

Novel Deep Learning Architectures for Enhanced COVID-19 Detection

NARENDRA KUMAR MISHRA



DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY DELHI

APRIL 2025

©Indian Institute of Technology Delhi (IITD), New Delhi, 2025

Novel Deep Learning Architectures for Enhanced COVID-19 Detection

A THESIS

submitted by

Narendra Kumar Mishra
2018EEZ8568

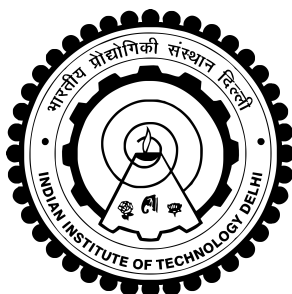
under the supervision of

Prof. Shiv Dutt Joshi, Indian Institute of Technology Delhi
Dr. Pushpendra Singh, Jawaharlal Nehru University New Delhi

for the award of the degree

of

Doctor of Philosophy



Department of Electrical Engineering
INDIAN INSTITUTE OF TECHNOLOGY DELHI

April 2025

Dedicated to
my kids
Devavrat and Aarunya

CERTIFICATE

This is to certify that the thesis titled **Novel Deep Learning Architectures for Enhanced COVID-19 Detection**, submitted by **Narendra Kumar Mishra (2018EEZ8568)**, to the Indian Institute of Technology, Delhi, for the award of the degree of **Doctor of Philosophy**, is a bona fide record of the research work done by him under our supervision. He has fulfilled the requirements for the submission of the thesis, which to the best of our knowledge has reached the required standard.

The material contained in the thesis has not been submitted in part or full to any other University or Institute for the award of any degree or diploma.

Shiv Dutt Joshi
Professor
Electrical Engineering
IIT-Delhi, 600 036

Pushpendra Singh
Associate Professor
School of Engineering
Jawaharlal Nehru University
New Delhi, 110 067

Place: New Delhi

Date: April 2025

ACKNOWLEDGEMENTS

I express my deepest gratitude to my supervisor, Professor Shiv Dutt Joshi, for his invaluable guidance and mentorship throughout my research journey. His unwavering support and consistent motivation, coupled with his never-failing patience and understanding, have been instrumental in navigating both academic and administrative challenges.

I am equally thankful to my co-supervisor, Dr. Pushpendra Singh, whose guidance and encouragement were pivotal in maintaining a balance between my professional responsibilities and research commitments. His steadfast support and positive reinforcement kept me focused and motivated.

I would like to extend my heartfelt appreciation to my present establishment, Weapons and Electronic Systems Engineering Establishment (WESEE), and my parent organization, Defence Research & Development Organisation (DRDO). I am profoundly grateful for the opportunity they provided me to pursue this PhD, enabling me to improve my expertise in the latest AI fields. Their belief in my potential and their commitment to my professional growth have empowered me to contribute more effectively to our collective goals. I am proud to be part of such forward-thinking organizations that prioritize continuous learning and innovation.

I am also grateful to the Student Review Committee (SRC) members for their invaluable insights and constructive suggestions, which have been essential in refining my research work.

Lastly, the unwavering support and endless encouragement from my family have been the foundation of my achievements. Their belief in my abilities and constant support made this challenging journey possible. I owe a profound debt of gratitude for their love and sacrifice.

New Delhi

Narendra Kumar Mishra

April 2025

2018EEZ8568

ABSTRACT

The COVID-19 pandemic has highlighted the need for efficient and reliable diagnostic methods. This thesis explores the integration of artificial intelligence (AI), deep learning, and signal processing techniques to enhance COVID-19 detection through medical imaging and acoustic data analysis.

The research begins with a comprehensive review of the literature, identifying significant advancements and challenges in AI-based COVID-19 detection. Convolutional neural network (CNN) models, utilizing transfer learning with architectures such as VGG16 and ResNet50, are designed to detect COVID-19 from CT scan images. These models are optimized through data augmentation and fine-tuning techniques, showing robust diagnostic capabilities validated by stratified 5-fold cross-validation.

