



**NEURAL NETWORK-BASED ECHOCARDIOGRAM  
VIDEO CLASSIFICATION BY INCORPORATING  
DYNAMIC INFORMATION AND ATTENTION**



**MECHANISM**

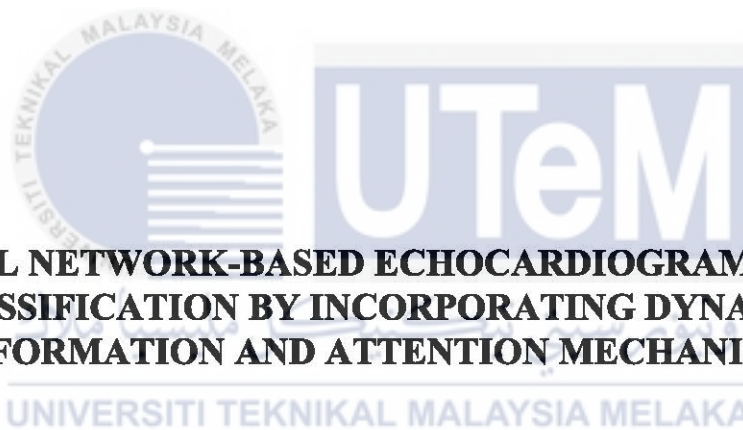
**YE ZI**

**DOCTOR OF PHILOSOPHY**

**2022**



**Faculty of Information and Communication Technology**



**NEURAL NETWORK-BASED ECHOCARDIOGRAM VIDEO  
CLASSIFICATION BY INCORPORATING DYNAMIC  
INFORMATION AND ATTENTION MECHANISM**

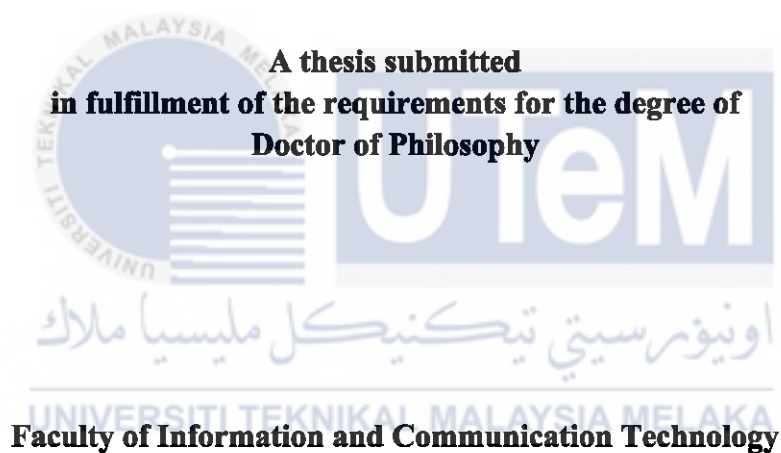
**YE ZI**

**Doctor of Philosophy**

**2022**

**NEURAL NETWORK-BASED ECHOCARDIOGRAM VIDEO  
CLASSIFICATION BY INCORPORATING DYNAMIC INFORMATION AND  
ATTENTION MECHANISM**

**YE ZI**

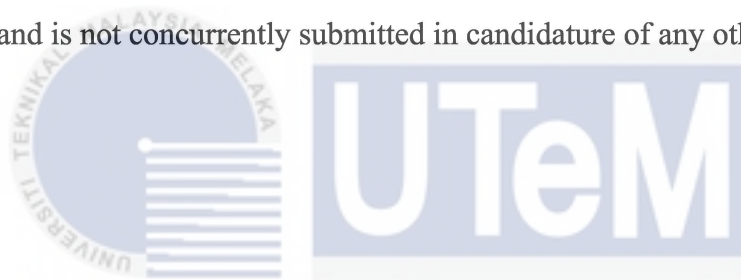


**UNIVERSITI TEKNIKAL MALAYSIA MELAKA**

**2022**

## DECLARATION

I declare that this thesis entitled “Neural Network-based Echocardiogram Video Classification by Incorporating Dynamic Information and Attention Mechanism” is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.



اونيورتي تيكنيكل ماليسيا ملاك  
Signature :.....  
UNIVERSITI TEKNIKAL MALAYSIA MELAKA  
Name : YE ZI  
Date : 28/09/2022

## APPROVAL

I hereby declare that I have read this thesis and in my opinion, this thesis is sufficient in terms of scope and quality for the award of the degree of Doctor of Philosophy.

