuspid atresia for a hypoplastic left heart or vice versa.



(a) **(b)**

Figure 1.1: Tricuspid Atresia(TA) of fetal heart **(a)** Gray-Scale **(b)** Colour Doppler

One of the most critical poor prognostic variables is the presence of ventriculoarterial discordance and/or a double-outlet connection. A narrow VSD is also a poor prognostic factor, indicating the need for surgery relatively soon.

Early and accurate diagnosis helps the clinicians to advise and perform appropriate surgical intervention in time which increases the 3 years survival of the baby by more than 80 percent.

2. Hypoplastic Left Heart Syndrome(HLHS)

   The term "HLHS" refers to a group of disorders characterized by significant hypoplasia of the left cardiac ventricle and left ventricular outflow tract. HLHS affects about one in every 10,000 live births. There are two types of HLHS. In the first variety, both mitral and aortic valves are atretic. The second type second has aortic atresia and a hypoplastic but patent mitral valve. Due to mitral atresia or imperforation, there is no contact between the left atrium and the left ventricle in the former kind. As a result, the left ventricle functions as a virtual cavity that is sometimes only visible from the outside of the heart in an autopsy sample owing to the route of the anterior and posterior coronary branches. In the latter case, left ventricle is hypoplastic in this case, but it does have an identifiable lumen [5].

   In the four-chamber view, a small globular ventricle with endocardial fibroelastosis or an aslitlike ventricle (mitral and aortic atresia) can be seen on the left side of the heart. The apex of the heart is never reached by the hypoplastic left ventricle Figure 1.2. The atretic or hypoplastic annulus of the mitral valve is noted. A threadlike ascending aorta and the highly hypoplastic annulus of the aortic valve can be seen with difficulty in the long-axis image of the left ventricle.

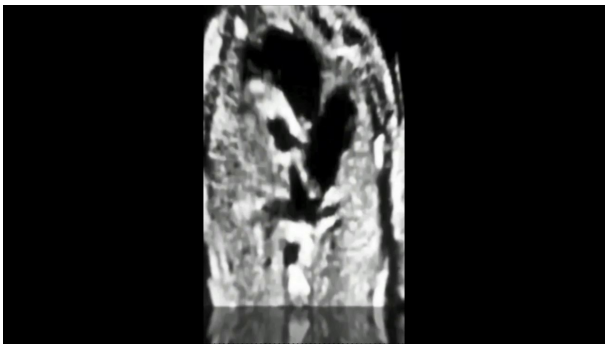


Figure 1.2: Hypoplastic left heart syndrome(HLHS) heart

   In patients with a patent and dysplastic mitral valve, color Doppler confirms the absence of left ventricle filling and displays reverse (left-to-right) flow across the foramen ovale and mitral regurgitation on the four-chamber image.

   It can be difficult to tell the difference between critical aortic stenosis, severe coarctation, and HLHS. The diagnosis of HLHS is more likely if there is no forward blood flow over the mitral or aortic valve.

Prognosis: HLHS remains the congenital heart abnormality with the highest fatality rate, notwithstanding recent brilliant accomplishments in pediatric cardio-surgery.

Despite all the technical advancements, there is a significant discrepancy in the sensitivities of ultrasonography for the detection of anomalies and the most likely reason for the same is the effect of the skill and experience of the monograph. In the RADIUS Trial, for example, there was a significant difference in the rates of anomaly detection between participating tertiary and non tertiary level centers [1]. Non tertiary level centers detected only 13% of congenital anomalies in the RADIUS Trial and were unable to detect any craniofacial, cardiac, gastrointestinal, or skeletal malformations. Tertiary level centers performed significantly better, detecting 35% of anomalies. Similarly, there is a significant variation in the skill levels of different doctors performing fetal echocradiography at different centers. Some other attributable factors include frequency of transducer, maternal build, period of gestation, abdominal scars, amniotic fluid index and position of the fetus.

The assessment of fetal echocardiography using cine loops is based on computer-assisted analysis of video frames in order to automatically monitor and evaluate fetuses without the need for manual intervention. Deep learning is a computer technology that generates feature maps based on artificial neural networks that can do specialized jobs with specified goals. Deep learning has made tremendous progress in the disciplines of computer vision. However, deep learning applications have not been extensively used in the field of fetal medicine as it's a relatively newer concept. Meanwhile, due to the diagnostic difficulty posed by a small and fast-beating fetal heart and caregivers' relatively little exposure to congenital heart disease, fetal echocardiography is very complicated, with inadequate/uneven skill in interpreting fetal cardiac pictures.

Artificial neural networks have an effective and promising structural innovation. Deep learning methods are continuously optimized. The study of fetal heart will be influenced by factors such as orientation, illumination, angle of view, and resolution.

Computer vision is a type of machine vision in which computers assist humans in recognizing, tracking, and understand visual objects. The rapid growth of computer vision and deep learning expands the uses of our vision while also deepening deep learning research. Manual annotation, on the other hand, is clearly inconvenient and time consuming. Because these labels are subjective, they cause bias and poor pattern categorization and retrieval accuracy.

The dataset used for training and validating deep learning models have been very important in estimating the precision and universality of the method. Hence, collection of data and its prepossessing has been very crucial and clearly defined in terms of

1. Selecting an appropriate study model.

2. Imaging modality.

3. Taking advice from experts in the respective field.

Preprocessing image datasets is a critical step in medical imaging that helps deep learning models perform better. There are various preprocessing techniques available for spatiotemporal image dataset. Cropping and selecting the regions of interest, scaling, filtering, and normalization of sequences are examples of such procedures. Data augmentation is another such technique which can be used to enhance spatiotemporal frames of time-series data.

1.2 Problem Statement

Deep learning is effective in video pattern recognition and has been applied to adult echocardiographic images, providing efficient classification of datasets that have been down sampled. Deep learning, however, has not been profoundly applied on fetal echocardiography cine loops either for classification into various standard planes or for anomaly detection.

Deep learning is the most effective method of completing the image recognition process. Although machine learning algorithms can recognize images, the majority of the feature selection work is done by humans.

The current deep learning method has shown wonderful results in a variety of domains, including image, voice, and natural language processing. Thus, I develop the research direction for deep learning in spatiotemporal fetal cardiac imaging. In this thesis, I purpose three deep learning algorithms to identify and classify the given fetal cardiac cine loop into standard planes and predict the CHD lesions. The three recognition models are CNN+LSTM, CNN+GRU and 3D CNN that will deal the visual time series problems in cine loops of fetal cardiac imaging. The methods need to be evaluated so as to achieve the best accuracy for fetal cardiac imaging.

I hypothesized that using input data curated according to clinical guidelines (i.e., selecting only the five screening cardiac views) would allow our models to detect diagnostic signals on small datasets and differentiate structurally normal fetal heart from the two major congenital heart abnormality i.e. Hypoplastic left heart syndrome (HLHS) and Tricuspid atresia (TA).

1.3 Objective

- To identify and classify the given cardiac cine loop into one of the five standard planes of fetal cardia used in fetal echocardiography.
- To demonstrate a deep learning based video processing tool for the prediction of common CHD lesions in particular Tricuspid atresia (TA) and Hypoplastic left heart syndrome (HLHS).

CHAPTER 2

LITERATURE REVIEW

2.1 Literature Review

There have been numerous video classification techniques proposed. Feature-based algorithms were used in older video classification techniques, while deep learning techniques have been used in the latest classification methods.

2.1.1 Feature Based Techniques

Classically text-based features, audio-based features and video-based features have been used perform the task of video classification. The methods employed by Rezeale and Wu was "closed captioning" [6] [7]. This is a useful method for hearing-impaired people in which the text of the speech is displayed on the screen as a text feature. Lin detected video clips using audio and video features, and these clips were then grouped into scenes [8]. To classify video, they used OCR (Optical Character Recognition) from detected clips and shut captioning. Brezeale used RGB color histograms as text-based features for video classification [6]. Jasinschi used probability values to represent six audio-based features: noise, speech, speech + speech, speech + noise, music, and speech + music. Hence, various feature based techniques have been implemented for video classification [9].

2.1.2 Deep Learning Methods

Deep Belief Networks (DBNs) were introduced by Hinton as a superior method of training each layer of the network [10]. This ignited research in deep neural networks resulting in numerous deep learning based models for video classification. A video clip was seen as a sequence of several frames in specific chronological order. Frame classification, a deep learning approach, is one of the most common video classification methods. In this method, features of each frame were extracted using Convolutional Neural Network(CNN). All frame-level features were averaged into a video-level representation as input to a classifier. Zha studied the performance of image-based video pattern recognition with features from different layers of deep learning models and several convolutional kernels for classification [11]. They showed this by integrating CNN features with kernel SVMs, they could achieve proper recognition performance. Frame classification methods do not account the relationship between frames and time which results in loss of temporal aspect and motion information.

CNNs are very successful at a wide range of activities since they learn features from data in the

form of an end-to-end pipeline focused at a given job. As a result, in contrast to image-based classification techniques, there are several works that specialize in using CNN models to discover hidden spatiotemporal feature patterns in extremely long videos. The temporal features in a video were used by end-to-end CNN architectures. Ji developed a 3D CNN model that runs on stacked video frames, extending the standard 2D CNN designed for images to the spatiotemporal domain [12]. Li introduced a 3D CNN algorithm that works on stacked sequence of video frames as an extension of the traditional 2D CNN design for images [13]. In 3D CNN, 3D kernels are used to extract spatiotemporal features. CNN training with 3D volume inputs is typically time-consuming. Ji introduced factorized spatiotemporal convolutional neural networks, in which the initial 3D convolution kernel learning is factorized as a sequential process of learning 2D spatial kernels within the bottom layer to efficiently handle 3D data [12].

Yi developed a two-stream technique after being motivated by the observation that movies may readily be split into spatial and temporal components [14]. This technique divided video representation learning into spatial and temporal cue feature learning. Using raw RGB frames as inputs, the authors initially gathered CNN to represent appearance information in a typical spatial context. Several frames Dense optical flows for temporal cues between neighboring frames were created, leading in temporal CNN training. The authors announced encouraging findings on two action recognition benchmarks. Ye considered a variety of alternatives, including dropout ratios and network topologies, because the two-stream method gave several implementation options that may impact performance [15].

Yi described a two-stream temporal CNN that collected motion information between consecutive frames [14]. However, the movements were only shown for a brief period of time. Because complex events in movies generally consist of numerous activities taking place over a lengthy period of time, the above-mentioned approach was insufficient for video analysis. This has motivated researchers to use RNN models in order to account for the temporal dynamics in movies. LSTM is one such RNN-based model that provides a good fit while avoiding the "vanishing gradient" effect. Its efficiency has been demonstrated in a variety of tasks, including image/video captioning. Akilan developed a combination of LSTM and CNN model outputs to simultaneously represent spatio-temporal cues for movie classification [16]. CNNs and LSTMs are highly complementary, according to the research. Feichtenhofer improved on the two-stream approach by investigating a more efficient method of combining spatial and temporal streams [17]. The two-stream approach was found to be superior at modeling correlations of spatial and temporal streams in their study [18] [19].

Kong suggested a technique for automated cardiac phase identification based on a CNN and RNN [20]. The researcher developed an LSTM to create a temporal regression that detects frames comprising the ED and ES phase throughout a cardiac cycle in their study. Xue and Debus, on the other hand, presented deep learning techniques to perform automated LV

evaluation and quantification directly from CMR cine sequences throughout the cardiac cycle [21] [22]. A deep multitask relationship learning network (DMTRL) for LV indices estimation and cardiac phase recognition was described in the study of Xue [21]. An LSTM was used in this study to represent the complicated temporal dynamics of the myocardium, which was then used to force the spatiotemporal consistency of successive frames. The design is made up of two parallel LSTMs, the first of which learns the LV deformation during the cardiac cycle to estimate the cardiac indices, and the latter of which records the spatiotemporal changes between successive frames to identify the cardiac phases. Similarly, Debus and Ferrante [28] carried out the identical tasks, but instead of employing an RNN, they added a spatiotemporal CNN that extracts spatial and temporal information using 3D convolutional kernels [22]. These 3D convolutions use the spatiotemporal information of successive frames to enhance quantification and cardiac phase detection accuracy [21] [22].

CHAPTER 3

RATIONALE OF THE STUDY

3.1 Rationale of the study

Fetal scanning and fetal echocardiography are specialties which require highly skilled expertise. Such training programs are very competitive and expensive. None such training programs are available in our country. As such very few doctors with the required skills are available in our country. The program which I have tried to develop can be used as a screening tool in any setup to identify high risk cases with fetus having possibility of having CCDs and they can be further be scanned by the experts for confirmation. This will reduce the work load on the experts and prevent diagnosis lapse in hospitals where experts are not available.

The main contribution of this thesis is to find out the best method for the identification and classification of the given fetal echocardiography into five standard planes and predict the CHD lesions. The evaluations are based on analyzing the optimal model. The current goal is aimed at finding the optimal algorithm. Multiple deep learning methods, i.e. CNN+LSTM, CNN+GRU and 3D CNN are proposed and presented on details. These three approaches based on deep learning will be justified and analyzed. Based on the outcomes of the experiments, the best recognition algorithm will be determined.

CHAPTER 4

METHODOLOGY

4.1 Dataset Collection and Description

Ultrasound (US) is still an important diagnostic technique for fetal echocardiography. This method allows for detailed morphological assessment of fetal cardiac structures. Ultrasound uses a physical medium such as air, water, or tissue to propagate high frequency sound waves from a transducer. Medical ultrasound devices employ pressure pulses with frequencies ranging from 1 to 15 MHz for diagnostic purposes. The examinations are often carried out in real time and with the ability to observe multiple planes of the fetal heart depending on the location of the transducer.

In this study, five standard screening views of normal fetal heart and CHD ultrasound data were used to diagnose fetal cardiac health. With waived consent, all data were de-identified to maintain the privacy of the person. The data set was split into three sections: training, validation, and testing, in the proportions of 60%, 20%, and 20%, respectively.

4.1.1 Dataset Collection

The data for this study was collected from several ultrasound centers. These data were recorded on Samsung HS40, GE Voluson S10, GE Voluson E8, GE Voluson E10 and Samsung HERA W10. It consists of videos of 5 different standard views of normal fetal heart and CHD lesions. The 5 canonical views used were – "Abdominal view (ABDO)", "4 Chamber (4C)", "3 Vessel View (3VV)", "Left ventricular outflow tract (LVOT)" and "Right ventricular outflow tract (RVOT)". Similarly, congenital heart diseases considered in this study were Tricuspid atresia(TA) and Hypoplastic left heart syndrome(HLHS) . There were a total of 564 videos. The video was recorded at a rate of 30 FPS(frames per second) by fetal medicine experts. The average length of each the video was about 6 to 10 seconds. Table 4.1 represents number of cine loops collected.

Table 4.1: Number of cine loops in each category of spatiotemporal fetal cardiac imaging

S. No.	Type	Number of Videos
1	3VV	40
2	4C	152
3	ABDO	60
4	LVOT	52
5	RVOT	60
6	TA	100
7	HLHA	100
Total		564

4.2 Baseline Models

4.2.1 Convolutional Neural Network(CNN)

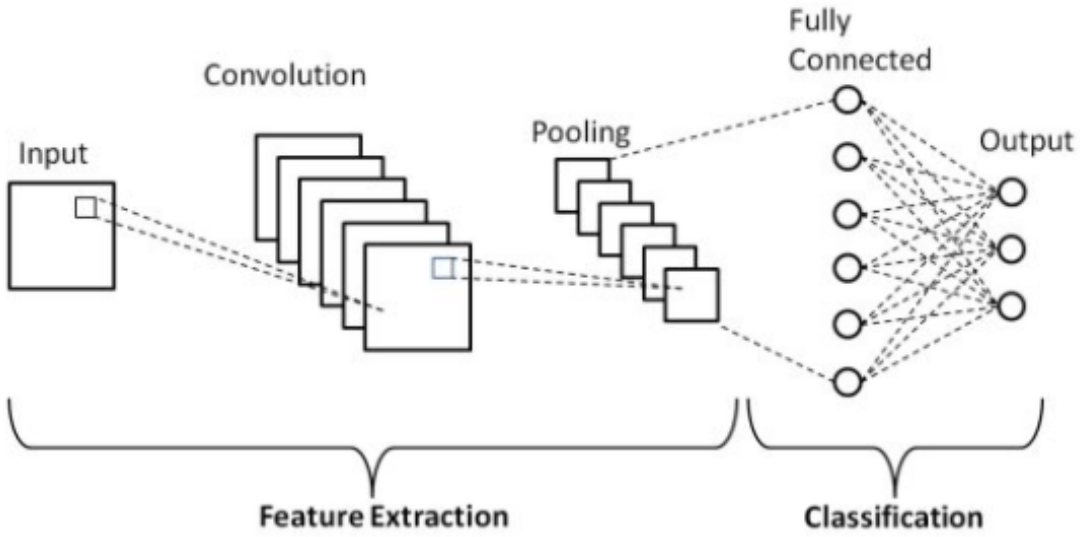


Figure 4.1: Block diagram Convolutional neural network (CNN)

CNNs are great at finding patterns and using them to classify images. They are made up of numerous hidden layers. It is made up of filters, kernels, or neurons with learnable weights and parameters. Each channel receives a few inputs and performs convolution as shown in Figure 4.1. The components of CNN comprise of :

i. Convolutional Layer

ii. Rectified Linear Unit (ReLU) Layer

iii. Pooling Layer

iv. Fully Connected Layer

Convolutional Layer

Convolution layer calculates the output of the neurons that depends only on a small number of inputs, each computing a dot product between a small number of inputs and their corresponding weights. This layer is the central component of a CNN and is responsible for the majority of the computational calculations. Its main function is feature extraction. It retains the spatial relation between pixel and input image. A simple example of Convolution layer is shown in Figure 4.2

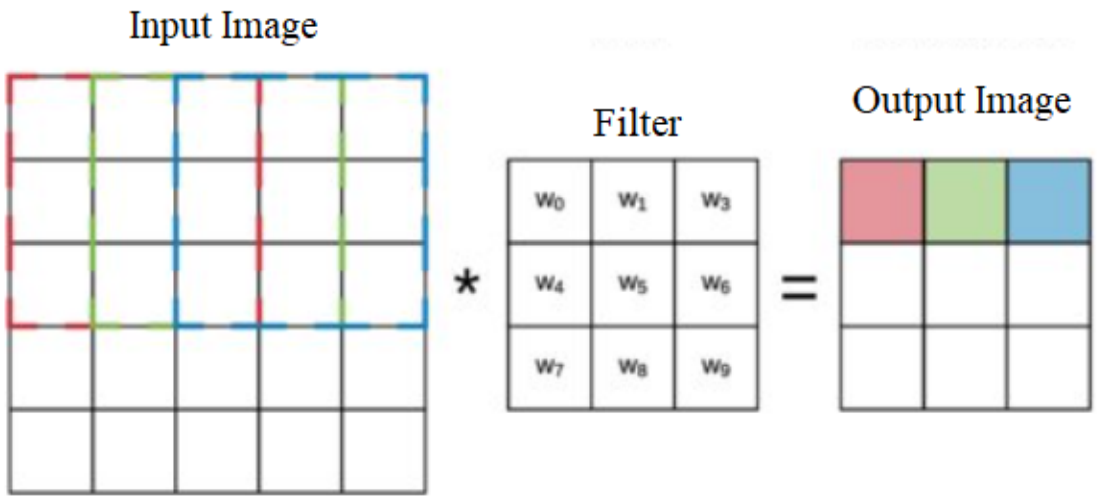


Figure 4.2: Convolution Layer for Model Building.