The study also investigates COVID-19 detection using cough sounds, exploring three strategies: machine learning with spectral features, deep learning with ResNet-50, and transfer learning with a Vision Transformer model. The experimental results demonstrate promising performance, suggesting potential applications for non-invasive detection.

A novel Probabilistic Pooling Convolutional Neural Network (PP-CNN) is introduced to enhance image classification. This model integrates probabilistic outputs from MaxPooling, MinPooling, and MaxMinPooling layers, achieving improved accuracy across diverse datasets, including CIFAR-10, CIFAR-100, CT scans, and X-ray images.

Additionally, an innovative approach to analyzing non-stationary signals using generalized constant-overlap-add (COLA) windows and the short-time discrete cosine transform (STDCT) is presented. This methodology enhances time-frequency representation and is applied to detect COVID-19 from cough sounds, validated through experiments on synthetic and real-life data.

This thesis provides significant contributions to AI-based COVID-19 detection, offering robust, non-invasive, and efficient diagnostic methods. The findings pave the way for future research in integrating AI and signal-processing techniques for broader medical applications, ultimately aiming to improve healthcare outcomes.

सार

कोविड-19 महामारी ने बेहतर और भरोसेमंद निदान विधियों की आवश्यकता को सामने लाया है। यह शोध-प्रबंध आर्टिफिशियल इंटेलिजेंस (एआई), डीप लर्निंग और सिग्नल प्रोसेसिंग तकनीकों के मेल से मेडिकल इमेजिंग और अकूस्टिक डेटा के माध्यम से कोविड-19 के पहचान को और प्रभावी बनाने पर केंद्रित है।

शोध की शुरुआत संबंधित साहित्य की समीक्षा से होती है, जिसमें एआई आधारित कोविड-19 पहचान के क्षेत्र में हुई प्रगति और चुनौतियों को पहचाना गया है। सीटी स्कैन इमेज से कोविड-19 की पहचान के लिए ट्रांसफर लर्निंग आधारित कन्वोल्यूशनल न्यूरल नेटवर्क मॉडल, जैसे वीजीजी16 और रेसनेट50, का प्रयोग किया गया है। इन मॉडलों को डेटा ऑगमेंटेशन और फाइन-ट्यूनिंग तकनीकों द्वारा बेहतर किया गया है, और इन्हें स्ट्रैटिफाइड फाइन-फोल्ड क्रॉस-वैलिडेशन द्वारा जांचा गया है।

खांसी की आवाज़ से कोविड-19 पहचान के लिए तीन रणनीतियों का विश्लेषण किया गया है: स्पेक्ट्रल फीचर के साथ मशीन लर्निंग, रेसनेट-50 के साथ डीप लर्निंग, और विज़न ट्रांसफार्मर मॉडल के साथ ट्रांसफर लर्निंग। प्रयोगों के नतीजे उत्साहवर्धक रहे हैं, जो बिना शरीर को छुए निदान की दिशा में संभावनाएं दिखाते हैं।

एक नया पीपी-सीएनएन मॉडल भी प्रस्तुत किया गया है, जो इमेज क्लासिफिकेशन को बेहतर बनाता है। यह मॉडल मैक्सपूलिंग, मिनपूलिंग और मैक्समिनपूलिंग लेयर्स के प्रायिक आउटपुट्स को मिलाकर बेहतर परिणाम देता है। इसका मूल्यांकन सिफार-10, सिफार-100, सीटी स्कैन और एक्स-रे जैसे विभिन्न डेटासेट्स पर किया गया है।

इसके अलावा, एसटीडीसीटी और कोला विंडो जैसे तरीकों के साथ नॉन-स्टेशनरी सिग्नल्स के विश्लेषण की एक नई पद्धति भी विकसित की गई है। यह पद्धति समय-आवृत्ति रिप्रेजेंटेशन को बेहतर बनाती है और इसका उपयोग खांसी की आवाज़ से कोविड-19 पहचान के लिए किया गया है। इस विधि को सिंथेटिक और वास्तविक डेटा दोनों पर सफलतापूर्वक परखा गया है।