 Signature :  .....

Supervisor Name : Dr. Yogan Jaya Kumar .....

Date : 28/09/2022 .....



اونيورسيتي تيكنيكل مليسيا ملاك

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

## DEDICATION

This thesis is dedicated to the people who have supported me throughout my education.

Thanks for making me see this adventure through to the end.



## ABSTRACT

Echocardiography, the use of ultrasound waves to investigate the heart's action, is the primary physiological test for cardiovascular disease diagnoses. The determination of the probe viewpoint forms an essential step in echocardiographic image analysis. Some of such views are identified as standard views due to the presentation and ease of the evaluations of their major cardiac structures. Finding valid cardiac views has traditionally been a laborious process and interpreted manually by the specialist, so there exists significant interest in providing an automated classification of echocardiograms in order to speed up the diagnosis process. However, the traditional machine learning methods require time-consuming and operator-dependent manual selection and annotation of features. Therefore, this study aims to simplify the diagnosis process by providing an automated classification of standard cardiac views based on deep learning technology. More importantly, our research considers and assesses some new neural network architectures driven by action recognition in video. For this aim, two classes of neural network architectures have been outlined: the CNN+BiLSTM model and the Spatiotemporal-BiLSTM model. It is finally verified that these methods aggregating dynamic information receive a more substantial classification effect. In addition, previous observations concluded that the most significant challenge lies in distinguishing among the various adjacent views. To this end, our study further aimed to adopt the attention mechanism for designing efficient neural networks. We proposed an ECHO-Attention architecture consisting of two parts. We first design an ECHO-ACTION block, which efficiently encodes Spatio-temporal features, channel-wise features, and motion features. Then, we insert this block into existing ResNet architectures, combined with a self-attention module to ensure its echocardiogram classification task-related focus, to form an effective ECHO-Attention network. All of these experiments are implemented on our privately collected dataset of 2693 videos acquired from 267 patients, which trained cardiologists have manually labeled. The evidence from this study showed that all the proposed methods yielded good results. The ECHO-Attention architecture provides the best classification performance (overall accuracy of 94.81%) on the entire video sample and achieved significant improvements on the classification of anatomically similar views (precision 88.65% and 81.70% for PSAX-AP and PSAX-MID on 30-frame clips, respectively).



# **KLASIFIKASI VIDEO EKO-KARDIOGRAM BERASASKAN RANGKAIAN NEURAL DENGAN MENGGABUNGGAN MAKLUMAT DINAMIK DAN MEKANISME PERHATIAN**

## **ABSTRAK**

*Ekokardiografi, yang menggunakan ultrabunyi untuk memeriksa fungsi jantung, adalah ujian fisiologi utama untuk diagnosis penyakit kardiovaskular. Penentuan sudut pandangan kuar adalah langkah penting dalam analisis imej ekokardiografi. Beberapa pandangan ini dikenal pasti sebagai pandangan piawai kerana tahap kemudahan penilaian dan pandangan struktur jantung utama. Secara tradisinya, mencari pandangan jantung secara berkesan menjadi satu proses yang sukar dan perlu ditafsirkan secara manual oleh pakar perubatan. Maka terdapat minat yang besar untuk menyediakan klasifikasi automatik ekokardiogram untuk mempercepatkan proses diagnostik. Walau bagaimanapun, kaedah pembelajaran mesin tradisional memakan masa dan bergantung pada pengendali untuk memilih dan menganotasi ciri secara manual. Oleh itu, kajian ini bertujuan untuk memudahkan proses diagnostik dengan menyediakan klasifikasi automatik pandangan jantung piawai berdasarkan teknik pembelajaran mendalam. Lebih penting lagi, kajian kami mempertimbangkan dan menilai beberapa seni bina rangkaian neural baharu yang didorong oleh pengecaman tindakan dalam video. Untuk tujuan ini, dua kelas seni bina rangkaian neural telah digariskan: model CNN+BiLSTM dan model Spatiotemporal-BiLSTM. Akhirnya, didapati kesan pengelasan yang diperolehi oleh kaedah mengumpulkan maklumat dinamik ini adalah lebih ketara. Tambahan pula, pendapat terdahulu menyimpulkan bahawa cabaran terbesar adalah dalam membezakan antara pelbagai sudut pandangan yang bersebelahan. Untuk tujuan ini, kajian kami selanjutnya menggunakan mekanisme perhatian untuk mereka bentuk rangkaian neural yang cekap. Kami mencadangkan seni bina ECHO-Attention yang terdiri daripada dua bahagian. Kami mula-mula mereka bentuk model ECHO-ACTION yang dapat mengekod ciri ruang-masa, ciri arah saluran dan ciri gerakan dengan cekap. Kemudian kami memasukkan blok ini ke dalam seni bina ResNet sedia ada, menggabungkan model perhatian sendiri untuk memastikan klasifikasi ekokardiografinya tugas berfokus, bagi membentuk rangkaian ECHO-Attention yang cekap. Semua eksperimen ini dilakukan pada set data yang dikumpul secara peribadi dan mengandungi 2693 video daripada 267 pesakit yang dilabel secara manual oleh pakar kardiologi terlatih. Bukti daripada kajian ini menunjukkan bahawa semua kaedah yang dicadangkan mencapai hasil yang baik. Seni bina ECHO-Attention memberikan prestasi klasifikasi terbaik (ketepatan keseluruhan 94.81%) pada keseluruhan sampel video dan mencapai peningkatan ketara pada klasifikasi pandangan yang serupa secara anatomi (ketepatan 88.65% dan 81.70% untuk PSAX-AP dan PSAX-MID pada 30 bingkai klip, masing-masing).*