Activation function

Activation function has the ability to add non-linearity to the network in order to learn complex patterns in the data. It converts the output from the previous cell into some other form that can be used as the input to the next cell. The most commonly used activation function is ReLU. It is a component wise operation that reconstitutes all negative values within the highlight outline to zero. It is defined as R(z) = max (0, z), is not differentiable at z = 0. The plot of ReLU activation function is shown in Figure 4.3

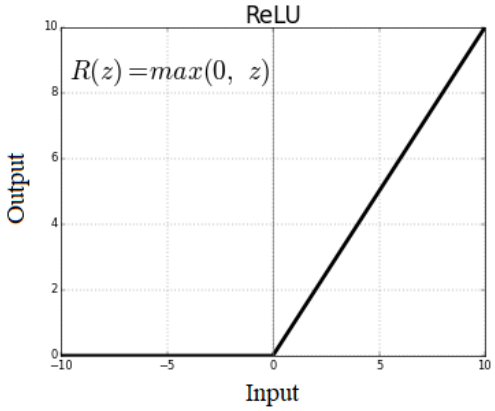


Figure 4.3: ReLU Activation Function.

Pooling Layer

Pooling layers are frequently used in convolution neural networks to reduce the dimension of the input for faster computation. It is applied to shrink the stack of images. To design pooling layer, we need hyperparameters as

f: filter size

s: stride

Here, Max Pooling is shown in Figure 4.4 with kernel size 2 and stride 2. Steps for implementing Max Pooling layer are

1. Choose a window size (2).
2. Select a stride.
3. Window should be walked across filtered images..
4. Take the maximum value for each window.

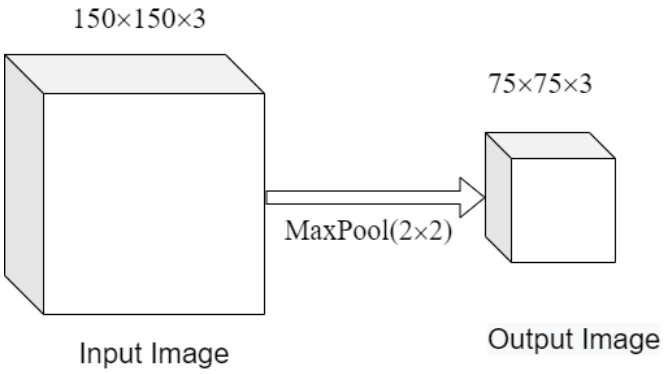


Figure 4.4: Max Pooling Layer.

There is no parameter to learn in pooling.

Fully Connected Layer

Classification of data into various classes can be done using fully connected neural network. The fully connected layer's primary function is to use the input from the previous convolution layer or pooling layer to classify the given image into its corresponding label. The previous layer's input is connected to the activation unit of the next layer in this layer. layer[23][24]. It can be implemented as

1. Fully connected input layer(Flatten):- The fully connected input layer takes input from the previous layer, flattens it, and converts it into a single vector that may be used as input to the following layer.
2. Fully Connected Layer:- The fully connected layer takes the previous layer's input and predicts the label based on feature analysis.
3. Fully Connected Output Layer:- The probabilities of each label in the data are computed by the fully connected output layer.

Softmax Classification

Softmax is an activation function often used to classify multiclass problem. It normalizes the output of a network into a probability distribution of each of possible outputs. It is also called a vector of probability distribution. It produces values in the range of 0-1, therefore used as the final layer in classification models[25].

$$\sigma(z)_i = \frac{e_i^z}{\sum_{j=1}^{K} e_j^z} \tag{1}$$

4.2.2 Recurrent Neural Network(RNN)

Recurrent neural networks are a type of neural network that can interpret information and can recognize patterns in ordinal, temporal or sequential data.

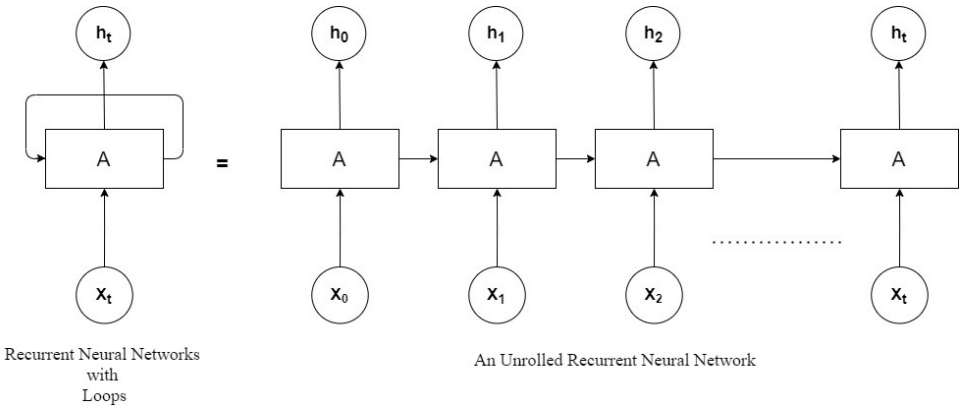


Figure 4.5: Recurrent Neural Network

Figure 4.5 shows that A is a neural network block that takes X_t as an input and produces h_t as an output. The loop transports information from one step of the network to the next. This loop may be regarded as many clones of the same network, each transmitting information to the next phase. But, RNN has difficulty in accessing information from the remote past-long-term dependency problem. It uses backpropagation algorithm for every timestamp, called back-propagation through time(BTT). Additionally, RNN has vanishing gradient problem and exploding gradient problem.

Long Short-Term Memory Networks(LSTM)

LSTM are a special type of RNN that are specifically designed to address the issue of long-term dependency. They can remember the given information for very long durations of time. So, they are mainly focused on resolving the vanishing gradient problem itself.

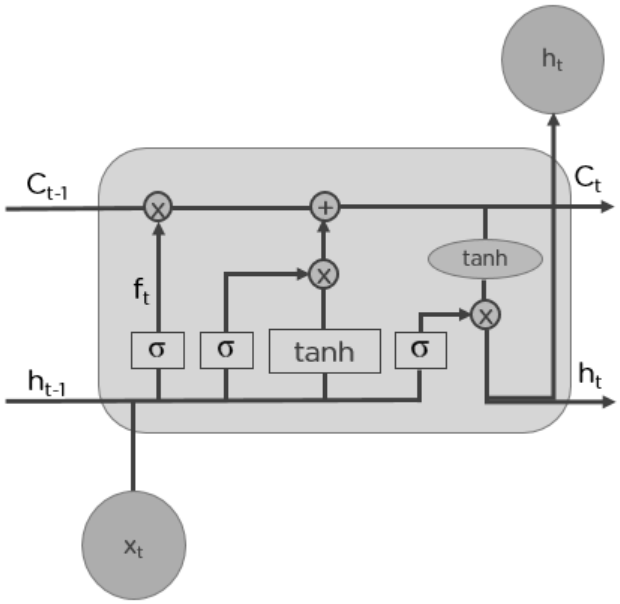


Figure 4.6: Block diagram of long short term memory(LSTM)

1. **Memory Cell(Cell State)**

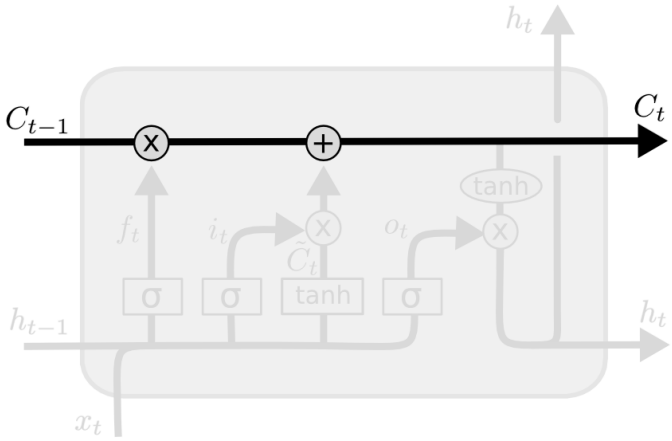


Figure 4.7: Memory cell of LSTM

The memory cell is used to remember and forget things based on the context of the input. It is the horizontal line running across the top of the diagram with some minor linear interactions as shown in Figure 4.7.

2. **Forget Cell**

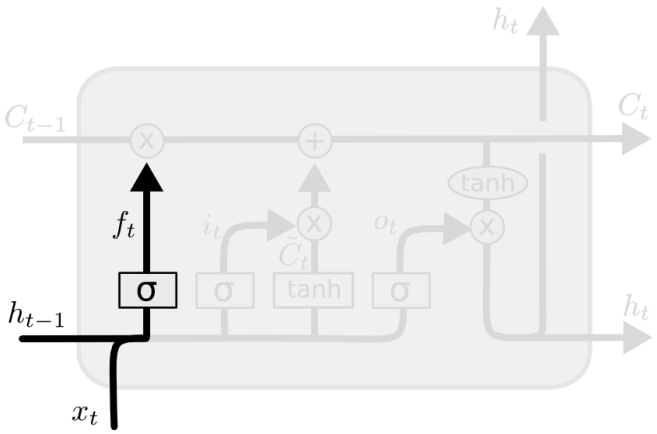


Figure 4.8: Forget gate of LSTM

The first step in LSTM is deciding which information should be removed from the memory cell. The forget gate sigmoidal layer makes this decision. This layer examines the previous state h_{t-1} and the current input x_t to compute a value between 0 and 1 for each number in the memory cell C_{t-1}.

$$f_t = \sigma(W_f * [h_{t-1}, x_t] + b_f)$$

where f_t = forget gate

3. **Input Gate**

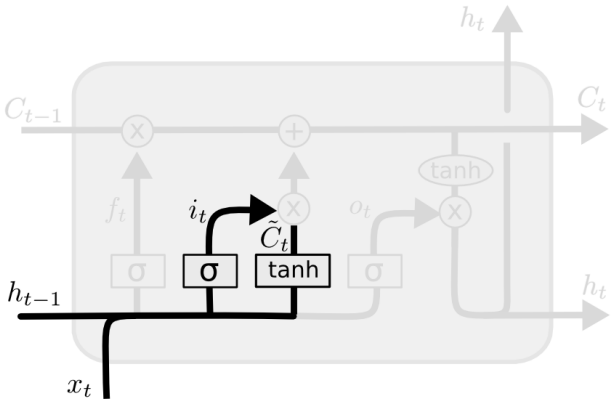


Figure 4.9: Input gate of LSTM

The quality and quantity of information stored in the memory state are determined by the input layer. This procedure is carried out in two stages.

(a) Sigmoidal function decides which values will be allowed to pass.

(b) A tanh function generates a vector of new candidate values c_t that assigns weightage to the values passed to it..

$$i_t = \sigma[W_i * [h_{t-1}, x_t] + b_i]$$
$$c_t = tanh[W_c * [h_{t-1}, x_t] + b_c]$$
where i_t = input gate

The old cell state, C_{t-1}, is now updated to the new cell state, C_t. The next step includes implementation of previously decided plan.

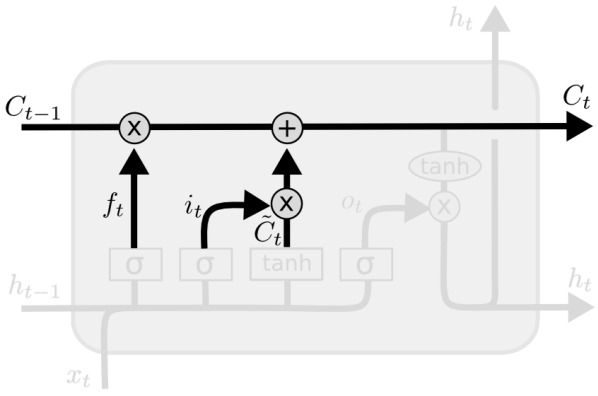


Figure 4.10: LSTM cell state update

And the updated equation is given as

$$C_t = f_t * C_{t-1} + i_t * C_t$$

4. **Output Gate**

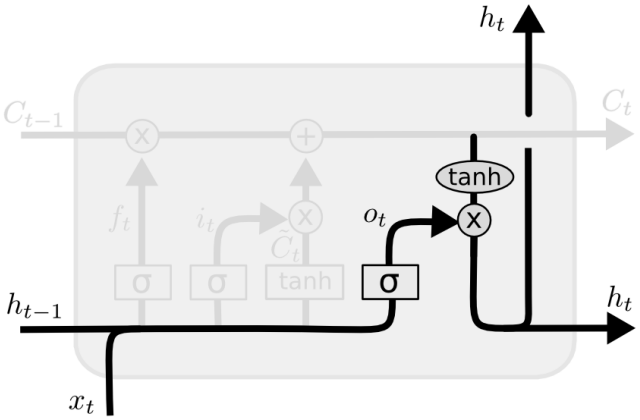


Figure 4.11: LSTM output gate

Concatenation of information is again passed through the sigmoidal function. Then the information from the memory cell is passed through the tanh function. To obtain the output, the point-by-point operation is now performed.

$$o_t = \sigma * (W_o * [h_{t-1}, x_t] + b_o)$$
$$h_t = o_t * tanh(C_t)$$
Where o_t = Output gate

Gated recurrent units (GRUs)

GRU is a relatively new addition to the Recurrent Neural Networks family and is very similar to LSTM.

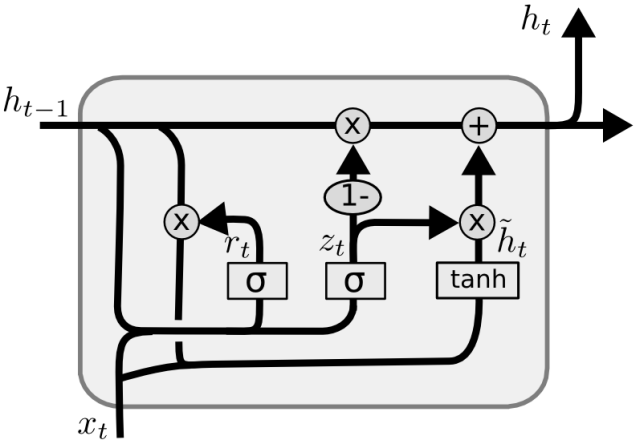


Figure 4.12: Diagram of gated recurrent units

It only has two gates: one for reset gate and another for update gate.

1. **Update gate**-combines the forget and input gates into the Update gate.

2. **Reset gate**-combines the hidden and the cell states to determine how much information to forget.

Mathematical expressions of GRU is given as

$$z_t = \sigma(W_z * [h_{t-1}, x_t])$$
$$r_t = \sigma(W_r * [h_{t-1}, x_t])$$
$$h_t = tanh(W * [r_t * h_{t_1}, x_t])$$
$$h_t = (1 - z_t) * h_{t-1} + z_t * h_t$$

GRU has fewer tensor operations than LSTM; hence , they are slightly faster to train in GRU than LSTM.

4.3 Main Models

Modern medical imaging technology allows for the recording of cine loops and time series sequences for analysis of the fetal heart and diagnosis of any abnormalities[18]. Figure 5.3 highlights the most widely utilized dynamic imaging modalities as well as therapeutic applications for deep learning in 3D/4D cardiac image sequences.

4.3.1 CNN+RNN Model

The spatiotemporal classification process consists of the following steps :

i. Data Collection .

ii. Pre-processing.

iii. Spatial Feature Extraction by CNN .

iv. Temporal Feature Extraction by RNN.

v. Classification

vi. Recognition.

