Novel Deep Learning Architectures for Enhanced COVID-19 Detection

NARENDRA KUMAR MISHRA

DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY DELHI

APRIL 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 ऑग्मेंटेशन और फाइन–ट्यूनिंग तकनीकों द्वारा बेहतर किया गया है, और इन्हें स्ट्रैटिफाइड फाइव–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

List of Tables

List of Figures

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