## ACKNOWLEDGEMENTS

Countless people supported my effort on this thesis. Thank you to my supervisors, Dr. Yogan Jaya Kumar and Prof. Ts. Dr. Goh Ong Sing, for your patience, guidance, and support. They provided invaluable feedback on my analysis and framing, at times responding to emails late at night and early in the morning. I have significantly benefited from your wealth of knowledge and meticulous editing.

Several other professors gave helpful advice as I wrote. Thank you to Fengyan Song, who generously took time out of his schedules to participate in my research and make this project possible. Thank you to Xianda Ni, for the professional assistance you have provided. Your encouraging words and thoughtful, detailed feedback have been very important to me. I would also like to acknowledge Pantai Hospital Ayer Keroh for valuable support regarding data resources.

Most importantly, I am grateful for my family's unconditional, unequivocal, and loving support.

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

## TABLE OF CONTENTS

	PAGE
<b>DECLARATION</b>	
<b>APPROVAL</b>	
<b>DEDICATION</b>	
<b>ABSTRACT</b>	i
<b>ABSTRAK</b>	ii
<b>ACKNOWLEDGEMENTS</b>	iii
<b>TABLE OF CONTENTS</b>	iv
<b>LIST OF TABLES</b>	viii
<b>LIST OF FIGURES</b>	xi
<b>LIST OF APPENDICES</b>	xv
<b>LIST OF ABBREVIATIONS</b>	xvi
<b>LIST OF SYMBOLS</b>	xix
<b>LIST OF PUBLICATIONS</b>	xx
<b>CHAPTER</b>	
<b>1. INTRODUCTION</b>	<b>1</b>
1.1 Background	1
1.1.1 Diagnostic modalities of heart	1
1.1.2 Two-dimensional echocardiography	4
1.1.3 Echocardiogram procedure	7
1.2 Problem statement	9
1.3 Research objective	11
1.4 Scope of research	14
1.5 Contribution of research	15
1.6 Thesis outline	15
<b>2. LITERATURE REVIEW</b>	<b>18</b>
2.1 Introduction	18
2.2 Traditional machine learning	19
2.2.1 Data pre-processing	19
2.2.2 Feature engineering	20
2.2.3 Training model	21
2.2.4 Classification	21
2.3 Neural network	21
2.3.1 Development of neural network	22
2.3.2 Single-layer perceptron	24
2.3.3 Multilayer perceptron	25
2.4 Convolutional neural network	27
2.4.1 Composition of convolutional neural network	29
2.4.2 Techniques related to convolutional neural network	38
2.4.3 The training of the neural network	41
2.4.4 Further explanation of CNN structure	48
2.4.5 Classic CNN architectures	51
2.5 Recurrent neural network	64
2.5.1 Long Short-Term Memory (LSTM)	65
2.5.2 Bi-LSTM network	66
2.6 Multiple-stream fusion network	67