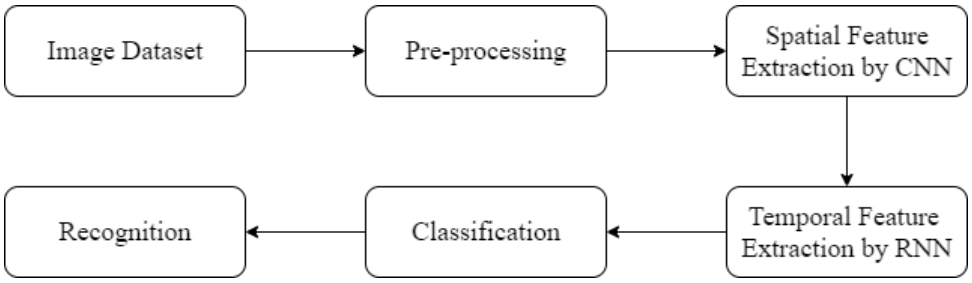


Figure 4.13: System Block Diagram for CNN+RNN Spatiotemporal Fetal Heart Imaging

Deep learning and machine learning have been utilized effectively in a variety of medical image analysis applications. The spatiotemporal analysis of fetal cardiac imaging sequences requires the extraction of spatial and temporal features in order to properly represent the relevant information throughout time. The CNN+ RNN (LSTM/GRU) architecture includes CNN layers for spatial feature extraction on 2D image input data combined with RNN(LSTM/GRU) to support temporal sequence predictions.

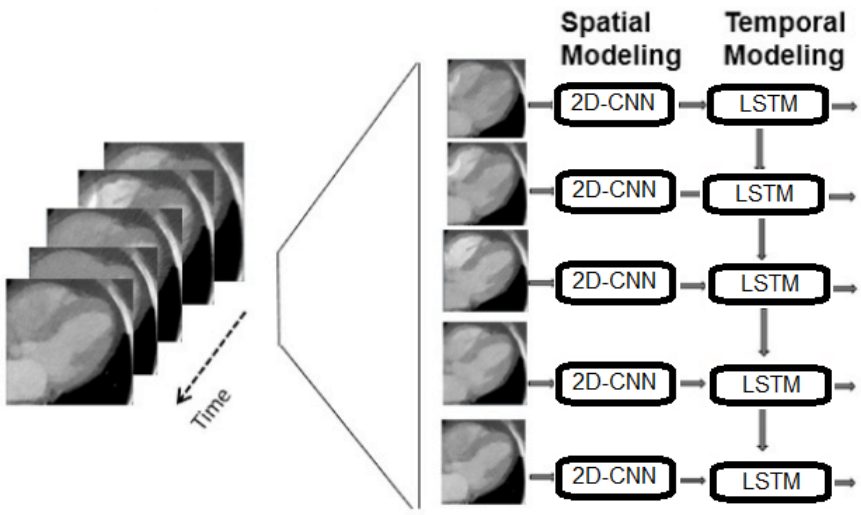


Figure 4.14: Spatiotemporal feature extraction of heart using CNN and RNN(LSTM)

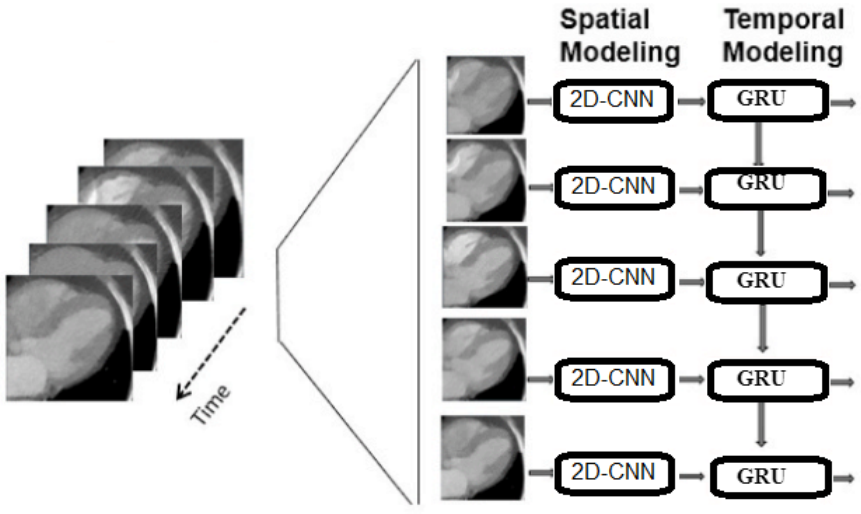


Figure 4.15: Spatiotemporal feature extraction of heart using CNN and GRU

CNNs are effective at finding spatial features and using them to classify images. They are made up of numerous hidden layers. They are made up of filters, kernels, or neurons with learnable weights and parameters. Each channel takes a few inputs and convolves them. Max pooling layer and fully-connected layers come next. Different CNN architectures have been utilized in clinical image analysis for anatomical structure classification and segmentation. Recurrent neural networks are special a type of neural network that can interpret information and can recognize patterns in ordinal, temporal or sequential data. But, as described earlier RNN has a long-term dependency problem and a vanishing gradient problem. To address these issues, LSTM and GRU were used for temporal sequence prediction. Therefore, CNN+LSTM and CNN+GRU model was used for spatiotemporal fetal cardiac imaging.

4.3.2 3D CNN Model

3D CNN are similar to 2D CNN. 2D Convolutional Neural Network (2D CNN) performs well for spatial feature extraction from 2D images, but it does not account for temporal information from the same image. Therefore, 3D CNN is used which is suitable for spatiotemporal feature extraction. Block diagram is shown in Figure 4.16. Some of the 3D CNN properties are

1. 3D CNN can model both spatial and temporal features simultaneously.
2. 3D filter of the size 3x3x3 includes spatial as well as temporal features.

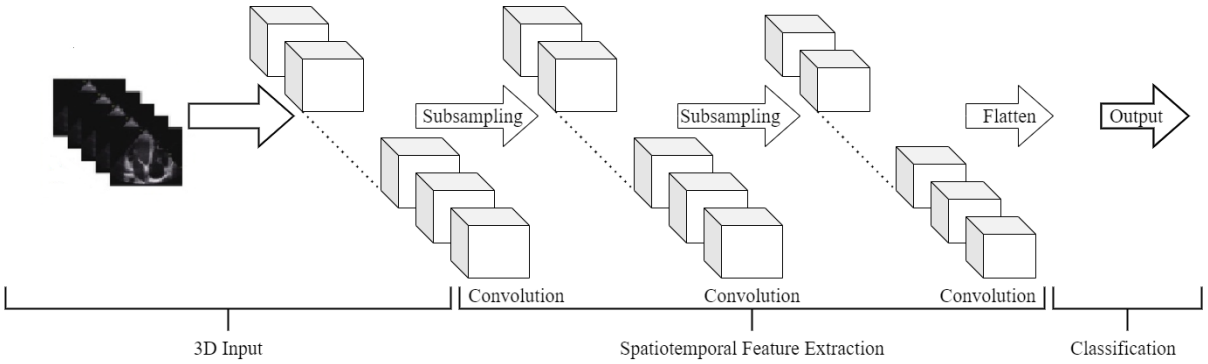


Figure 4.16: Block diagram of 3D CNN

The components of 3D CNN are similar to that of 2D CNN and is given as

1. **Convolution Layer** : Voxels are used in 3D CNN in place of pixels which were used in 2D CNN. The operation of 3D CNN is illustrated in Figure 4.17.

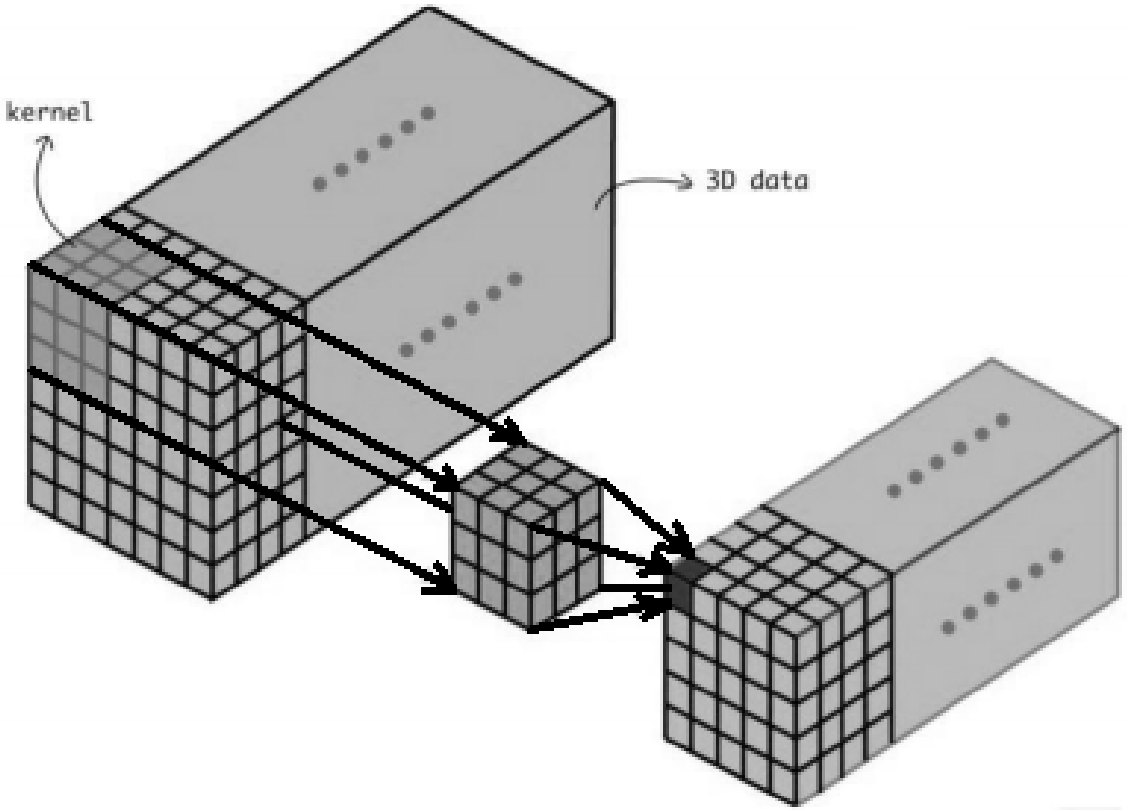


Figure 4.17: Graphic representation of 3D-convolution layer

2. **3D maxpool(2x2x2)** : In 3D Maxpool layer, the maximum value within the volume is selected. In this case, the maximum element within the cube of size 2 is selected.

3. **Fully Connected Layer** : A fully connected layer is also known as a feed forward neural network. The output of the last pooling layer or convolutional layer is flattened and then fed into the fully connected layer as its input.

4.4 Model Evaluation

The performance of a classification model is measured by the number of test records correctly and erroneously predicted by the model. The confusion matrix gives a more detailed view that includes not only the performance of a predictive model, but also which classes are predicted correctly and erroneously, as well as the types of errors produced.

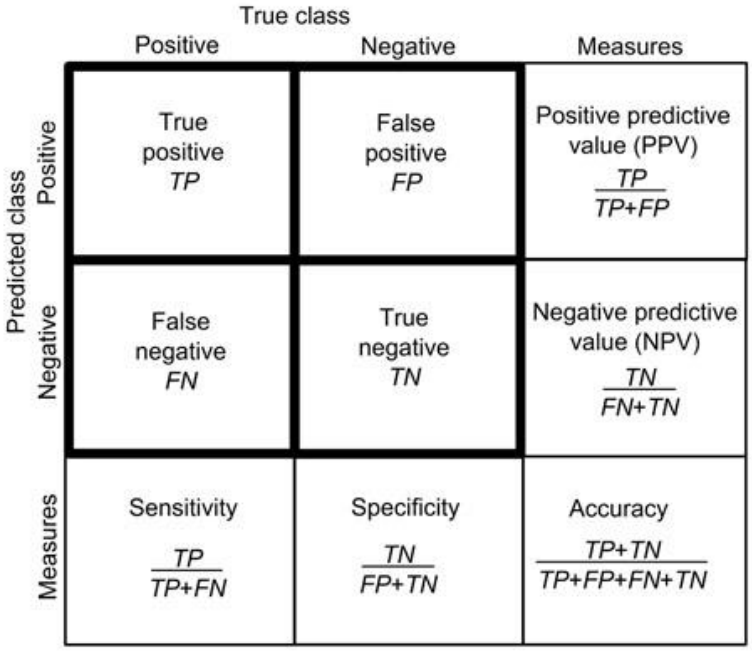


Figure 4.18: Confusion Matrix for evaluating spatiotemporal fetal cardiac models.

1. True Positive (TP): The model classifies the example as positive, and also positive for the actual mark.

2. False Positive (FP): The model categorizes the example as positive, but negative is the actual mark.

3. Real Negative (TN): The model categorizes the example as negative, and negative is also the actual mark.

4. False Negative (FN): The model labels the example as negative, but the mark remains positive.

4.4.1 Sensitivity, Specificity and Accuracy

Sensitivity tests the percentage of positives recognized correctly (e.g. the proportion of images that are correctly identified as true images). It is calculated as,

$$Sensitivity = TP/(FN + TP)$$

Specificity tests the number of negatives correctly recognized (e.g. the percentage of false images which are correctly identified as false images). It is expressed as,

$$Specificity = TN/(FP + TN)$$

Thus, the accuracy that reflects a better measurement metric can be obtained as:

$$Accuracy = Sensitivity * Prevalence + Specificity * (1 - Prevalence)$$

or

$$Accuracy = (TN + TP)/(FP + FN + TP + TN)$$

Where prevalence is defined as the ratio of the number of anomaly cases in a given population..

4.4.2 The AUC ROC curve (Receiver Operating Characteristics) (Area Under The Curve)

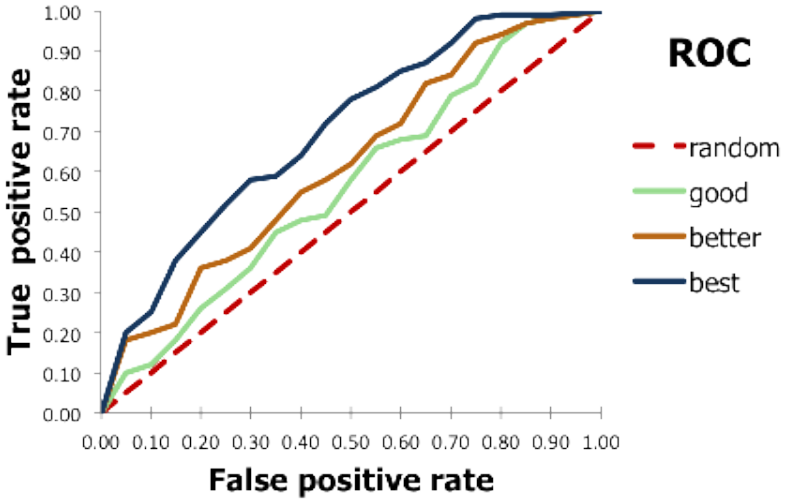


Figure 4.19: ROC Curve for evaluating a given model

The given model will predict the true positives or negatives and all other metrics mentioned above based on the assumption of a threshold that dictates what output label is considered positive and negative. A receiver operating characteristics (ROC) curve, which is developed at various threshold settings by plotting, is a method that helps us see how the threshold plays out the model's decision. False Positive Rate vs. True Positive Rate [26].

AUC ROC provides us the measure of goodness of fit, summarizes the model output across all thresholds, and provides a good sense of the discriminative power of a given model.

F1 Score

The F1 score is a measure that combines precision and recall into a single rating. It is a harmonic mean of sensitivity and precision[27]. It is preferable to obtain a single score that reflects both sensitivity and precision.

$$F1 = \frac{2}{\frac{1}{Recall} + \frac{1}{Precision}}$$

4.5 Tools and Resources

The following tools and resources will be used in this study:

1. **Python**: Python is a high-level, interpreted programming language that was employed in this study.
2. **Anaconda**: Anaconda is a data science platform to simplify package management and deployment for scientific computing.
3. **Jupyter Notebook**: Jupyter notebook is a computational notebook that allows us to write and execute the python code in local web browsers.
4. **Tensorflow**: Tensorflow is a machine learning library that is available as open-source software. Its library can be run on all kind of computers irrespective of their processing powers.
5. **Keras**: Keras is an open source library which supports almost all types of neural network models.
6. **OpenCV**: OpenCV is a free and open source computer vision library that may be used in real-time.
7. **Numpy**: Numpy is a MATLAB-style library for numerical operations that is well optimized.
8. **Tflearn**: Tflearn is a high level API built on top of Tensorflow which supports deep learning models.
9. **Matplotlib**: Matplotlib is a Python-based data visualization and plotting library.
10. **Scikit Learn**: Scikit-learn is a Python machine learning library for analysis and prediction of scientific computing.
11. **Theano**: Theano is a Python library that allows you to define, optimize, and evaluate multi-dimensional arrays quickly.
12. **TensorBoard**: TensorBoard is a tensorflow visualization toolkit.