यह शोध कोविड-19 की पहचान के लिए एआई और सिग्नल प्रोसेसिंग आधारित नई, प्रभावी और गैर-संपर्क विधियों का योगदान देता है। इसके निष्कर्ष इन तकनीकों के भविष्य के चिकित्सकीय प्रयोगों के लिए शोध के नए रास्ते खोलते हैं, जिनका उद्देश्य स्वास्थ्य सेवाओं को और सशक्त बनाना है।

Contents

ACKNOWLEDGEMENTS	i
ABSTRACT	ii
ABSTRACT	iii
LIST OF TABLES	x
LIST OF FIGURES	xiv
ABBREVIATIONS	xv
1 INTRODUCTION	1
1.1 Motivation	1
1.2 Statement of the Problem	2
1.3 Objectives and Scope	4
1.4 Significance and Relevance	5
1.5 Ethical Considerations in AI-based Healthcare	5
1.6 Structure of the Thesis	6
2 BASIC PRELIMINARIES AND LITERATURE SURVEY	9
2.1 Introduction	9
2.2 Overview of Artificial Intelligence Techniques for COVID-19 Detection	10
2.3 Overview of Pipeline for COVID-19 Image Classification	12
2.3.1 Data Pre-processing	13
2.3.2 Feature Extraction and Classification	14
2.3.3 Performance Evaluation	15
2.3.4 Explanation of Model Prediction	15
2.4 Medical Imaging Modalities	16

2.4.1	CXR dataset	16
2.4.2	CT dataset	19
2.4.3	CT and CXR dataset	20
2.4.4	CT vs CXR	20
2.5	Deep Learning in Medical Image Analytics	21
2.5.1	Transfer Learning work	22
2.5.2	Custom CNN Work	28
2.5.3	Self-Supervised Learning	28
2.6	Challenges and Opportunities in COVID-19 Detection	30
2.6.1	Challenges	30
2.6.2	Opportunities	36
2.7	Conclusion	39
3	DEEP LEARNING-BASED COVID-19 DETECTION FROM CT SCANS	41
3.1	Introduction	41
3.2	Experimental Design	43
3.2.1	Dataset	43
3.2.2	Proposed Deep Learning Architecture	44
3.3	Results analysis and Discussions	49
3.4	Conclusion	60
4	COUGH SOUND ANALYSIS FOR COVID-19 DETECTION USING SPECTRAL FEATURES AND DEEP NEURAL NETWORKS	63
4.1	Introduction	63
4.1.1	Literature Survey	65
4.2	Designing the Experimental Framework for COVID-19 Detection	66
4.2.1	Data Collection	66
4.2.2	Data Splitting	67
4.2.3	Voice Activity Detection	68
4.2.4	Segmentation	69
4.2.5	Shuffling	69
4.2.6	Spectral Features Computation	69

4.2.7	MelSpectrograms Computation	71
4.3	Proposed Architecture	73
4.3.1	Model-1: Machine Learning based COVID-19 Detection from Cough	74
4.3.2	Model-2: Deep Learning based COVID-19 Detection from Cough	74
4.3.3	Model-3: Transfer Learning based COVID-19 Detection from Cough	75
4.4	Results Analysis and Discussion	77
4.5	Conclusion	84
5	PROBABILISTIC POOLING CNN FOR ENHANCED IMAGE CLASSIFICATION	86
5.1	Introduction	86
5.2	Methodology	90
5.2.1	Theoretical Framework	91
5.2.2	Theoretical Framework	92
5.2.3	Activation Function Strategy	93
5.2.4	Probabilistic Output Fusion	93
5.2.5	Dataset Selection Rationale	94
5.2.6	Architectural Design	94
5.2.7	Structural Block diagram	96
5.2.8	Evaluation Metrics	96
5.3	Implementation	98
5.4	Results Analysis and Discussion	100
5.5	Conclusion	106
6	GENERALIZED COLA WINDOWS AND STDCT FOR IMPROVED COVID-19 DETECTION	108
6.1	Introduction	108
6.2	Preliminaries	111
6.3	Proposed methodology	112
6.3.1	Perfect Reconstruction COLA window	112
6.3.2	TFR using STFT and the proposed COLA windows	114