2.7	Visual attention mechanism	69
2.7.1	Hard and Soft	70
2.7.2	Spatial attention module	71
2.7.3	Channel attention module	72
2.7.4	Hybrid attention module	73
2.7.5	Classification of video using visual attention	75
2.8	Review on automatic cardiac view classification	77
2.8.1	General machine learning methods	78
2.8.2	Convolutional neural network methods	79
2.8.3	Fused network methods	80
2.8.4	Attention-based methods	86
2.9	Summary	88
<b>3.</b>	<b>METHODOLOGY</b>	<b>91</b>
3.1	Introduction	91
3.2	Research design	91
3.3	Research operational framework	93
3.3.1	Phase 1: Preliminary study and data acquisition	96
3.3.2	Phase 2: Data processing	100
3.3.3	Phase 3: Cardiac view classification using classic 2D CNN architectures	102
3.3.4	Phase 4: Dynamic information enhancement for echo video classification	103
3.3.5	Phase 5: Echo classification using a hybrid attention-based approach	104
3.3.6	Phase 6: Report writing	105
3.4	Experimental evaluation	106
3.5	Statistical significance tests for comparing models	107
3.5.1	The problem of model selection	108
3.5.2	Paired sample t-test	109
3.6	Summary	110
<b>4.</b>	<b>CARDIAC VIEW CLASSIFICATION USING CLASSIC 2D CNN ARCHITECTURES</b>	<b>111</b>
4.1	Introduction	111
4.2	Overview of approach	112
4.3	Classic 2D CNNs architectures	113
4.4	Experiment setting	116
4.4.1	Experimental requirements	116
4.4.2	Data augmentation	116
4.4.3	Transfer learning	117
4.4.4	Hyperparameter setting	117
4.5	Experiment results	118
4.6	Discussion	124
4.7	Summary	127
<b>5.</b>	<b>DYNAMIC INFORMATION ENHANCEMENT FOR ECHO VIDEO CLASSIFICATION</b>	<b>128</b>
5.1	Introduction	128
5.2	Overview of approach	128

5.3	CNN+BiLSTM architecture	132
5.4	Spatiotemporal-BiLSTM architecture	135
5.4.1	Convolutional-to-Fully Connected	135
5.4.2	Fully Connected-to-BiLSTM	135
5.4.3	BiLSTM-to-Classification	136
5.5	Experiment setting	136
5.5.1	Transform learning	136
5.5.2	Hyperparameter setting	137
5.6	Experiment results	138
5.7	Discussion	147
5.7.1	Dynamic information effects	147
5.7.2	Optical flow effects	148
5.7.3	Video frames effects	149
5.7.4	Additional factors	150
5.7.5	Impact on the domain	150
5.7.6	Study limitation	151
5.8	Summary	151
<b>6.</b>	<b>CLASSIFICATION OF ECHOCARDIOGRAM USING A HYBRID ATTENTION-BASED APPROACH</b>	<b>153</b>
6.1	Introduction	153
6.2	Overview of approach	154
6.3	ECHO-Attention model	157
6.3.1	Spatio-Temporal Module and Channel Module	157
6.3.2	Motion Module	159
6.3.3	ECHO-ACTION Block	162
6.3.4	Whole architecture	162
6.4	Experiment setting	165
6.4.1	Data augmentation	165
6.4.2	Training approach	165
6.4.3	Loss function	166
6.5	Experiment results	166
6.6	Discussion	172
6.6.1	Why attention mechanism	173
6.6.2	How attention mechanism works	175
6.6.3	How dynamic effects	177
6.6.4	Additional findings	177
6.7	Summary	178
<b>7.</b>	<b>CONCLUSION AND RECOMMENDATIONS</b>	<b>180</b>
7.1	Introduction	180
7.2	Proposed methods	180
7.2.1	Cardiac view classification using classic 2D CNN architectures	181
7.2.2	Dynamic information enhancement for echo video classification	182
7.2.3	Classification of echocardiogram using a hybrid attention-based approach	182
7.3	Contribution of the study	183
7.4	Future works	185