CHAPTER 5

METHODOLOGY IMPLEMENTATION

5.1 Implementation

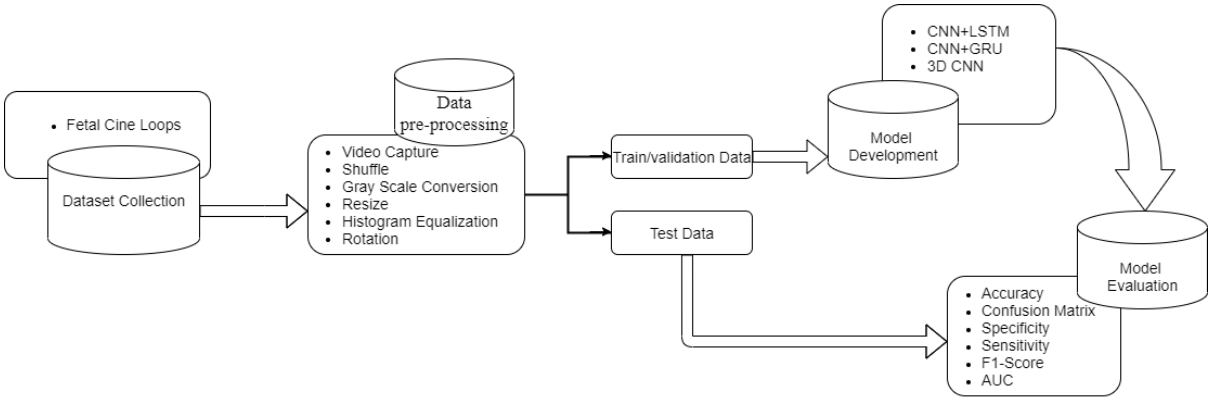


Figure 5.1: The Overall proposed system architecture of spatiotemporal fetal cardiac imaging system

Figure 5.1 illustrated the overall system for fetal cardiac cine loops identification and classification along with the prediction of common CHD lesions. 30 FPS cine loops of varying lengths (6-10 seconds) were converted into images and then passed through the preprocessing pipeline. Image capturing, shuffling, gray scale conversion, resizing, histogram equalization and rotation were done during the preprocessing method. The preprocessed data was then divided into three groups: "training", "validation", and "testing", in proportions of 60%, 20%, and 20% respectively. The models CNN+LSTM, CNN+GRU and 3D CNN were trained using the training dataset and training accuracy and loss were calculated after each epoch or training steps. The validation loss and accuracy were also obtained during the training phase. Finally, the performances of the trained models were evaluated with evaluation metrics, confusion matrix, accuracy, sensitivity, specificity, AUC and F1-score.

Google Colab and anaconda 4.10.1 has been used to implement the deep learning model. Colab is a cloud environment for Jupyter notebook that includes GPUs and TPUs and it is used for high computation. Python programming has been used to create the code. For data pre-processing opencv-python, numpy, pandas and theano has been used to generate feature matrix and target vector. The resulting datasets were split into training, validation and testing using scikit-learn. The Keras and Tensorflow library have been used to implement the spatiotemporal CNN+LSTM and CNN+GRU model in google colab. Tensorflow 2.2.0 and keras 2.3.1 have been used to implement 3D CNN model in local anacoda enviroment. Scikit-learn has been used for model evaluation. For graph visualization, the Matplotlib and tensorboard have been used.

5.1.1 Pre-processing

A video is a collection of set of images called frames, arranged in specific chronological order. For spatiotemporal classification, 30 FPS cine loops of varying lengths (6-10 seconds) were converted into images then passed through the preprocessing pipeline. Image capturing, shuffling, gray scale conversion, resizing, histogram equalization and rotation were done during the preprocessing.

The network architecture and the data format have been crucial while constructing an efficient neural network model. The number of images, number of channels, image depth, image height, and image width were all common parameters.

Pre-processing for CNN+RNN Model

All images collected from different sources, including RGB images were converted to grayscale images (size - 150 X 150). The histogram equalization technique was used to improve image contrast by stretching the intensity range. Following that, image augmentation was done. Before resizing the image, the region of interest was extracted to emphasize the five standard cardiac planes as shown in Figure 5.2 .

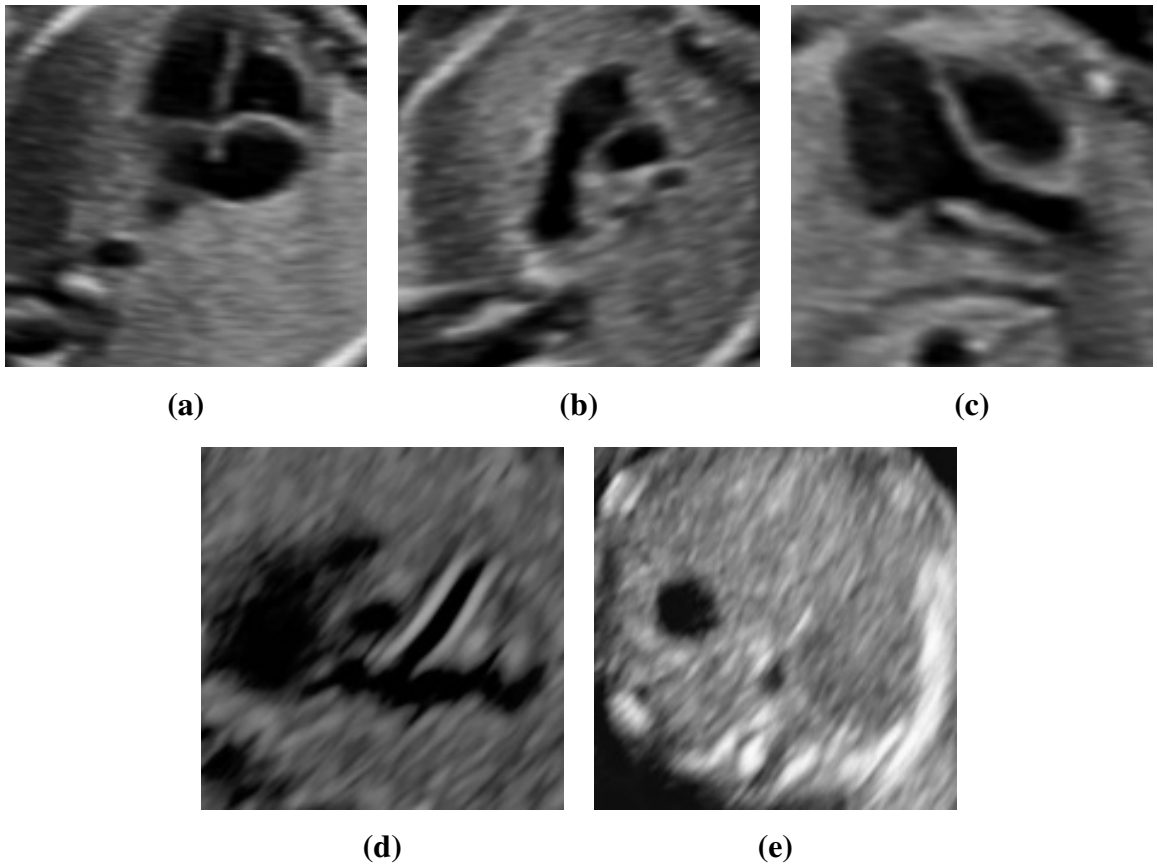


Figure 5.2: Five standard views of fetal heart **(a)** Four Chamber (4Chamber) **(b)** Three Vessel View (3VV) **(c)** Left Ventricular Outflow Tract(LVOT) **(d)** Right Ventricular Outflow Tract(RVOT) **(e)** Abdominal(ABDO)

Pre-processing for 3D CNN Model

Videos were of varying number of frames per second so it was normalized to a sequence duration of 30 frames. This was accomplished by erasing or repeating equidistantly spaced frames. Following that, each video frame was converted into grayscale frame. Finally, both channels of each sequence (Grayscale and Depth) were normalized to have a mean of zero and a variance of one unit. This was done so the system could converge faster. Further, images were resized to 150 x 150 x 43 voxel to enable them to fit in the 3D-CNN model as shown in Figure 5.3.

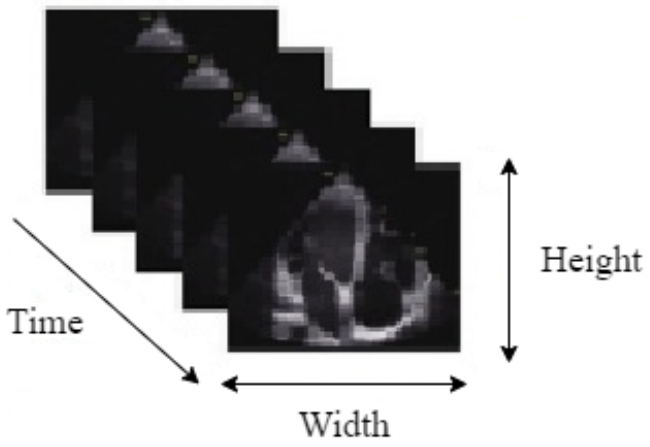


Figure 5.3: Spatiotemporal Cardiac image sequences (Voxel)

The fetal heart rate ranges between 110-160 beats per minute. On an average, I have considered 130 beats per minutes which signifies that 1 fetal cardiac cycle cycle is completed in 0.47 seconds. To ensure reliability, 3 cardiac cycles were considered in a sequence averaging 1.41 seconds. As, there are 30 frames in 1 second, there will be 43 frames in 1.41 second (1.41x30=42.3). Hence, voxel size of 43 frames was considered in this study.

5.2 Development of Models

The proposed models were developed as CNN+LSTM, CNN+GRU and 3D CNN for fetal cardiac cine loops identification and classification, as well as the prediction of common CHD lesions.

5.2.1 Identification and classification into five standard planes

Combined CNN+RNN(LSTM/GRU) models

A combined method was developed to automatically identify and classify five types fetal cardiac images. The CNN+ RNN (LSTM/GRU) architecture includes 2D CNN layers for spatial feature extraction on 2D image input data combined with RNN(LSTM/GRU) to support temporal sequence predictions.

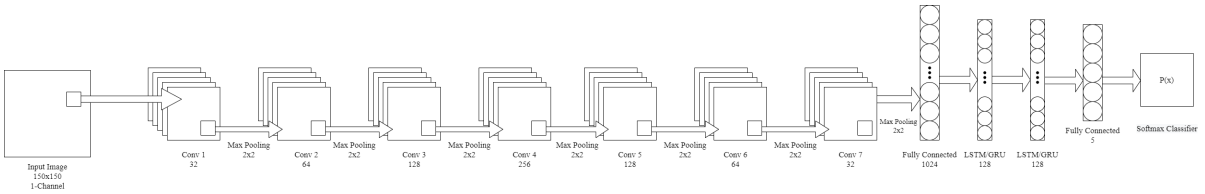


Figure 5.4: CNN+RNN(LSTM/GRU) architecture for standard five view classification of fetal heart.

Figure 5.4 illustrates the combined CNN+RNN(LSTM/GRU) models for spatiotemporal fetal cardiac imaging. The network has 7 convolutional layers followed by pooling layer, one fully-connected layer, two LSTM/GRU layer and one output layer with the softmax function. The convolutional layer is used for feature extraction that is activated by the RELU function. The max pooling layer with size of 2x2 kernels is used for dimensionality reduction of the given input image. In the last part, the function map is transferred to the LSTM layer followed second LSTM layer to extract temporal information. After analyzing the temporal features, the architecture sorts the fetal cardiac images through a fully connected layer to predict whether they belong under any of the five categories.

3D CNN model

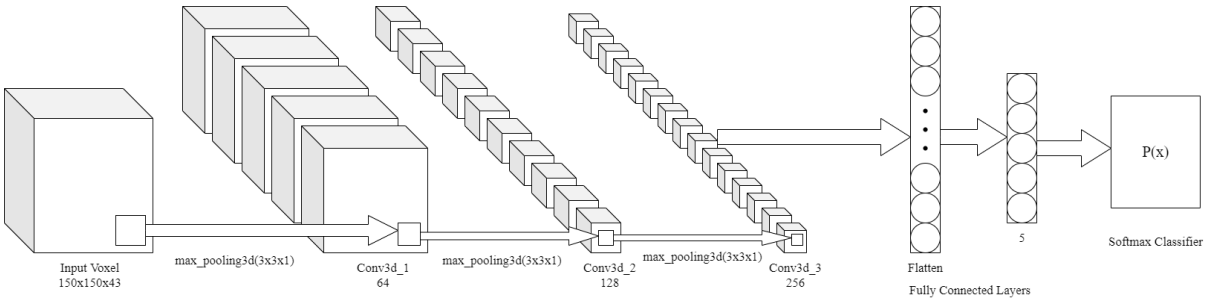


Figure 5.5: 3D CNN architecture for standard five view classification of fetal heart.

Figure 5.5 illustrates 3D CNN model for spatiotemporal fetal cardiac imaging. The network has three 3D convolutional layers followed by pooling layer, one fully-connected layer and one output layer with the softmax function. The 3D convolutional layer is used for spatiotemporal

feature extraction that is activated by the RELU function. The 3D max pooling layer with size of 3x3x1 kernels is used for dimensionality reduction of the given input image. In the final layer, the architecture sorts the fetal cine loops through a fully connected layer to predict whether they belong under any of the five categories.

5.2.2 Diagnosis of Normal Heart vs. CHD lesions

Combined CNN+RNN(LSTM/GRU) models

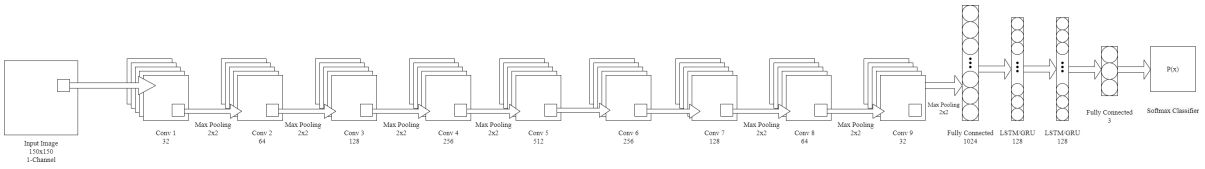


Figure 5.6: CNN+RNN(LSTM/GRU) architecture for normal heart vs. CHD lesions.

Figure 5.6 illustrates the combined CNN+RNN(LSTM/GRU) models for spatiotemporal fetal cardiac imaging. The network has 9 convolutional layers followed by pooling layer, one fully-connected layer, two LSTM/GRU layer and one output layer with the softmax function. The convolutional layer is used for feature extraction that is activated by the RELU function. The max pooling layer with size of 2x2 kernels is used for dimensionality reduction of the given input image. In the last part, the function map is transferred to the LSTM layer followed second LSTM layer to extract temporal information. After analyzing the temporal features, the architecture sorts the fetal cardiac images through a fully connected layer to predict whether they belong under any of the three categories.

3D CNN model

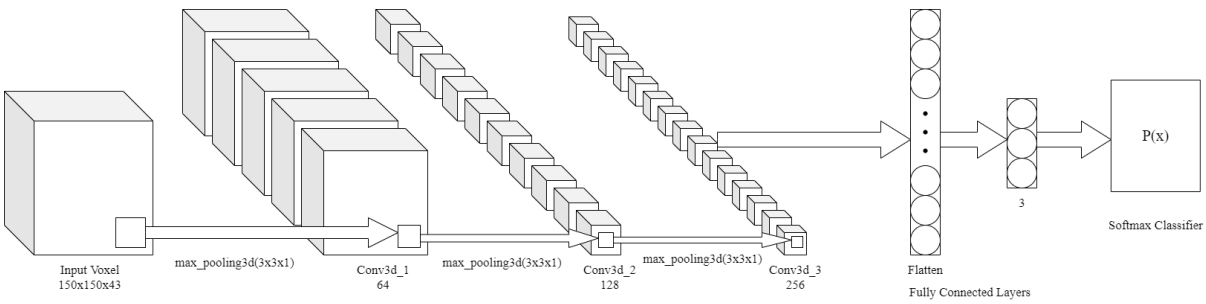


Figure 5.7: 3D CNN architecture for Normal Heart vs. CHD lesions.

Figure 5.7 illustrates 3D CNN model for spatiotemporal fetal cardiac imaging. The network has three 3D convolutional layers followed by pooling layer, one fully-connected layer and one output layer with the softmax function. The 3D convolutional layer is used for spatiotempoal

feature extraction that is activated by the RELU function. The 3D max pooling layer with size of 3x3x1 kernels is used for dimensionality reduction of the given input image. In the final layer, the architecture sorts the fetal cine loops through a fully connected layer to predict whether they belong under any of the three categories.

CHAPTER 6

RESULTS AND DISCUSSION

A fetal cardiac imaging system with fully functional Spatiotemporal deep learning models CNN+LSTM, CNN+GRU, and 3D CNN has been developed. The data was split into three categories: training, validation, and testing, with a 60%, 20%, and 20% split, respectively.

6.1 Classification and performance evaluation of five standard view fetal heart

The three systems were able to classify 364 fetal cardiac cine loops into five standard fetal heart views, namely ABDO, 4C, 3VV, LVOT, and RVOT with an accuracy of 97.39%, 98.02% and 91.35% respectively.

- Total Samples of videos : 364
- Number of epochs : 10
- Learning Rate : 0.001
- Optimizer : Adam Optimizer
- Training Accuracy :
    1. CNN+LSTM : 97.39%
    2. CNN+GRU : 98.02%
    3. 3DCNN : 91.35%

A random test result of the classification of five standard views of fetal heart is shown in Figure 6.1

Figure 6.1: Random test result of the five standard screening views of the fetal heart

The training accuracy of CNN+LSTM, CNN+GRU, and 3D CNN is 97.39%, 98.02%, and 91.35% respectively, with corresponding crossentropy losses of 0.14912, 0.10478, and 5.8198. Figure 6.18, Figure 6.19 and Figure 6.20 depicts a plot of accuracy vs. epochs (training steps) for respective models. The training accuracy summary is explained using the bar diagram Figure 6.2 . The given figures indicate that the model performance has been improving over time leading to conclusion that the models are learning.