6.3.3	TFR using Short-time DCT and proposed COLA windows	116
6.4	Experimental Results	117
6.4.1	Analysis of multi-component signal	117
6.4.2	Analysis of amplitude and frequency modulated signals	118
6.4.3	Analysis of a heart sound signal	118
6.4.4	Detection of COVID-19 from cough sounds	119
6.5	Conclusion	126
7	CONCLUSION AND FUTURE SCOPE	127
7.1	Conclusion	127
7.2	Future Scope	128
7.2.1	Expanding and Diversifying Data	128
7.2.2	Model Optimization and Efficiency	128
7.2.3	Multimodal Analysis	128
7.2.4	Generative AI for Comprehensive Multimodal Diagnostics	128
7.2.5	Interdisciplinary Collaboration and Interpretability	129
7.2.6	Integration and Clinical Validation of Deep Learning-Based COVID-19 Detection	129
7.2.7	Comparative Analysis with Traditional Diagnostic Methods	130
7.2.8	Importance of Ablation Studies in Future Research	130
7.2.9	Future Work on CNN and Signal Processing	131
A	Public Imaging Datasets for COVID-19 Diagnosis	132
B	Summary of state-of-art Deep Learning techniques used for the COVID-19 classification using Radio Images	136

List of Tables

2.1	Techniques for visual explanation of Deep CNN	17
3.1	Details of the number of images used for training, validation and testing of the proposed COVID-19 detection model.	44
3.2	Details of convolutional building blocks: VGG16 architecture (Left) and ResNet50 architecture (Right).	46
3.3	Details of Hyperparameters (Left) and data augmentation (Right) used in the proposed COVID-19 detection model.	49
3.4	Details of models and various experiments conducted in the proposed COVID-19 detection model.	51
3.5	Performance measures for Experiment-1 and Experiment-2 on the test dataset.	51
3.6	Performance measures for Experiment-3 and Experiment-4 on the test dataset.	51
3.7	Evaluation Matrix of binary classification, COVID-19 vs. Normal, using VGG16 based architecture (Experiment-1) and using ResNet50 based architecture (Experiment-2).	52
3.8	Evaluation Matrix of multiclass classification, COVID-19 vs. Normal vs. Pneumonia, using VGG16 based architecture (Experiment-3) and using ResNet50 based architecture (Experiment-4).	53
4.1	Details of cough dataset used for development of COVID-19 detection model.	68
4.2	The spectral features computed from cough segments for development of machine learning-based COVID-19 detection model.	71
4.3	Parameters for computation of MelSpectrograms from cough segments. . .	73
4.4	Evaluation Metrics for Binary Classification of Cough Sounds (Healthy vs. COVID-19) Using Three Models. Model-1 : Spectral Feature-Based Machine Learning, Model-2 : Deep Learning via Fine-Tuning of the ResNet50 Model, Model-3 : Transfer Learning via Pre-trained Vision Transformer Model. . .	81
4.5	A comparative study of our proposed model's performance with the state-of-the-art results.	83
4.6	Details of Proposed Model's experimental Hyperparameters.	84