7.5	Limitation and challenges	187
7.6	Summary	189
<b>REFERENCES</b>		<b>190</b>
<b>APPENDICES</b>		<b>208</b>



## LIST OF TABLES

TABLE	TITLE	PAGE
2.1	The published journals within the past five years focus on traditional Machine Learning approaches regarding classifying echocardiographic views automatically	82
2.2	The published journals within the past five years focus on Deep Learning methods regarding classifying echocardiographic views automatically	83
2.3	The published journals within the past five years focus on the fused network regarding classifying echocardiographic views automatically	85
2.4	The comparison table of output categories, input data size and overall recognition accuracies among existing models	89
3.1	The overview of the dataset	99
3.2	Table demonstrating classic 2D CNNs assessed in this study	102
3.3	Table demonstrating dynamic Neural Networks assessed in this study	104
4.1	The overall accuracy on the test set	118
4.2	Table of precision and recall for VGG16 model	119
4.3	Table of precision and recall for Inception-Resnet-V2 model	119
4.4	Table of precision and recall for Xception model	119
4.5	Confusional matrix for VGG16 model, demonstrating actual view labels on the y-axis, and neural network-predicted view labels on the x-axis by view category for video classification	120
4.6	Confusional matrix for Inception-Resnet-V2 model	120
4.7	Confusional matrix for Xception model	120
4.8	Classification accuracy on testing videos of different fixed frames	121

4.9	Computing time on testing videos of different fixed frames (Seconds for each sample)	121
4.10	Paired samples test between Xception and Inception-Resnet-V2	124
4.11	Paired samples test between Xception and VGG16	124
5.1	The overview of the datasets of 30-frame	129
5.2	Table demonstrating the overall accuracy for each of the networks on the test set	138
5.3	Table demonstrating the computing time for each of the networks on the test set	139
5.4	Table of precision and recall for VGG16+BiLSTM model	143
5.5	Table of precision and recall for Inception-Resnet-V2+BiLSTM model	143
5.6	Table of precision and recall for Xception+BiLSTM model	143
5.7	Table of precision and recall for Spatiotemporal-BiLSTM model	144
5.8	Confusional matrix for Xception+BiLSTM model	144
5.9	Confusional matrix for VGG16+BiLSTM model	144
5.10	Confusional matrix for Inception-Resnet-V2+BiLSTM model	145
5.11	Confusional matrix for Spatiotemporal-BiLSTM model	145
5.12	Comparison of CNN+BiLSTM structure and Spatiotemporal-BiLSTM structure	146
5.13	Classification accuracy on various video frames	146
5.14	Paired samples test between Xception+BiLSTM and Spatiotemporal-BiLSTM	147
6.1	The overview of the datasets of 30-frame	156
6.2	Table demonstrating the overall accuracy for each of the networks on the datasets of 30-frame	167
6.3	Classification accuracy on entire videos	168
6.4	Computing time of different models in terms of various frames (seconds/sample)	169
6.5	Precision and recall results using ECHO-Attention network (ResNet-50) on the datasets of 30-frame segments	169



6.6	Precision and recall results using SpatioTemporal-BiLSTM (ResNet-50) on the datasets of 30-frame segments	169
6.7	Precision and recall results using CNN+BiLSTM (ResNet-50) on the datasets of 30-frame segments	170
6.8	Precision and recall results using ECHO-Attention network (ResNet-50) on the datasets of entire videos	170
6.9	Confusion matrices for echo view classification. The ECHO-Attention network (ResNet-50) was evaluated on the datasets of 30-frame segments	170
6.10	Confusion matrices for echo view classification. The SpatioTemporal-BiLSTM (ResNet-50) was evaluated on the datasets of 30-frame segments	171
6.11	Confusion matrices for echo view classification. The CNN+BiLSTM (ResNet-50) was evaluated on the datasets of 30-frame segments	171
6.12	Confusion matrices for echo view classification. The ECHO-Attention network (ResNet-50) was evaluated on the datasets of entire videos	171
6.13	Paired samples test between Xception+BiLSTM and ECHO-Attention network (ResNet-50)	172