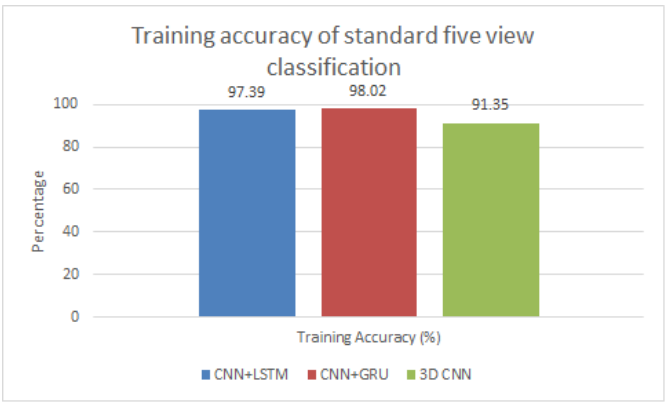
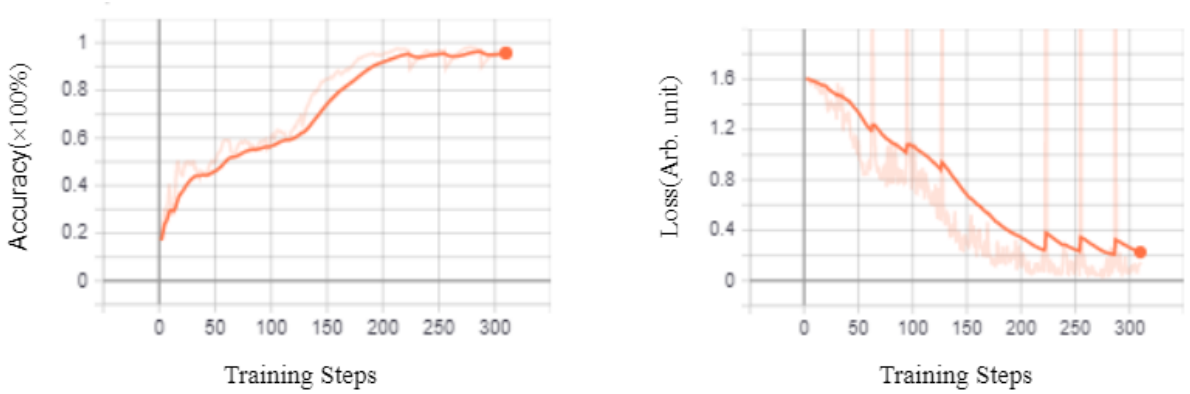
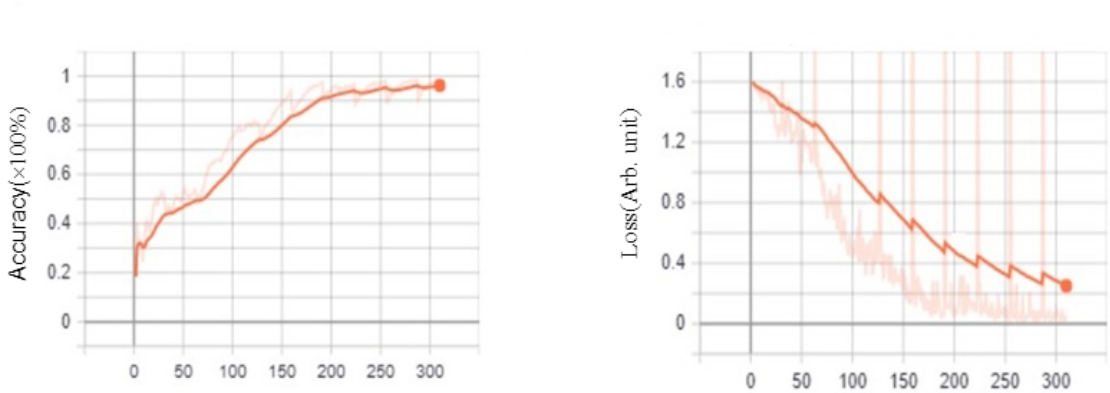


Figure 6.2: Bar diagram of training accuracy of CNN+LSTM, CNN+GRU and 3D CNN models



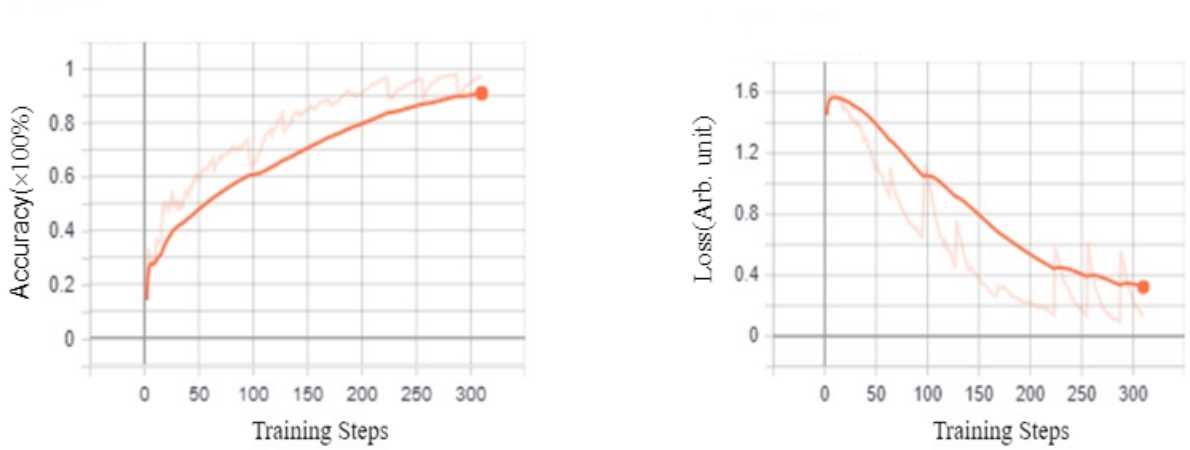
(a) Training accuracy curve (b) Training loss curve

Figure 6.3: Training Accuracy and loss curves for CNN+LSTM model



(a) Training accuracy curve (b) Training loss curve

Figure 6.4: Accuracy and loss curves for CNN+GRU model



(a) Training accuracy curve (b) Training loss curve

Figure 6.5: Accuracy and loss curves for 3D CNN model

Confusion matrices for five standard views of spatiotemporal fetal cardiac classifier are shown in Figures 6.6,Figure 6.7 and Figure 6.8. They were used to describe a classifier's performance on test samples with known labels. It addressed the class imbalance problem.

The rows of confusion matrix represented the true labels and columns represented the predicted labels. The trained model did predictions on the test data. The confusion matrix parameters were calculated based on these predictions as true positives, true negatives, false positives, and false negatives. CNN+LSTM, CNN+GRU and 3D CNN spatiotemporal models predicted on unseen data with an accuracy of 92.63%, 94.99% and 82.69%.

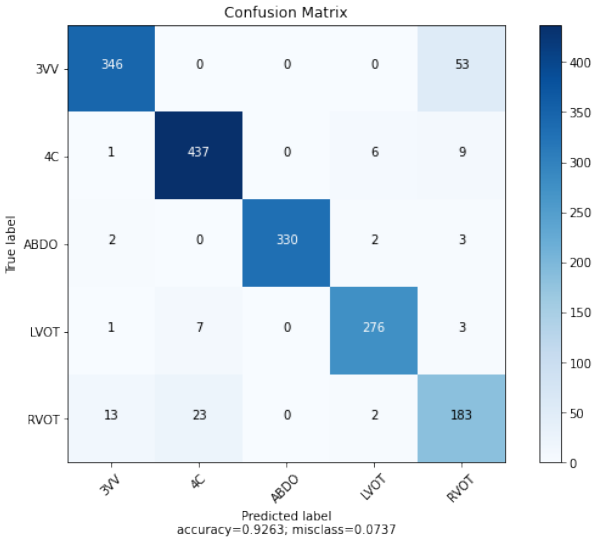


Figure 6.6: CNN+LSTM model confusion matrix

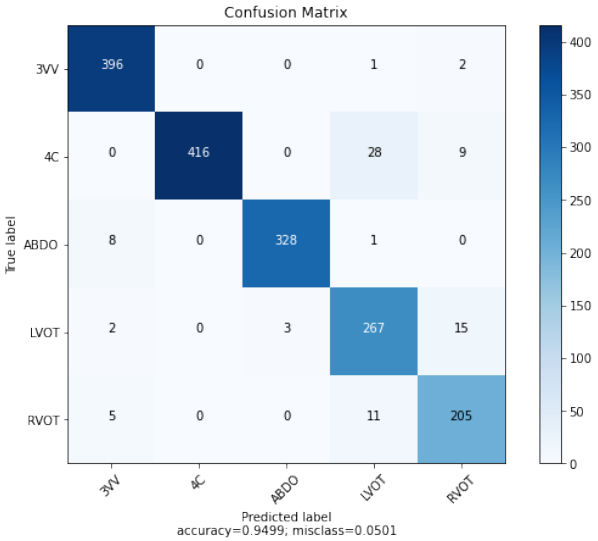


Figure 6.7: CNN+GRU model confusion matrix

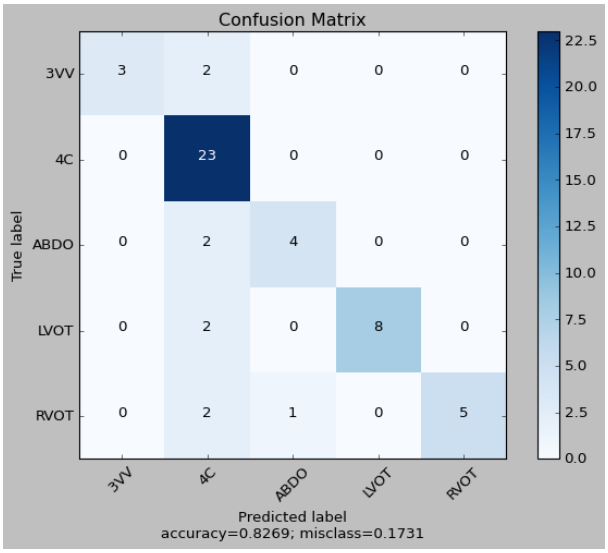


Figure 6.8: 3D CNN model confusion matrix

The calculation and distribution of evaluation matrix for five standard planes of spatiotemporal fetal cardiac images with their corresponding values are shown in table 6.1, table 6.2, table 6.3 and its bar diagram are showon in Figure 6.9, Figure 6.10, Figure 6.11.

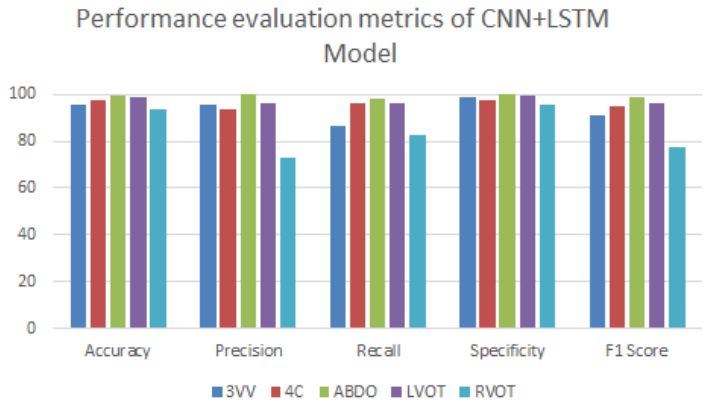


Figure 6.9: Bar chart for performance evaluation metrics of a CNN+LSTM model for five Views of fetal heart

Table 6.1: Performance evaluation metrics of a CNN+LSTM model for five Views of fetal heart

	Labels	True Positives	True Negatives	False Positives	False Negatives	Accuracy	Precision	Recall	Specificity	F1 Score
0	3VV	346	1281	17	53	0.958751	0.953168	0.867168	0.986903	0.908136
1	4C	437	1214	30	16	0.972893	0.935760	0.964680	0.975884	0.950000
2	ABDO	330	1360	0	7	0.995875	1.000000	0.979228	1.000000	0.989505
3	LVOT	276	1400	10	11	0.987625	0.965035	0.961672	0.992908	0.963351
4	RVOT	183	1408	68	38	0.937537	0.729084	0.828054	0.953930	0.775424

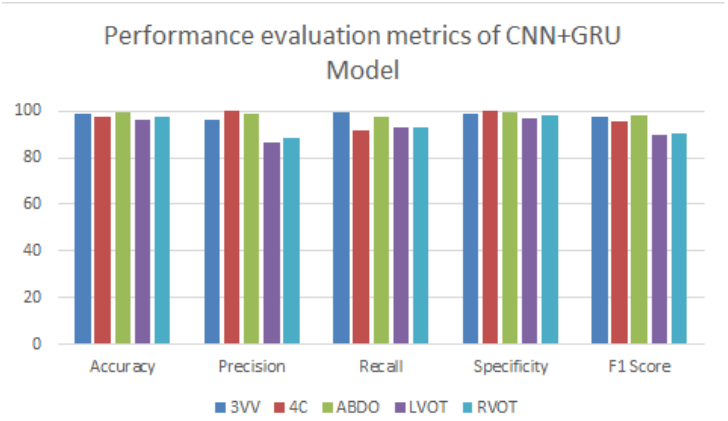


Figure 6.10: Bar chart for performance evaluation metrics of a CNN+GRU model for five Views of fetal heart

Table 6.2: Performance evaluation metrics of CNN+GRU model for five Views of fetal heart

	Labels	True Positives	True Negatives	False Positives	False Negatives	Accuracy	Precision	Recall	Specificity	F1 Score
0	3VV	396	1283	15	3	0.989393	0.963504	0.992481	0.988444	0.977778
1	4C	416	1244	0	37	0.978197	1.000000	0.918322	1.000000	0.957422
2	ABDO	328	1357	3	9	0.992929	0.990937	0.973294	0.997794	0.982036
3	LVOT	267	1369	41	20	0.964054	0.866883	0.930314	0.970922	0.897479
4	RVOT	205	1450	26	16	0.975250	0.887446	0.927602	0.982385	0.907080

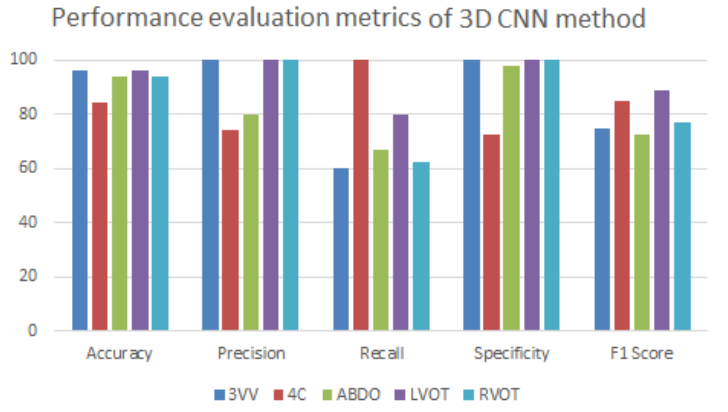


Figure 6.11: Bar chart for performance evaluation metrics of a CNN+GRU model for five Views of fetal heart

Table 6.3: Performance evaluation metrics of 3D CNN model for five Views of fetal heart

	Labels	True Positives	True Negatives	False Positives	False Negatives	Accuracy	Precision	Recall	Specificity	F1 Score
0	3VV	3	47	0	2	0.961538	1.000000	0.600000	1.000000	0.750000
1	4C	23	21	8	0	0.846154	0.741935	1.000000	0.724138	0.851852
2	ABDO	4	45	1	2	0.942308	0.800000	0.666667	0.978261	0.727273
3	LVOT	8	42	0	2	0.961538	1.000000	0.800000	1.000000	0.888889
4	RVOT	5	44	0	3	0.942308	1.000000	0.625000	1.000000	0.769231

Finally, the ROC AUC curves for the five standard spatiotemporal classification views were plotted, with the X-axis representing the false positive rate and the Y-axis representing the true positive rate, as shown in Figure 6.12, Figure 6.13 and Figure 6.14. The areas under the curve for the given CNN+LSTM, CNN+GRU and 3D CNN models were plotted as shown in above Figures . These ROC AUC values represent the probability that the model will be able to differentiate between positive and negative classes.