5.1	Summary of recent pooling techniques used in CNNs. Each method is evaluated on different datasets, highlighting the key features of the pooling strategy and corresponding specific benefits. The symbols $\uparrow$ and $\downarrow$ represent an increase and a decrease, respectively.	89
5.2	Details of medical images dataset used for the development of COVID-19 detection model using the proposed MaxMinPooling method.	94
5.3	Experimental Hyperparameter specifications for the proposed MaxMinPooling model.	101
5.4	Comparison of performance metrics for various approaches on CIFAR-10 and CIFAR-100 datasets.	103
5.5	Performance metrics of the experimented methods on X-ray and CT Scan Dataset	104
6.1	The energy (E_x) of the signal (6.24) in the time and TF domains obtained by different approaches.	118
6.2	Performance Comparison in % for detection of COVID-19 from Cough Sounds (Healthy vs. COVID-19)	125
A.1	Public Imaging Datasets used for COVID-19 Diagnosis	132
B.1	Summary of state-of-art DL techniques used for the COVID-19 classification using CXR Abbreviations: Acc.- Accuracy, BP-Bacterial Pneumonia, C-COVID-19, CAM- Class Activation Maps, CAP- Community Acquired Pneumonia, CN- COVID-19 negative, FPN- Feature Pyramid Network, HU- Hounsfield Units, Infl.- Influenza, LT- Lung Tumor, N-Normal, NF- No Findings, P- Pneumonia, Rad.- Radiologist, SARS- Severe Acute Respiratory Syndrome, Seg.- Segmentation, VP- Viral Pneumonia, Sen.- Sensitivity, Spe.- Specificity	137
B.2	Summary of state-of-art DL techniques used for the COVID-19 classification using CT Abbreviations: Acc.- Accuracy, BP-Bacterial Pneumonia, C-COVID-19, CAM- Class Activation Maps, CAP- Community Acquired Pneumonia, CN- COVID-19 negative, FP- Fungal Pneumonia, FPN- Feature Pyramid Network, HU- Hounsfield Units, Infl.- Influenza, LT- Lung Tumor, MP- Mycoplasma Pneumonia, N-Normal, NF- No Findings, P- Pneumonia, Rad.- Radiologist, SARS- Severe Acute Respiratory Syndrome, Seg.- Segmentation, VP- Viral Pneumonia, Sen.- Sensitivity, Spe.- Specificity	151

B.3 Summary of state-of-art DL techniques used for the COVID-19 classification using Multimodality	
Abbreviations: Acc.- Accuracy, BP- Bacterial Pneumonia, C-COVID-19, CAM- Class Activation Maps, CAP- Community Acquired Pneumonia, CN- COVID-19 negative, FP- Fungal Pneumonia, FPN- Feature Pyramid Network, HU- Hounsfield Units, Infl.- Influenza, LC- Lung Cancer, LT- Lung Tumor, MP- Mycoplasma Pneumonia, N-Normal, NF- No Findings, P- Pneumonia, Rad.- Radiologist, SARS- Severe Acute Respiratory Syndrome, Seg.- Segmentation, VP- Viral Pneumonia, Sen.- Sensitivity, Spe.- Specificity	159

List of Figures

2.1	Most common classes considered for labelling of CXR and CT-scan images, where SARS stands for Severe Acute Respiratory Syndrome and MERS stands for Middle East Respiratory Syndrome	10
2.2	CXR images of (2.2a) a COVID-19, (2.2b) a bacterial pneumonia, (2.2c) a viral pneumonia, and (2.2d) a healthy subject	10
2.3	CT-scan images of (2.3a) a COVID-19 and (2.3b) a healthy subject	10
2.4	AI based Techniques and Modalities for COVID-19 Detection.	11
2.5	A workflow of Deep learning based COVID-19 detection pipeline	12
2.6	Schematic representation of a typical Convolutional Neural Network architecture	15
2.7	Schematic representation of Transfer Learning approach	16
2.8	Top 10 Datasets Most Frequently Used for Validating COVID-19 Detection Models	29
2.9	Distribution of Published Papers Utilizing Various Deep Learning Architectures for COVID-19 Detection	30
3.1	Sample CT scan images used for training and evaluation of the proposed COVID-19 detection model. First row: Normal; second row: COVID-19, third row: Pneumonia.	44
3.2	Process flow architecture of the proposed neural network models: (A) Model-1, based on VGG16 architecture. (B) Model-2, based on ResNet50 architecture.	48
3.3	Illustration of the stratified 5-fold cross-validation approach utilized in the proposed model for COVID-19 detection.	50
3.4	Learning curves for binary classification (COVID-19 vs. Normal) using the VGG16 architecture in Experiment-1: (A) Dataset from Fold-1, (B) Dataset from Fold-2, (C) Dataset from Fold-3, (D) Dataset from Fold-4, and (E) Dataset from Fold-5.	54
3.5	Learning curves for binary classification (COVID-19 vs. Normal) with ResNet50-based architecture (Experiment-2): using (A) Fold-1 dataset, (B) Fold-2 dataset, (C) Fold-3 dataset, (D) Fold-4 dataset, and (E) Fold-5 dataset. . .	55