## LIST OF FIGURES

FIGURE	TITLE	PAGE
1.1	Anatomical diagram of the heart (Michigan, 2019)	2
1.2	Echocardiographic imaging planes (Michigan, 2019)	5
1.3	The echocardiographic diagnosis procedures	9
2.1	Traditional machine learning flow chart	20
2.2	Development of neural network	23
2.3	Basic neurons mathematical model (Karimboyevich et al., 2022)	25
2.4	Structure diagram of multilayer perceptron (Agahian et al., 2022)	26
2.5	Convolutional neural network (Li et al., 2021)	29
2.6	Schematic diagram of convolutional layer processing without fill operation (Albawi et al., 2017)	30
2.7	Schematic diagram of convolutional layer processing with zero filling operation (Albawi et al., 2017)	31
2.8	Schematic diagram of maximum and average pooling operations (Kernbach et al., 2022)	32
2.9	Schematic diagram of the fully connected layer (Kalaycı et al., 2022)	33
2.10	Schematic diagram of the Sigmoid function (Szandała, 2021)	34
2.11	Schematic diagram of the Tanh function (Fan, 2000)	34
2.12	Schematic diagram of the ReLU function (Szandała, 2021)	35
2.13	Schematic diagram of local network connection	39
2.14	Schematic diagram of receptive field (Liu et al., 2022)	40
2.15	Schematic diagram of receptive field calculation	41

2.16	Output characteristics of a fingerprint through the middle layer of CNN (Cao et al., 2015)	49
2.17	Output characteristics of ECG through the middle tier of CNN (Jin et al., 2015)	50
2.18	The structure of AlexNet (Xiao et al., 2017)	52
2.19	The structure of VggNet (Simonyan et al., 2014)	54
2.20	The structure of residual block (He et al., 2016)	55
2.21	The comparison of linear convolutional filter (left) and MLP convolutional filter (right) (Lin et al., 2013)	56
2.22	Inception module I (Szegedy et al., 2015)	59
2.23	Inception module II (Ioffe et al., 2015)	59
2.24	Inception module III (Szegedy et al., 2016)	60
2.25	The Xception architecture (Chollet, 2017)	61
2.26	An “Extreme” version of the Inception module (Chollet, 2017)	61
2.27	The structure of RNN (Sherstinsky, 2020)	64
2.28	The structure of LSTM (Lindemann et al., 2021)	65
2.29	The structure of Bi-LSTM (Xiang et al., 2020)	67
2.30	The structure of the two-stream neural network (Simonyan et al., 2014)	68
2.31	The basic soft-attention network structure	70
2.32	Non-local neural networks (Wang et al., 2018)	71
2.33	The calculation process of channel attention (Hu et al., 2018)	72
2.34	Parallel hybrid attention mechanism (Fu et al., 2019)	74
2.35	Serial hybrid attention mechanism (Fu et al., 2019)	74
2.36	ACTION module consists of three sub-modules, Spatio-Temporal Excitation (STE) module, Channel Excitation (CE) module, and Motion Excitation (ME) module (Wang et al., 2021)	77
2.37	An example of spatial attention (Niu et al., 2021)	88
2.38	The scatter plot visualizes the relationship between experimental results and scale of dataset of different existing proposed approaches for cardiac view identification.	90

3.1	Overview of research study	94
3.2	Research operational framework	95
3.3	Inclusion of 9 different anatomical echocardiographic views: apical two-chamber(A2C), apical three-chamber(A3C), apical four-chamber(A4C), apical five-chamber(A5C), parasternal long-axis(PLAX), parasternal short-axis aortic valve focused(PSAX-AV), parasternal short-axis apical focused(PSAX-AP), parasternal short-axis mitral valve focused(PSAX-MV), and parasternal short-axis papillary focused(PSAX-MID); A class of "others" is also investigated for other cardiac aspects	98
3.4	The numbers of patients for every ten viewpoints in the database and applied for training, validation, and testing correspondingly	99
3.5	The numbers of videos for every ten viewpoints in the database and applied for training, validation, and testing correspondingly	100
3.6	Procedures used for removing the class label and segmenting relevant visual structures	101
3.7	2-class Classification Confusion Matrices	106
4.1	Workflow of cardiac view classification using classic 2D CNNs architectures	113
4.2	Modified VGG architecture	114
4.3	Modified Inception-Resnet-V2 architecture	115
4.4	Modified Xception architecture	115
4.5	Model evaluating approach	116
4.6	Screenshot of Xception when testing on 30-frame samples	121
4.7	Screenshot of VGG16 when testing on 30-frame samples	122
4.8	Screenshot of Inception-Resnet-V2 when testing on 30-frame samples	123
5.1	Workflow of cardiac view classification using dynamic information enhancement architectures	130
5.2	Testing procedure for dynamic information enhancement architectures	131
5.3	Investigation on testing video samples of various frames with training on 30-frame clips	132
5.4	Investigation on testing video samples of various frames with training on 15-frame clips	132