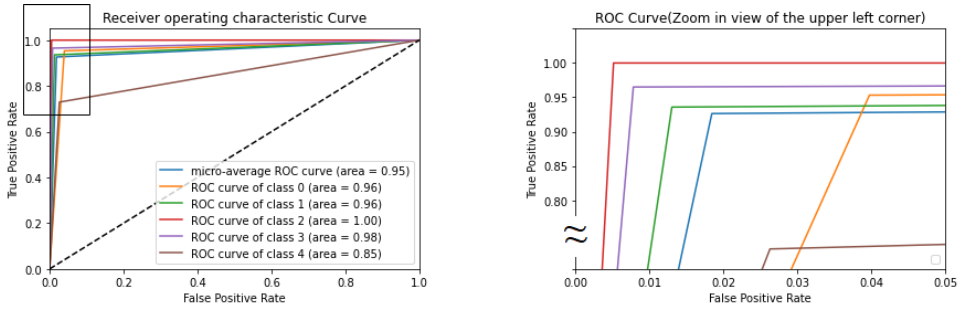


Figure 6.12: CNN+LSTM ROC (Receiver operating characteristics) curve for five standard spatiotemporal views of the fetal heart

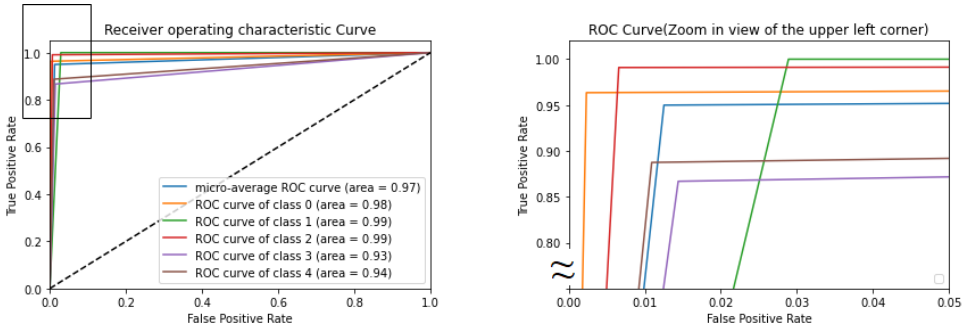


Figure 6.13: CNN+GRU ROC (Receiver operating characteristics) curve for five standard spatiotemporal views of the fetal heart

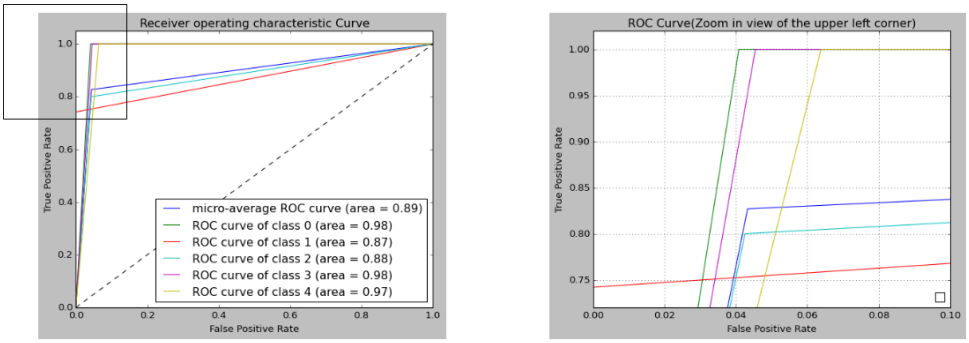


Figure 6.14: 3D CNN ROC (Receiver operating characteristics) curve for five standard spatiotemporal views of the fetal heart

6.1.1 Comparison of 3D CNN, CNN+GRU, and CNN+LSTM Models for five standard view of fetal heart

In this section, comparisons are made between implemented 3D CNN, CNN+GRU, and CNN+LSTM models. The table shows 6.4 performance of the three models. Specific bar diagram is shown in Figure 6.15

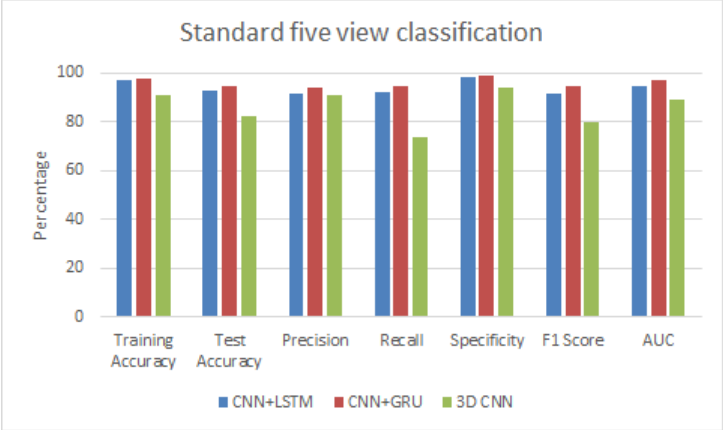


Figure 6.15: Bar chart of performance evaluation of the three models

Table 6.4: Comparison of 3D CNN, CNN+GRU, and CNN+LSTM models

Labels	Training Accuracy	Test Accuracy	Precision	Recall	Specificity	F1 Score	AUC
CNN+LSTM	97.39	92.63	91.662	92.018	98.192	91.726	95
CNN+GRU	98.02	94.99	94.174	94.84	98.79	94.436	97
3D CNN	91.35	82.69	90.838	73.834	94.048	79.746	89

For five standard views spatiotemporal classification, the CNN+LSTM, CNN+GRU, and 3D CNN models achieved training accuracy of 97.39%, 98.02%, and 91.35%, respectively, and

testing accuracy of 92.63%, 94.99%, and 82.69%. CNN+GRU model has higher value of recall (94.174%) and specificity (98.79%). As we know, high value of recall gives low false negatives and high value of specificity gives high true negatives. Precision and F1 score of CNN+GRU model is also higher compared to other two models.

The CNN+GRU model performed better than the other two models as its testing accuracy on test dataset is better than other two models. CNN proved to be good at spatial feature extraction while RNN(LSTM and GRU) was efficient in temporal feature extraction. The training performance of LSTM and GRU were comparable but the testing performance of GRU was higher than LSTM though it had fewer gates than LSTM. Furthermore, while 3D CNN's training performance was comparable to the other models, its testing accuracy was significantly lower than the other two models. The probable reason for this lower performance was lower datasets.

Hence, CNN+GRU performed better than the other two spatiotemporal models.

6.1.2 Comparison of CNN, CNN+GRU, and CNN+LSTM models for five standard view of fetal heart

In this section, the results of CNN, CNN+LSTM and CNN+GRU models for classification of five standard view of fetal heart is presented (Table 6.5). Its bar diagram is shown in Figure 6.16

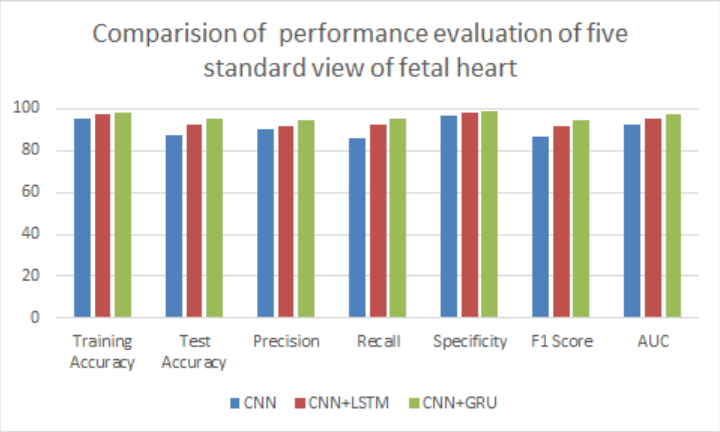


Figure 6.16: Bar chart of performance evaluation of the three models

Table 6.5: Comparison of CNN, CNN+GRU, and CNN+LSTM models

Labels	Training Accuracy	Test Accuracy	Precision	Recall	Specificity	F1 Score	AUC
CNN	95	87.57	89.97608	85.5603	96.84896	86.36036	92
CNN+LSTM	97.39	92.63	91.662	92.018	98.192	91.726	95
CNN+GRU	98.02	94.99	94.174	94.84	98.79	94.436	97

The experimental outcomes of CNN, CNN+LSTM and CNN+GRU is shown in Table 6.5. The Overall performance of CNN+GRU was shown to be better than the other two models with a test accuracy of 94.99%. The performance of CNN+LSTM model was less than CNN+GRU but it supersheded the CNN model. Furthermore, when the performance of CNN was compared with CNN+LSTM and CNN+GRU, the performance of CNN+LSTM and CNN+GRU came out to be much better than the CNN model. The resason for this better result being that the CNN+LSTM and CNN+GRU can deal with sequential data and resolve long term dependency problem along with spatial feature detection.

6.2 Classification and performance evaluation of normal fetal heart vs. Congenital heart disease(CHD) lesions

The three systems were able to classify 352 fetal cardiac cine loops into three categories namely normal-4C, HLHS, and TA with an accuracy of 94.61%, 91.99% and 86.54% respectively.

- Total Samples of videos : 352
- Number of epochs : 10
- Learning Rate : 0.001
- Optimizer : Adam Optimizer
- Training Accuracy :
    1. CNN+LSTM : 98.60%
    2. CNN+GRU : 98.63%
    3. 3DCNN : 97.03%

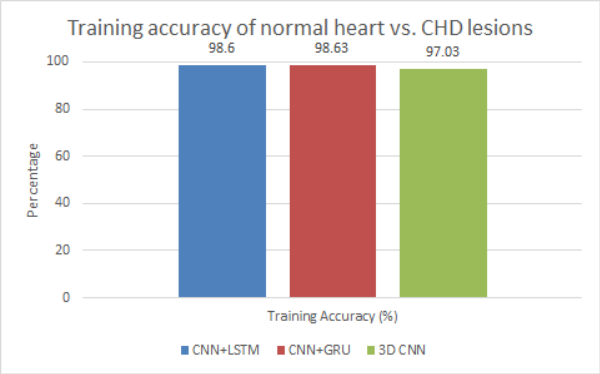
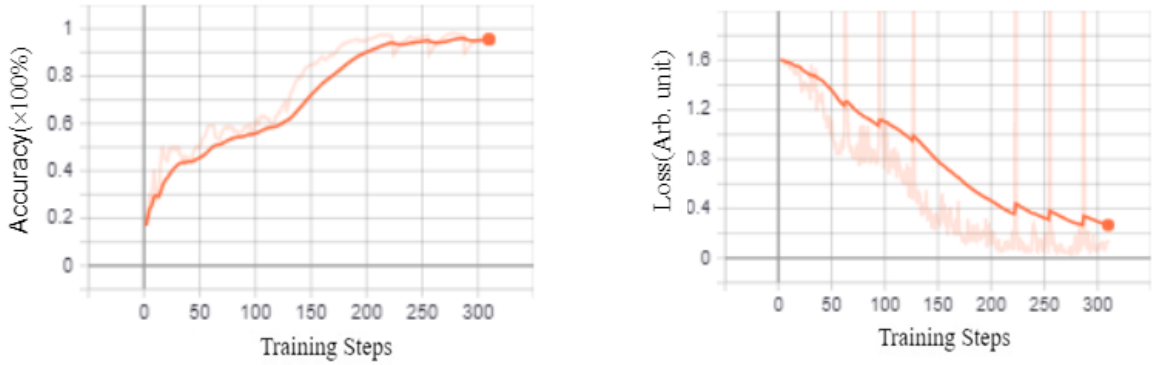


Figure 6.17: Bar chart for the training accuracy of normal fetal heart vs. CHD lesions

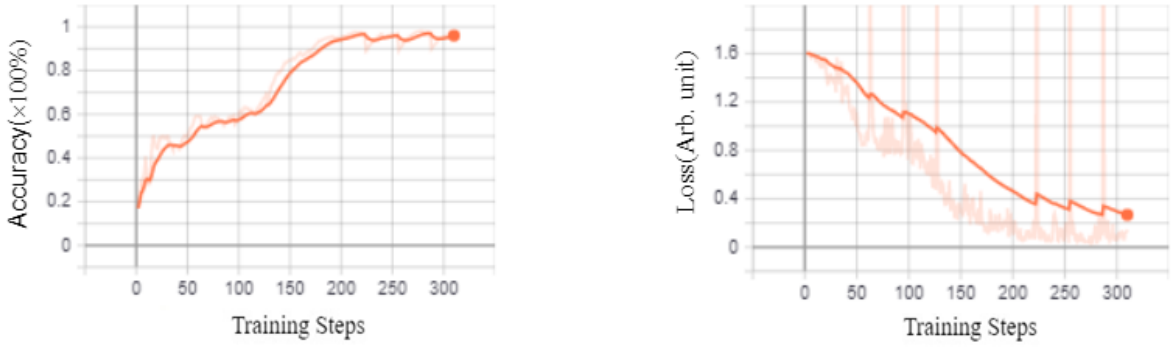
CNN+LSTM, CNN+GRU, and 3D CNN had training accuracy of 98.60%, 98.63%, and 97.03%,

respectively, with crossentropy losses of 0.08566, 0.07725, and 0.7429. Figure 6.18, Figure 6.19 and Figure 6.20 represents a plot of accuracy vs epochs(training steps) for the respective models and training accuracy summary is shown in Figure 6.17. The graphs shows the model's performances have been improving over time, which means that the models are learning.



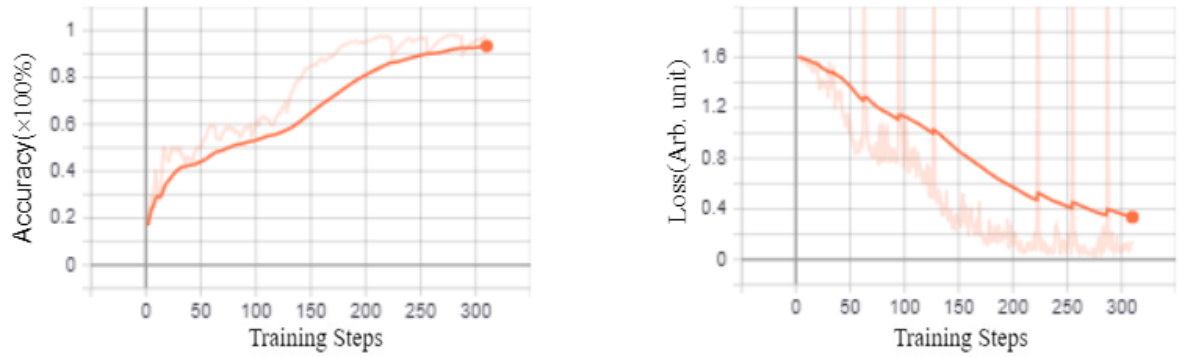
(a) Training accuracy curve (b) Training loss curve

Figure 6.18: Accuracy and loss curves for CNN+LSTM model



(a) Training accuracy curve (b) Training loss curve

Figure 6.19: Accuracy and loss curves for CNN+GRU model



(a) Training accuracy curve (b) Training loss curve

Figure 6.20: Accuracy and loss curves for 3D CNN model

Confusion matrices for three views normal heart vs. CHD lesions of spatiotemporal fetal cardiac classifier are shown in Figures 6.21, Figure 6.22 and Figure 6.23. They are used to describe a classifier's performance on test samples with known labels. Class imbalance problem was also addressed by this classifier.

The rows of confusion matrix represented the true labels and columns represented the predicted labels. The trained model did predictions on the test data. The confusion matrix parameters were calculated based on these predictions as true positives, false negatives, true negatives, and false positives. CNN+LSTM, CNN+GRU and 3D CNN spatiotemporal models made predictions on the test data with an accuracy of 94.61%, 91.99% and 86.54%.

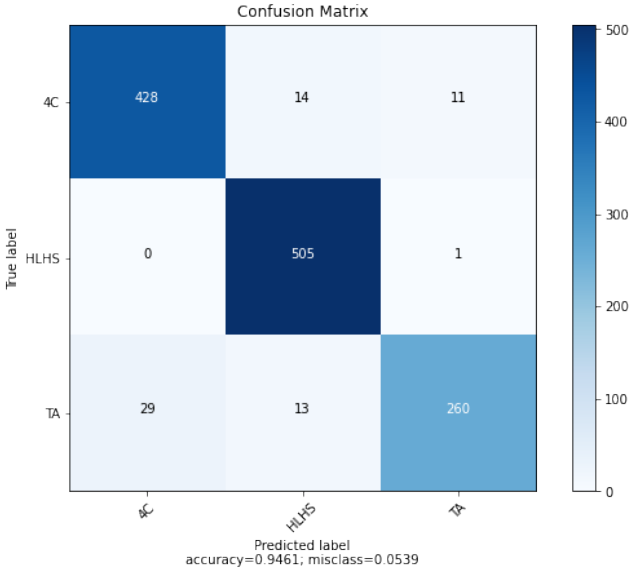


Figure 6.21: CNN+LSTM model confusion matrix

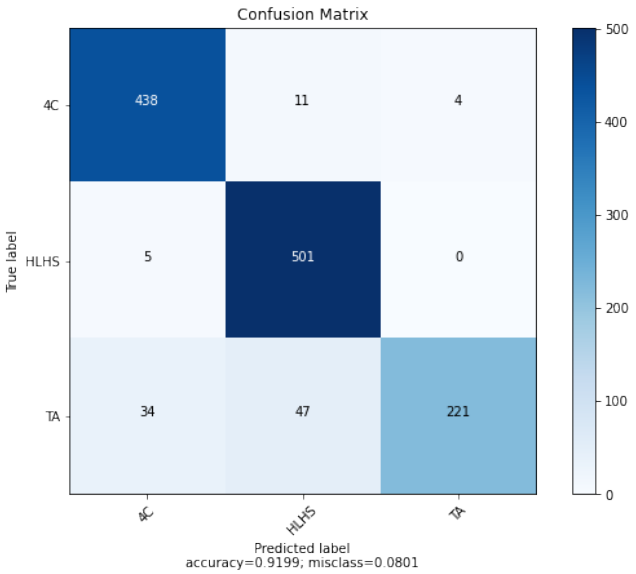


Figure 6.22: CNN+GRU model confusion matrix

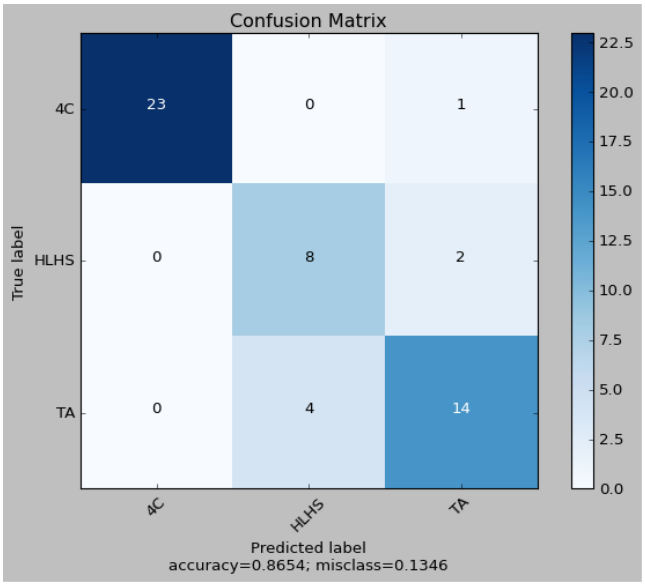


Figure 6.23: 3D CNN model confusion matrix

The calculation and distribution of performance evaluation for normal heart vs CHD lesions with their corresponding values are shown in table 6.1, table 6.2, table 6.3 and its bar diagrams are shown in Figure 6.24, Figure 6.25, Figure 6.26.