3.6	Learning Curves of multiclass classification, COVID-19 vs. Normal vs. Pneumonia, using VGG16 based architecture (Experiment-3): using (A) Fold-1 dataset, (B) Fold-2 dataset, (C) Fold-3 dataset, (D) Fold-4 dataset, (E) Fold-5 dataset.	56
3.7	Learning Curves of multiclass classification, COVID-19 vs. Normal vs. Pneumonia, using ResNet50 based architecture (Experiment-4): using (A) Fold-1 dataset, (B) Fold-2 dataset, (C) Fold-3 dataset, (D) Fold-4 dataset, (E) Fold-5 dataset.	57
3.8	Confusion Matrices for binary classification, COVID-19 vs. Normal, using VGG16 based architecture (Experiment-1)	58
3.9	Confusion Matrices for binary classification, COVID-19 vs. Normal, using ResNet50 based architecture (Experiment-2)	59
3.10	Confusion Matrices for multiclass classification, COVID-19 vs. Normal vs. Pneumonia, using VGG16 based architecture (Experiment-3)	60
3.11	Confusion Matrices for multiclass classification, COVID-19 vs. Normal vs. Pneumonia, using ResNet50 based architecture (Experiment-4)	61
4.1	Data processing and model implementation flow for COVID-19 detection from cough sounds.	67
4.2	Samples of COVID-19 and Healthy cough sounds: (A) original recording, (B) voice detection, (C) segmentation.	72
4.3	MelSpectrogram of (a) COVID-19 segment, (b) Healthy segment	73
4.4	A fully connected neural network for detection of COVID-19 using spectral features from Cough.	75
4.5	ResNet50 based Deep Learning architecture for detection of COVID-19 using MelSpectrogram from Cough.	76
4.6	Vision Transformer (ViT) Architecture for detection of COVID-19 using MelSpectrogram from cough.	77
4.7	Model Summary of pre-trained ViT model employed for detection of COVID-19.	78
4.8	(a) Performance Learning Curves for Model-1 for COVID-19 Detection. (b) Corresponding Confusion Matrix.	79
4.9	(a) Performance Learning Curves for Model-2 for COVID-19 Detection. (b) Corresponding Confusion Matrix.	80
4.10	(a) Performance Learning Curves for Model-3 for COVID-19 Detection. (b) Corresponding Confusion Matrix.	80
5.1	Sample images from CIFAR-10, CIFAR-100, X-ray, and CT scan datasets.	95

5.2	Block diagram of the proposed method incorporating different pooling strategies, activation functions and probabilistic averaging for enhanced image classification.	96
5.3	Structural block diagram of the proposed architecture, detailing the sequence of operations. The diagram shows the flow from the convolutional operation, through the Positive and Negative ReLU activation functions, followed by block-wise concatenation, and finally the MaxPooling and MinPooling layers with interleaved concatenation.	97
5.4	A block diagram of the custom CNN model for classifying images from (a) the CIFAR-10 dataset and (b) the CIFAR-100 dataset. Conv2D refers to Convolutional Layer 2D followed by activation function, and BN stands for Batch Normalization.	99
5.5	A block diagram of the custom CNN model for classifying X-ray and CT scan images into three categories: COVID-19, Normal, and Pneumonia. Conv2D refers to Convolutional Layer 2D followed by activation function, and BN stands for Batch Normalization.	100
5.6	GradCAM analysis of the MaxMinPooling model on six CT scan images. Each row contains two pairs of images: the original CT scan (left) and the corresponding heatmap (right), highlighting the decision-making process of the model for effective pattern interpretation.	103
5.7	Comparison of model Accuracies for CIFAR-10, CIFAR-100, X-ray and CT scan datasets.	104
5.8	Performance learning curves of the proposed MaxMinPooling method for CIFAR-10, CIFAR-100, X-ray and CT scan datasets.	105
5.9	Confusion Matrix for CIFAR-10 image classification using the proposed MaxMinPooling method.	106
5.10	Confusion Matrices for detection of COVID-19 using the proposed MaxMinPooling method.	106
6.1	(a) A set of generalized Gaussian functions, and (b) the corresponding PR COLA windows obtained using (6.5) and (6.7).	115
6.2	(a) Fourier spectrum of multi-component synthetic FM signal $x(t)$ (6.24); TFR of the signal $x(t)$ based on (b) existing Gaussian window-based STFT, (c) STFT with proposed Gaussian COLA window, (d) STDCT with proposed Gaussian COLA window, (e) STFT with proposed Gaussian COLA window and GSO, (f) STDCT with proposed Gaussian COLA window and GSO, (g) EMD-HS, and (h) FSST.	116
6.3	(a) Fourier spectrum of multi-component synthetic AFM signal $y(t)$ (6.25); TFR of the signal $y(t)$ based on (b) existing Gaussian window-based STFT, (c) STFT with proposed Gaussian COLA window, (d) STDCT with proposed Gaussian COLA window, (e) STFT with proposed Gaussian COLA window and GSO, (f) STDCT with proposed Gaussian COLA window and GSO, (g) EMD-HS, and (h) FSST.	119