5.5	The structure of CNN+BiLSTM architectures for video classification	134
5.6	The structure of Spatiotemporal-BiLSTM architecture for video classification	134
5.7	Screenshot of Xception+BiLSTM	139
5.8	Screenshot of VGG16+BiLSTM	140
5.9	Screenshot of Inception-Resnet-V2+BiLSTM	141
5.10	Screenshot of Spatiotemporal-BiLSTM model	142
5.11	Examples of the short-axis view of the papillary muscle and that of the apical level	151
6.1	Workflow of cardiac view classification using a hybrid attention-based approach	156
6.2	Testing procedure for entire videos	157
6.3	(a) Spatio-Temporal Module (STM), (b) Channel Module (CM)	158
6.4	Motion Module (MM)	161
6.5	ECHO-ACTION block	162
6.6	ECHO-Attention for ResNet-50 architecture	164
6.7	The screenshot of ECHO-Attention (ResNet-50)	168
6.8	The papillary muscles of PSAX-MID	174
6.9	The schematic diagram of attention mechanism using heat map visualization (Hafiz et al., 2021).	174
6.10	The structure of the self-attention mechanism (Gaihua et al., 2021).	176
7.1	The current Cardiac CAD research focuses on four parts	186

## LIST OF APPENDICES

APPENDIX	TITLE	PAGE
A	The main code for Classic 2D CNN (Xception), using Pytorch	208
B	The main code for Xception+BiLSTM model, using Pytorch	209
C	The main code for Spatiotemporal-BiLSTM model, using Pytorch	211
D	The main code for ECHO-Attention model (ResNet-50), using Pytorch	214



## LIST OF ABBREVIATIONS

2D	-	Two-Dimensional
A2C	-	Apical Two-Chamber View
A3C	-	Apical Three-Chamber View
A4C	-	Apical Four-Chamber View
A5C	-	Apical Five-Chamber View
aCMQ	-	Automated Cardiac Motion Quantification
ACTION	-	spAtio-temporal, Channel and moTion excitatION
ADAM	-	Adaptive Moment Estimation
AI	-	Artificial Intelligence
AVA	-	Aortic Valve Opening Area
BiLSTM	-	Bidirectional Long Short-Term Memory
BN	-	Batch Normalization
BoW	-	Bag-Of-Words
BP	-	Background Propagation
CAD	-	Computer-Aided Diagnosis
CE	-	Channel Excitation
CM	-	Channel Module
CNN	-	Convolutional Neural Network
CPU	-	Central Processing Unit
CT	-	Computed Tomography
CVD	-	Cardiovascular Diseases



DICOM	-	Digital Imaging And Communications In Medicine
DL	-	Deep Learning
ECG	-	Electrocardiogram
EF	-	Ejection Fraction
FC	-	Fully Connected Layer
FN	-	False Negative
FoV	-	Field Of View
FP	-	False Positive
GLM	-	Generalized Linear Model
GPU	-	Graphics Processing Unit
IRB	-	Institutional Review Board
LSTM	-	Long Short-Term Memory
MB-SGD	-	Mini-Batch Stochastic Gradient Descent
ME	-	Motion Excitation
MINIST	-	Modified National Institute Of Standards And Technology
ML-DNN	-	Multi-Loss Deep Neural Network
MLP	-	Multilayer Perceptron
MM	-	Motion Module
NIN	-	Network In Network
NLP	-	Natural Language Processing
PLAX	-	Parasternal Long-Axis View Of The Left Ventricle
PSAX	-	Parasternal Short-Axis View
PSAX-AP	-	Parasternal Short-Axis View Of The Left Ventricle At The Apical Level
PSAX-AV	-	Parasternal Short-Axis View Of The Aorta
PSAX-MID	-	Parasternal Short-Axis View Of The Left Ventricle At Papillary Muscle Level
PSAX-MV	-	Parasternal Short-Axis View Of The Left Ventricle At The Mitral

		Valve Level
RNN	-	Recurrent Neural Network
SC	-	Subcostal
SSN	-	Suprasternal
STE	-	Spatio-Temporal Excitation
STM	-	Spatio-Temporal Module
STN	-	Spatial Transformer Network
SVHN	-	Street View House Numbers
SWWAE	-	Stacked What-Where Auto-Encoders
TN	-	True Negative
TP	-	True Positive
US	-	Ultrasound
VGG	-	Visual Geometry Group
VTN	-	Video Transformer Network