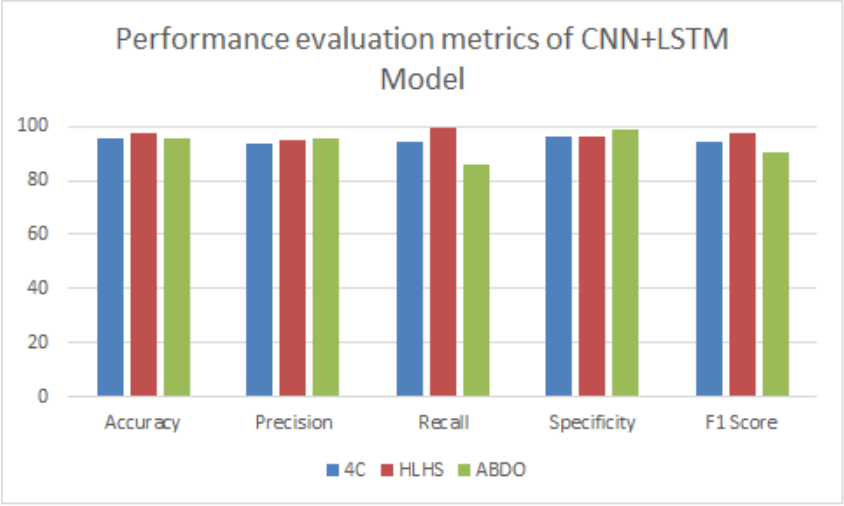


Figure 6.24: Bar chart of performance evaluation metrics of a CNN+LSTM model for normal fetal heart vs. CHD lesions

Table 6.6: Performance evaluation metrics of a CNN+LSTM model for normal fetal heart vs. CHD lesions

	Labels	True Positives	True Negatives	False Positives	False Negatives	Accuracy	Precision	Recall	Specificity	F1 Score
0	4C	428	779	29	25	0.957177	0.936543	0.944812	0.964109	0.940659
1	HLHS	505	728	27	1	0.977795	0.949248	0.998024	0.964238	0.973025
2	TA	260	947	12	42	0.957177	0.955882	0.860927	0.987487	0.905923

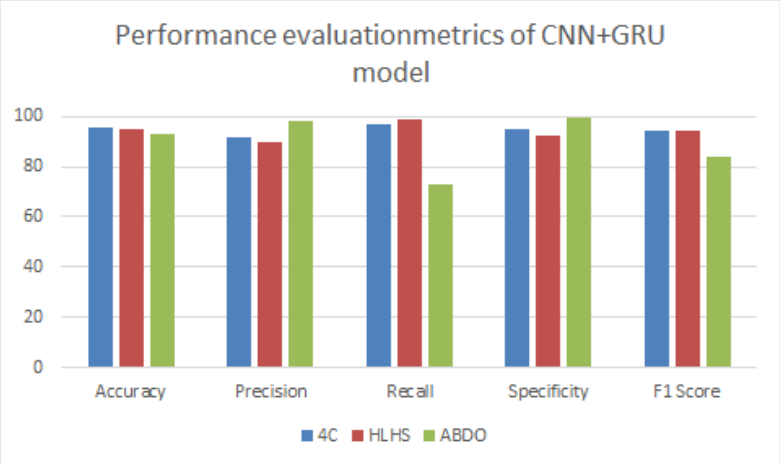


Figure 6.25: Bar chart of performance evaluation metrics of a CNN+GRU model for normal fetal heart vs. CHD lesions

Table 6.7: Performance evaluation metrics of CNN+GRU model for normal fetal heart vs. CHD lesions

	Labels	True Positives	True Negatives	False Positives	False Negatives	Accuracy	Precision	Recall	Specificity	F1 Score
0	4C	438	769	39	15	0.957177	0.918239	0.966887	0.951733	0.941935
1	HLHS	501	697	58	5	0.950040	0.896243	0.990119	0.923179	0.940845
2	TA	221	955	4	81	0.932593	0.982222	0.731788	0.995829	0.838710

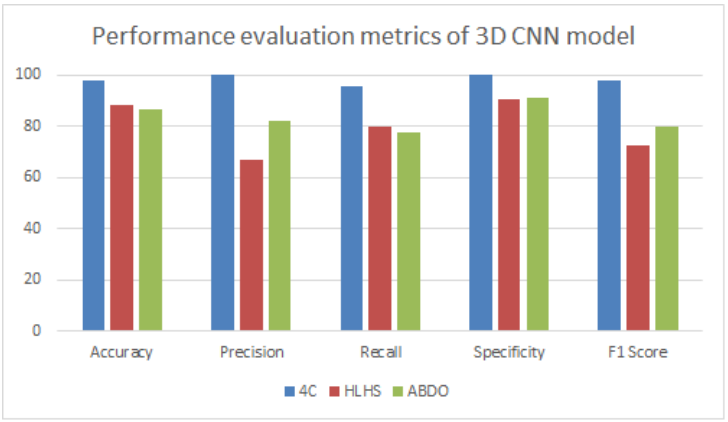


Figure 6.26: Bar chart of performance evaluation metrics of a 3D CNN model for normal fetal heart vs. CHD lesions

Table 6.8: Performance evaluation metrics of 3D CNN model for normal fetal heart vs. CHD lesions

	Labels	True Positives	True Negatives	False Positives	False Negatives	Accuracy	Precision	Recall	Specificity	F1 Score
0	4C	23	28	0	1	0.980769	1.000000	0.958333	1.000000	0.978723
1	HLHS	8	38	4	2	0.884615	0.666667	0.800000	0.904762	0.727273
2	TA	14	31	3	4	0.865385	0.823529	0.777778	0.911765	0.800000

Finally, the ROC AUC curves for normal heart vs. CHD lesions spatiotemporal classification views were plotted, with the X-axis representing the false positive rate and the Y-axis representing the true positive rate, as shown in Figure 6.27, Figure 6.28 and Figure 6.29. The areas under the curve for the given three models were plotted as shown in above figures . These ROC AUC values represent the probability that the model will be able to differentiate between positive and negative classes.

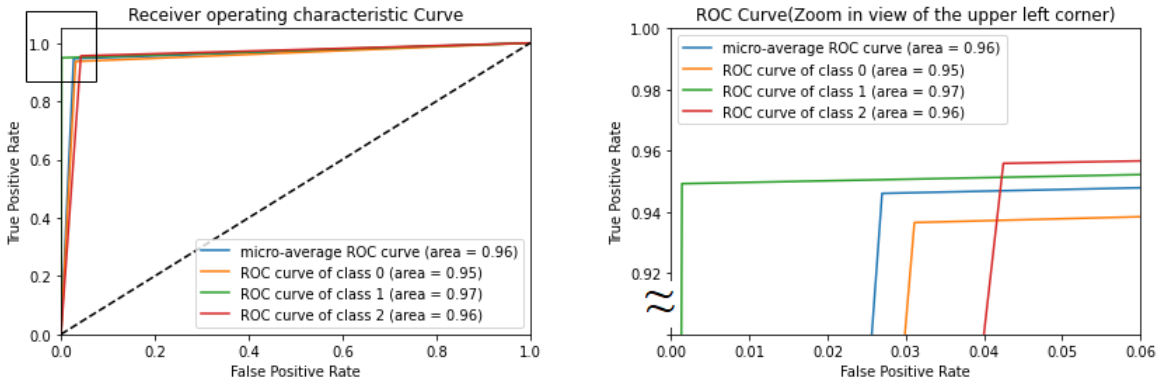


Figure 6.27: CNN+LSTM ROC (Receiver operating characteristics) curve for normal fetal heart vs. CHD lesions

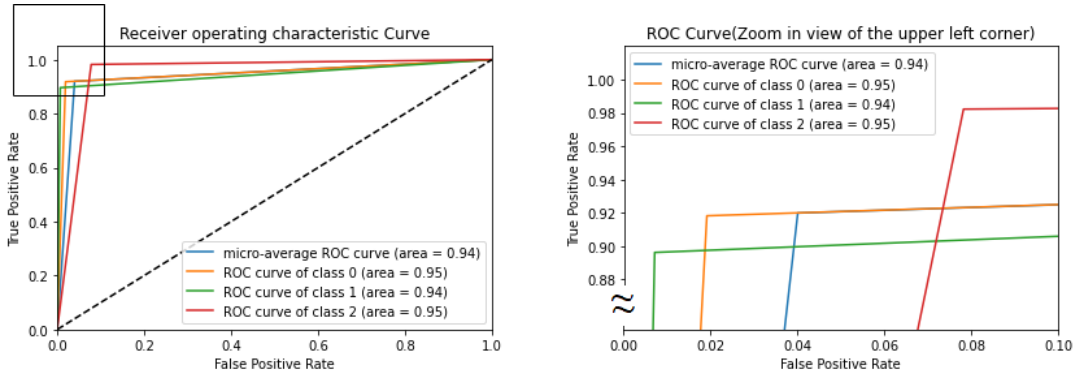


Figure 6.28: CNN+GRU ROC (Receiver operating characteristics) curve for normal fetal heart vs. CHD lesions

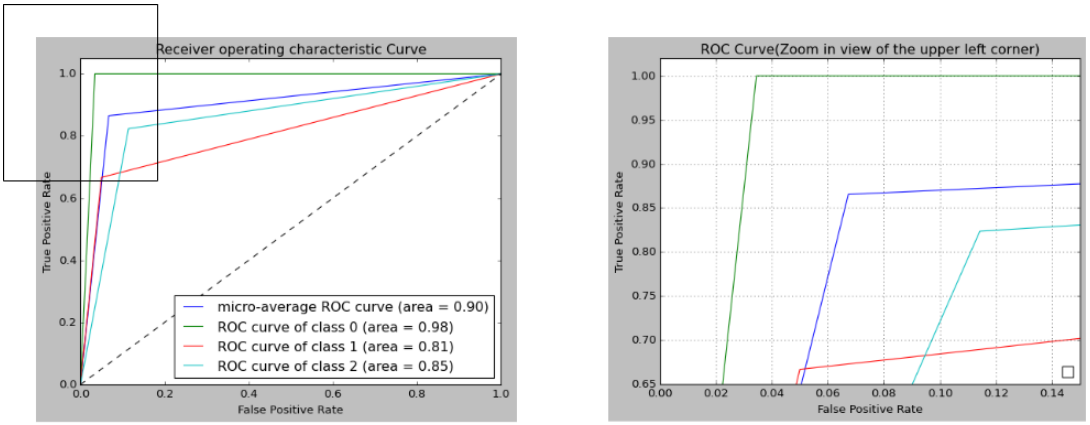


Figure 6.29: 3D CNN ROC (Receiver operating characteristics) curve for normal fetal heart vs. CHD lesions

6.2.1 Comparison of 3D CNN, CNN+GRU, and CNN+LSTM Models for the classification of normal fetal heart vs. CHD lesions

In the section mentioned below, comparisons are made between implemented 3D CNN, CNN+GRU, and CNN+LSTM models. Bar diagram shown in Figure 6.30 shows the graphical representation of the data shown in table 6.9.

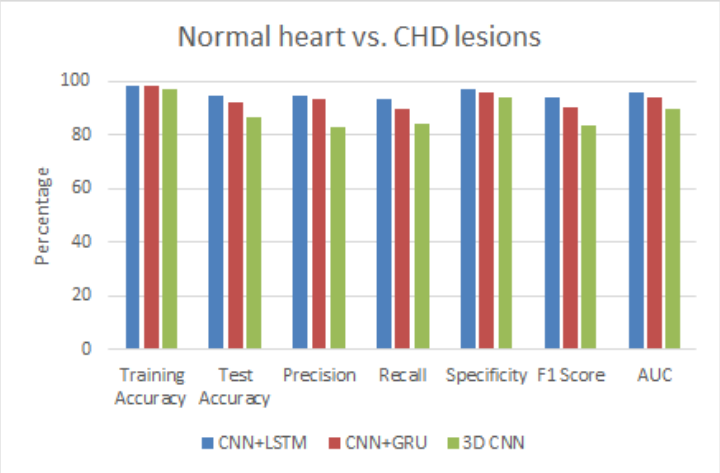


Figure 6.30: Bar chart of performance evaluation of the three models

Table 6.9: Comparison of 3D CNN, CNN+GRU, and CNN+LSTM models

Models	Training Accuracy	Test Accuracy	Precision	Recall	Specificity	F1 Score	AUC
CNN+LSTM	98.6	94.61	94.72	93.456	97.193	93.986	96
CNN+GRU	98.63	91.99	93.22	89.626	95.69	90.713	94
3D CNN	97.03	86.54	83.01	84.536	93.886	83.533	90

For normal vs. CHD lesions spatiotemporal classification, the CNN+LSTM, CNN+GRU, and 3D CNN models gained training accuracy of 98.60%, 98.63%, and 97.30% and testing accuracy of 94.61%, 91.99%, and 86.54%, respectively. Table 6.9 shows that the overall performace of CNN+LSTM model for normal vs. CHD lesions is better than CNN+GRU and 3D CNN.

The CNN+LSTM model performed better than the other two models as its testing accuracy is higher. CNN proved to be good at spatial feature extraction while RNN(LSTM and GRU) was efficient in temporal feature extraction. The training performance of LSTM and GRU were comparable but the testing performance of GRU was lower than LSTM as GRU had fewer gates than LSTM. Furthermore, while 3D CNN's training performance was comparable to the other models, its testing accuracy was significantly lower than the other two models. The probable reason for this lower performance was lower datasets.

Hence, CNN+LSTM outperformed than the other two spatiotemporal models.

6.2.2 Comparison of CNN, CNN+GRU, and CNN+LSTM Models for the classification of normal fetal heart vs. CHD lesions

The outcomes of the CNN, CNN+LSTM, and CNN+GRU models for classification of normal fetal heart vs. CHD lesions are reported in this section (Table 6.10). Figure 6.31 depicts its bar diagram.

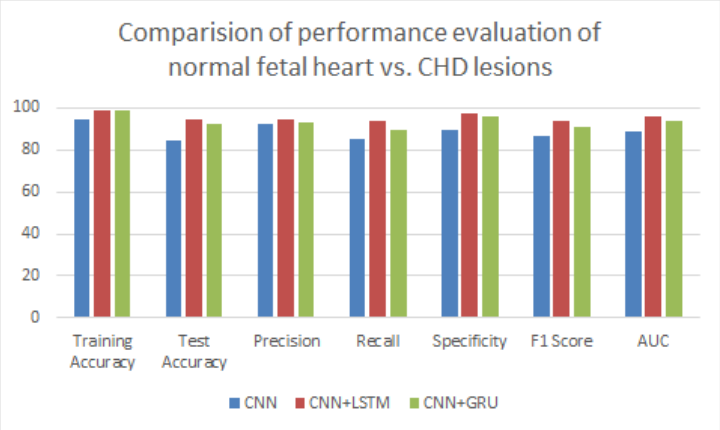


Figure 6.31: Bar chart of performance evaluation of the three models

Table 6.10: Comparison of CNN, CNN+GRU, and CNN+LSTM models

Models	Training Accuracy	Test Accuracy	Precision	Recall	Specificity	F1 Score	AUC
CNN	94.29	84.68	92.304	84.802	89.569	86.256	89
CNN+LSTM	98.6	94.61	94.72	93.456	97.193	93.986	96
CNN+GRU	98.63	91.99	93.22	89.626	95.69	90.713	94

The experimental outcomes of CNN, CNN+LSTM and CNN+GRU is shown in Table 6.10. With a test accuracy of 94.61%, CNN+LSTM outperformed the other two models. The performance of the CNN+GRU model was lower than that of the CNN+LSTM model, but it outperformed the CNN model. Furthermore, when CNN's performance was compared to CNN+LSTM and CNN+GRU, CNN+LSTM and CNN+GRU performed significantly better than the CNN model. The reason for this improved performance being that CNN+LSTM and CNN+GRU can deal with sequential data and handle long-term dependency problems while detecting spatial features.

CHAPTER 7

CONCLUSION

Hence, the proposed spatiotemporal deep learning models i.e. CNN+LSTM, CNN+GRU and 3D CNN were able to classify the fetal cardiac images into the five standard views with an accuracy of 92.63%, 94.99% and 82.69% respectively. CNN+GRU outperforms the other two models in terms of accuracy when compared to experimental findings.

Similarly, these algorithms were able to identify and diagnose normal heart vs. Tricuspid atresia and HLHS with an accuracy of 94.61%, 91.99% and 86.54% respectively. When compared to the experimental findings, CNN+LSTM surpasses the other two models.

CNN+LSTM and CNN+GRU performed significantly better than the CNN model. The reason for this improved performance being that CNN+LSTM and CNN+GRU can deal with sequential data and handle long-term dependency problems while detecting spatial features.

REFERENCES

[1] C. Ferencz, J. D. Rubin, R. J. Mccarter, J. I. Brenner, C. A. Neill, L. W. Perry, S. I. Hepner, and J. W. Downing, "Congenital heart disease: Prevalence at livebirth: The baltimore-washington infant study," *American journal of epidemiology*, vol. 121, no. 1, pp. 31–36, 1985.

[2] J. P. Crane, M. L. LeFevre, R. C. Winborn, J. K. Evans, B. G. Ewigman, R. P. Bain, F. D. Frigoletto, D. McNellis, R. S. Group, *et al.*, "A randomized trial of prenatal ultrasonographic screening: Impact on the detection, management, and outcome of anomalous fetuses," *American journal of obstetrics and gynecology*, vol. 171, no. 2, pp. 392–399, 1994.