6.4	(a) A normal human heart sound signal; TFR of the heart sound signal based on (b) existing Gaussian window-based STFT, (c) STFT with proposed Gaussian COLA window, (d) STDCT with proposed Gaussian COLA window, (e) STFT with proposed Gaussian COLA window and GSO, (f) STDCT with proposed Gaussian COLA window and GSO, (g) EMD-HS, and (h) FSST.	120
6.5	Block diagram for COVID-19 detection model using COLA windows based STDCT and STFT.	121
6.6	Sample cough sounds from the training dataset. (a) COVID-19 (b) Healthy.	122
6.7	COLA windows based STDCT of a sample cough sound from the training dataset. (a) COVID-19 (b) Healthy.	122
6.8	COLA windows based STFT of a sample cough sound from the training dataset. (a) COVID-19 (b) Healthy.	123
6.9	ResNet50 model architecture for COVID-19 detection from cough sounds, utilizing COLA windows-based STDCT and STFT. The blue-highlighted area indicates that only the final convolutional block of ResNet50 was retrained.	123
6.10	Performance learning curves of the COVID-19 detection model trained using COLA windows based (a) STDCT images (b) STFT images.	124
6.11	Confusion matrices of the COVID-19 detection model trained using COLA windows based on (a) STDCT images and (b) STFT images.	125

ABBREVIATIONS

AI	Artificial Intelligence
ANN	Artificial Neural Network
CNN	Convolutional Neural Network
COVID-19	Corona Virus Disease 2019
CT	Computed Tomography
COLA	Constant-OverLap-Add
CXR	Chest X-Ray
DL	Deep Learning
EMD	Empirical Mode Decomposition
FCN	Fully Convolutional Network
FDM	Fourier Decomposition Method
FN	False Negative
FP	False Positive
FSST	Fourier SynchroSqueezed Transform
GAN	Generative Adversarial Networks
GPGPU	General Purpose Graphics Processing Unit
GNN	Graph Neural Networks
GRU	Gated Recurrent Unit
LLM	Large Language Model
LSTM	Long Short-Term Memory
MFCC	Mel Frequency Cepstral Coefficients
ML	Machine Learning
MRI	Magnetic Resonance Imaging
NLP	Natural Language Processing
ResNet	Residual Neural Network
RNN	Recurrent Neural Network
ROC	Receiver Operating Characteristic
RT-PCR	Reverse Transcription-Polymerase Chain Reaction
SARS	Severe Acute Respiratory Syndrome
STDCT	Short-Time Discrete Cosine Transform
STFT	Short-Time Fourier Transform
TFR	Time-Frequency Representation
TN	True Negative
TP	True Positive

VP	Viral Pneumonia
ViT	Vision Transformer
VGG	Visual Geometry Group
WHO	World Health Organization