[3] M. Abu-Harb, E. Hey, and C. Wren, "Death in infancy from unrecognised congenital heart disease.," *Archives of disease in childhood*, vol. 71, no. 1, pp. 3–7, 1994.

[4] L. J. Salomon, Z Alfirevic, V Berghella, C Bilardo, E Hernandez-Andrade, S. Johnsen, K Kalache, K.-Y. Leung, G Malinger, H Munoz, *et al.*, "Practice guidelines for performance of the routine mid-trimester fetal ultrasound scan," *Ultrasound in Obstetrics & Gynecology*, vol. 37, no. 1, pp. 116–126, 2011.

[5] D. Paladini and P. Volpe, *Ultrasound of congenital fetal anomalies: differential diagnosis and prognostic indicators*. CRC press, 2018.

[6] D. Brezeale and D. J. Cook, "Automatic video classification: A survey of the literature," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 38, no. 3, pp. 416–430, 2008.

[7] Z. Wu, X. Wang, Y.-G. Jiang, H. Ye, and X. Xue, "Modeling spatial-temporal clues in a hybrid deep learning framework for video classification," in *Proceedings of the 23rd ACM international conference on Multimedia*, 2015, pp. 461–470.

[8] W.-H. Lin and A. Hauptmann, "News video classification using svm-based multimodal classifiers and combination strategies," in *Proceedings of the tenth ACM international conference on Multimedia*, 2002, pp. 323–326.

[9] R. S. Jasinschi, N. Dimitrova, T. McGee, L. Agnihotri, J. Zimmerman, and D. Li, "Integrated multimedia processing for topic segmentation and classification," in *Proceedings 2001 International Conference on Image Processing (Cat. No. 01CH37205)*, IEEE, vol. 3, 2001, pp. 366–369.

[10] G. E. Hinton, "Deep belief networks," *Scholarpedia*, vol. 4, no. 5, p. 5947, 2009.

[11] S. Zha, F. Luisier, W. Andrews, N. Srivastava, and R. Salakhutdinov, "Exploiting image-trained cnn architectures for unconstrained video classification," *arXiv preprint arXiv:1503.04144*, 2015.

[12] S. Ji, W. Xu, M. Yang, and K. Yu, "3d convolutional neural networks for human action recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 1, pp. 221–231, 2012.

[13] S. Ji, C. Zhang, A. Xu, Y. Shi, and Y. Duan, "3d convolutional neural networks for crop classification with multi-temporal remote sensing images," *Remote Sensing*, vol. 10, no. 1, p. 75, 2018.

[14] S. Jasechko, Z. D. Sharp, J. J. Gibson, S. J. Birks, Y. Yi, and P. J. Fawcett, "Terrestrial water fluxes dominated by transpiration," *Nature*, vol. 496, no. 7445, pp. 347–350, 2013.

[15] W. Ye, J. Cheng, F. Yang, and Y. Xu, "Two-stream convolutional network for improving activity recognition using convolutional long short-term memory networks," *IEEE Access*, vol. 7, pp. 67 772–67 780, 2019.

[16] T. Akilan, Q. J. Wu, A. Safaei, J. Huo, and Y. Yang, "A 3d cnn-lstm-based image-to-image foreground segmentation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 3, pp. 959–971, 2019.

[17] C. Feichtenhofer, H. Fan, J. Malik, and K. He, "Slowfast networks for video recognition," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 6202–6211.

[18] K. A. L. Hernandez, T. Rienmüller, D. Baumgartner, and C. Baumgartner, "Deep learning in spatiotemporal cardiac imaging: A review of methodologies and clinical usability," *Computers in Biology and Medicine*, p. 104 200, 2020.

[19] R. Arnaout, L. Curran, E. Chinn, Y. Zhao, and A. Moon-Grady, "Deep-learning models improve on community-level diagnosis for common congenital heart disease lesions," *arXiv preprint arXiv:1809.06993*, 2018.

[20] B. Kong, Y. Zhan, M. Shin, T. Denny, and S. Zhang, "Recognizing end-diastole and end-systole frames via deep temporal regression network," in *International conference on medical image computing and computer-assisted intervention*, Springer, 2016, pp. 264–272.

[21] W. Xue, G. Brahm, S. Pandey, S. Leung, and S. Li, "Full left ventricle quantification via deep multitask relationships learning," *Medical image analysis*, vol. 43, pp. 54–65, 2018.

[22] A. Debus and E. Ferrante, "Left ventricle quantification through spatio-temporal cnns," in *International Workshop on Statistical Atlases and Computational Models of the Heart*, Springer, 2018, pp. 466–475.

[23] K. Liu, G. Kang, N. Zhang, and B. Hou, "Breast cancer classification based on fully-connected layer first convolutional neural networks," *IEEE Access*, vol. 6, pp. 23 722–23 732, 2018.

[24] Y Zhao, S Edington, J Fleenor, E Sinkovskaya, L Porche, and A Abuhamad, "Fetal cardiac axis in tetralogy of fallot: Associations with prenatal findings, genetic anomalies and postnatal outcome," *Ultrasound in Obstetrics & Gynecology*, vol. 50, no. 1, pp. 58–62, 2017.

[25] R. Memisevic, C. Zach, M. Pollefeys, and G. E. Hinton, "Gated softmax classification," *Advances in neural information processing systems*, vol. 23, pp. 1603–1611, 2010.

[26] D. K. McClish, "Analyzing a portion of the roc curve," *Medical Decision Making*, vol. 9, no. 3, pp. 190–195, 1989.

[27] D. Chicco and G. Jurman, "The advantages of the matthews correlation coefficient (mcc) over f1 score and accuracy in binary classification evaluation," *BMC genomics*, vol. 21, no. 1, pp. 1–13, 2020.

APPENDIX

1

ORIGINALITY REPORT

17%	8%	12%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Paladini, . "Congenital heart disease", Ultrasound of Congenital Fetal Anomalies Differential Diagnosis and Prognostic Indicators, 2007. Publication	1%
2	"Frontiers of Multimedia Research", Association for Computing Machinery (ACM), 2017 Publication	1%
3	www.fhwa.dot.gov Internet Source	1%
4	www.mallinckrodt.com Internet Source	1%
5	Submitted to University College London Student Paper	1%
6	www.coursehero.com Internet Source	<1%
7	Paladini, . "Congenital heart disease", Ultrasound of Congenital Fetal Anomalies, 2014.	<1%

Publication

8	Shi-Joon Yoo, Young Ho Lee, Kyoung Sik Cho, Dae-Young Kim. "Sequential segmental approach to fetal congenital heart disease", Cardiology in the Young, 2008 Publication	<1 %
9	Ximing Huai, Siriaraya Panote, Dongeun Choi, Noriaki Kuwahara. "Chapter 36 Heart Sound Recognition Technology Based on Deep Learning", Springer Science and Business Media LLC, 2020 Publication	<1 %
10	www.umiacs.umd.edu Internet Source	<1 %
11	Submitted to Mansoura University Student Paper	<1 %
12	doctorpenguin.com Internet Source	<1 %
13	Submitted to University of Stirling Student Paper	<1 %
14	Georgios Petmezas, Kostas Haris, Leandros Stefanopoulos, Vassilis Kilintzis et al. "Automated Atrial Fibrillation Detection using a Hybrid CNN-LSTM Network on Imbalanced ECG Datasets", Biomedical Signal Processing and Control, 2021 Publication	<1 %

15	Submitted to University of Lincoln Student Paper	<1 %
16	www.hindawi.com Internet Source	<1 %
17	Karen Andrea Lara Hernandez, Theresa Rienmüller, Daniela Baumgartner, Christian Baumgartner. "Deep Learning in Spatiotemporal Cardiac Imaging: A Review of Methodologies and Clinical Usability", Computers in Biology and Medicine, 2020 Publication	<1 %
18	Submitted to Cranfield University Student Paper	<1 %
19	Submitted to Indian Institute of Technology, Madras Student Paper	<1 %
20	Jiaying Li, Weidong Wang, Zheng Han. "A variable weight combination model for prediction on landslide displacement using AR model, LSTM model, and SVM model: a case study of the Xinming landslide in China", Environmental Earth Sciences, 2021 Publication	<1 %
21	Submitted to King's College Student Paper	<1 %

22	Huaijun Wang, Jing Zhao, Junhuai Li, Ling Tian, Pengjia Tu, Ting Cao, Yang An, Kan Wang, Shancang Li. "Wearable Sensor-Based Human Activity Recognition Using Hybrid Deep Learning Techniques", Security and Communication Networks, 2020 Publication	<1 %
23	baadalsg.inflibnet.ac.in Internet Source	<1 %
24	"Cybersecurity and Secure Information Systems", Springer Science and Business Media LLC, 2019 Publication	<1 %
25	"Proceedings of International Conference on Sustainable Expert Systems", Springer Science and Business Media LLC, 2021 Publication	<1 %
26	Submitted to Swinburne University of Technology Student Paper	<1 %
27	Keke Geng, Guodong Yin. "Using Deep Learning in Infrared Images to Enable Human Gesture Recognition for Autonomous Vehicles", IEEE Access, 2020 Publication	<1 %
28	hdl.handle.net Internet Source	<1 %

29	jaimeargudo.com Internet Source	<1 %
30	"Information Systems Design and Intelligent Applications", Springer Science and Business Media LLC, 2018 Publication	<1 %
31	Submitted to Xianjiatong-Liverpool University Student Paper	<1 %
32	socratic.org Internet Source	<1 %
33	www.guoyanbin.com Internet Source	<1 %
34	trepo.tuni.fi Internet Source	<1 %
35	"International Conference on Innovative Computing and Communications", Springer Science and Business Media LLC, 2021 Publication	<1 %
36	Submitted to The British College Student Paper	<1 %
37	Submitted to University of Essex Student Paper	<1 %
38	Submitted to University of Southampton Student Paper	<1 %
	www.mcponline.org	

39	Internet Source	<1 %
40	"Hypoplastic Left Heart Syndrome", Springer Science and Business Media LLC, 2003 Publication	<1 %
41	Submitted to Institute of Engineering, Pulchock Campus, Tribhuvan University Student Paper	<1 %
42	Submitted to King Abdulaziz University Student Paper	<1 %
43	"Data Engineering and Intelligent Computing", Springer Science and Business Media LLC, 2021 Publication	<1 %
44	orca.cf.ac.uk Internet Source	<1 %
45	Submitted to Hacettepe University Student Paper	<1 %
46	Submitted to Heriot-Watt University Student Paper	<1 %
47	Submitted to Imperial College of Science, Technology and Medicine Student Paper	<1 %
48	Kwangjo Kim, Muhamad Erza Aminanto, Harry Chandra Tanuwidjaja. "Network Intrusion	<1 %

	Detection using Deep Learning", Springer Science and Business Media LLC, 2018 Publication	
49	arxiv.org Internet Source	<1 %
50	dissertations.umi.com Internet Source	<1 %
51	docplayer.net Internet Source	<1 %
52	mafiadoc.com Internet Source	<1 %
53	Submitted to The University of Buckingham Student Paper	<1 %
54	Submitted to kingswaynz Student Paper	<1 %
55	open.library.ubc.ca Internet Source	<1 %
56	vdocuments.mx Internet Source	<1 %
57	"Databases and Information Systems", Springer Science and Business Media LLC, 2020 Publication	<1 %
58	Dinh, Howard, Jeffrey Goldman, and Michael Poon. "CMR in the assessment of pulmonary	<1 %

hypertension", Cardiovascular Magnetic Resonance Established and Emerging Applications, 2004.
Publication

59	Gaël Manson, Sid-Ahmed Berrani. "Content-Based Video Segment Reunification for TV Program Extraction", 2009 11th IEEE International Symposium on Multimedia, 2009 Publication	<1 %
60	eprints.utar.edu.my Internet Source	<1 %
61	scholarworks.uark.edu Internet Source	<1 %
62	www.csd.uoc.gr Internet Source	<1 %
63	www.karger.com Internet Source	<1 %
64	www.sonographr.com Internet Source	<1 %
65	"Gesture Recognition", Springer Science and Business Media LLC, 2017 Publication	<1 %
66	A. Chakladar. "Dangers of listening to the fetal heart at home", BMJ, 11/05/2009 Publication	<1 %
67	Submitted to Anglia Ruskin University	

	Student Paper	<1 %
68	Guang Yang, HwaMin Lee, Giyeol Lee. "A Hybrid Deep Learning Model to Forecast Particulate Matter Concentration Levels in Seoul, South Korea", Atmosphere, 2020 Publication	<1 %
69	Submitted to Instituto de Empress S.L. Student Paper	<1 %
70	Jin-Chung Shih, Shu-Chien Huang, Chia-Hui Lin, Tzu-Hung Lin et al. "Diagnosis of Transposition of the Great Arteries in the Fetus", Journal of Medical Ultrasound, 2012 Publication	<1 %
71	Md Zahangir Alom, Tarek M. Taha, Chris Yakopcic, Stefan Westberg et al. "A State-of-the-Art Survey on Deep Learning Theory and Architectures", Electronics, 2019 Publication	<1 %
72	Program: electronic library and information systems, Volume 48, Issue 3 (2014-09-16) Publication	<1 %
73	creativecommons.org Internet Source	<1 %
74	dl.mehrsys.ir Internet Source	<1 %

75	etheses.whiterose.ac.uk Internet Source	<1 %
76	ipfs.io Internet Source	<1 %
77	"Intelligent Computing Paradigm and Cutting-edge Technologies", Springer Science and Business Media LLC, 2021 Publication	<1 %
78	Abhilash Majumder. "Deep Reinforcement Learning in Unity", Springer Science and Business Media LLC, 2021 Publication	<1 %
79	Chen, Chi-Tsong. "Analog and Digital Control System Design", Oxford University Press Publication	<1 %
80	Darin Brezeale, Diane J. Cook. "Learning video preferences from video content", Proceedings of the 8th international workshop on Multimedia data mining (associated with the ACM SIGKDD 2007) - MDM '07, 2007 Publication	<1 %
81	Darin Brezeale. "", IEEE Transactions on Systems Man and Cybernetics Part C (Applications and Reviews), 5/2008 Publication	<1 %
82	Irfan Javid, Ahmed Khalaf, Rozaida Ghazali. "Enhanced Accuracy of Heart Disease	<1 %

Prediction using Machine Learning and Recurrent Neural Networks Ensemble Majority Voting Method", International Journal of Advanced Computer Science and Applications, 2020
Publication

---

83 Jeremy Tan, Anselm Au, Qingjie Meng, Sandy FinesilverSmith et al. "Chapter 24 Automated Detection of Congenital Heart Disease in Fetal Ultrasound Screening", Springer Science and Business Media LLC, 2020 <1 %
Publication

---

84 Nabila Khodeir. "Bi-GRU Urgent Classification for MOOC Discussion Forums Based on BERT", IEEE Access, 2021 <1 %
Publication

---

85 Saugat Aryal, Anjana Shriwantha Porawagama, Makulle Gedara Sajeeva Hasith, Sachini Chloe Thoradeniya et al. "Using Pre-trained Models As Feature Extractor To Classify Video Styles Used In MOOC Videos", 2018 IEEE International Conference on Information and Automation for Sustainability (ICIAfS), 2018 <1 %
Publication

---

86 Su P. Sone, Janne J. Lehtomaki, Zaheer Khan. "Wireless Traffic Usage Forecasting Using Real Enterprise Network Data: Analysis and <1 %

Methods", IEEE Open Journal of the Communications Society, 2020
Publication

87	V. Nagadevara. "Development of Hybrid Classification Methodology for Mining Skewed Data Sets ? A Case Study of Indian Customs Data", IEEE International Conference on Computer Systems and Applications 2006, 2006 Publication	<1 %
88	Xin Xu, Hubo Cai, JungHo Jeon. "Classifying Semantic Relationships of Utility-Specific Terminologies via LSTM Networks along Shortest Dependency Paths", Construction Research Congress 2020, 2020 Publication	<1 %
89	dokumen.pub Internet Source	<1 %
90	livrepository.liverpool.ac.uk Internet Source	<1 %
91	pubs.rsna.org Internet Source	<1 %
92	www.icommercecentral.com Internet Source	<1 %
93	www.ijeat.org Internet Source	<1 %

94	www.mdpi.com Internet Source	<1 %
95	www.medrxiv.org Internet Source	<1 %
96	www.rezrov.net Internet Source	<1 %
97	Sadegh Arefnezhad, Sajjad Samiee, Arno Eichberger, Matthias Frühwirth, Clemens Kaufmann, Emma Klotz. "Applying deep neural networks for multi-level classification of driver drowsiness using Vehicle-based measures", Expert Systems with Applications, 2020 Publication	<1 %
98	"Hypoplastic Left Heart Syndrome", Springer Science and Business Media LLC, 2005 Publication	<1 %
99	Sepp Hochreiter, Jürgen Schmidhuber. "Long Short-Term Memory", Neural Computation, 1997 Publication	<1 %
100	Suleiman Y. Yerima, Mohammed K. Alzaylaee, Annette Shajan, Vinod P. "Deep Learning Techniques for Android Botnet Detection", Electronics, 2021 Publication	<1 %

101	Theerawit Wilaiprasitporn, Apiwat Ditthapron, Karis Matchaparn, Tanaboon Tongbuasirilai et al. "Affective EEG-Based Person Identification Using the Deep Learning Approach", IEEE Transactions on Cognitive and Developmental Systems, 2020 Publication	<1 %
102	zero.inf.cs.cmu.edu Internet Source	<1 %

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		