

**Robustness and Imperceptibility Improvement of Transform  
Domain Digital Audio Watermarking using Neural Networks**

न्यूरल नेटवर्क की सहायता से फ्रीक्वेंसी डोमेन डिजिटल ऑडियो वाटरमार्किंग की  
मजबूती और अगोचरता में सुधार।

A Thesis

Submitted for the Award of the Ph.D. degree of  
**PACIFIC ACADEMY OF HIGHER EDUCATION AND  
RESEARCH UNIVERSITY**

By

**MR. PATIL ABHIJIT JAYPRAKASH**

श्री पाटील अभिजित जयप्रकाश

Under the Supervision of

**DR. DILENDRA HIRAN**

Professor

Faculty of Computer Science  
Pacific University, Udaipur

**DR. SHELKE RAMESH D.**

Associate Professor

Shivajirao, Jondhale College of Engineering,  
Dombivali (E) Thane



**FACULTY OF ENGINEERING**

**Department of Computer Engineering**

**PACIFIC ACADEMY OF HIGHER EDUCATION AND  
RESEARCH UNIVERSITY UDAIPUR**

**2024**

## DECLARATION

I, **PATIL ABHIJIT JAYPRAKASH**, S/o **SHRI PATIL JAYPRAKASH SHANTARAM** resident of B-302, Shree Mahavir C.H.S., Sector No. 5A, New Panvel (East), Panvel Raigad - 410206, Maharashtra, India, hereby declare that the research work incorporated in the present thesis entitled “**Robustness and Imperceptibility Improvement of Transform Domain Digital Audio Watermarking using Neural Networks**” is my own work and is original. This work (in-part or in full) has not been submitted to any University for the award of a Degree or a Diploma. I have properly acknowledged the material collected from secondary sources wherever required. I solely own the responsibility for the originality of the entire content.

Date :    /    /2024

**(Mr. Patil Abhijit Jayprakash)**

Signature of the Candidate

# CERTIFICATE

It gives me immense pleasure in certifying that the thesis entitled “ **Robustness and Imperceptibility Improvement of Transform Domain Digital Audio Watermarking using Neural Networks**” and submitted by **Mr. Patil Abhijit Jayprakash** is based on the work research carried out under my guidance. He has completed the following requirements as per Ph.D. regulations of the University;

- i. Course work as per University rules.
- ii. Residential requirements of the University.
- iii. Regularly presented Half Yearly Progress Report as prescribed by the University.
- iv. Published/ accepted minimum of two research paper in a refereed research journal.

I recommend the submission of thesis as prescribed/ notified by the University.

Date :    /    /2024

**DR. DILENDRA HIRAN**  
Professor  
Faculty of Computer Science  
Pacific University, Udaipur

# CERTIFICATE

It gives me immense pleasure in certifying that the thesis entitled “**Robustness and Imperceptibility Improvement of Transform Domain Digital Audio Watermarking using Neural Networks**” and submitted by **Mr. Patil Abhijit Jayprakash** is based on the work research carried out under my guidance. He has completed the following requirements as per Ph.D. regulations of the University;

- i. Course work as per University rules.
- ii. Residential requirements of the University.
- iii. Regularly presented Half Yearly Progress Report as prescribed by the University.
- iv. Published/ accepted minimum of two research paper in a refereed research journal.

I recommend the submission of thesis as prescribed/ notified by the University.

Date :    /    /2024

**DR. SHELKE RAMESH D.**

Associate Professor

Shivajirao Jondhale College of Engineering,  
Dombivali (E) Thane

# COPYRIGHT

I, **PATIL ABHIJIT JAYPRAKASH**, hereby declare that the Pacific Academy of Higher Education and Research University, Udaipur Rajasthan, shall have the right to preserve, use and disseminate the thesis entitled “**Robustness and Imperceptibility Improvement of Transform Domain Digital Audio Watermarking using Neural Networks**” in print or electronic format for academic research.

Date :    /    /2024

**(Mr. Patil Abhijit Jayprakash)**

Signature of the Candidate

## ACKNOWLEDGEMENT

The accomplishment of this thesis would never have come to light without the support of many people; First and foremost, I would like to thank **Almighty God** for giving me wit and wisdom. I am highly indebted to my Research Supervisor **Dr. Dilendra Hiran**, Professor, Faculty of Computer Science, Pacific University, Udaipur for being constant support, motivation and valuable inputs encouraged me to pursue with my research. I express my deep and sincere gratitude for his valuable guidance, immense help and time devotion at every step, without which this work would not have been possible. I would like to express my sincere gratitude to my Co-Supervisor **Dr. Ramesh Shelke**, for his continuous support and valuable guidance throughout my Ph.D. research work, I am thankful to him for his patience, motivation, enthusiasm, and immense knowledge in the subject. I would like to take this opportunity to say thanks to **Prof. Hemant Kothari**, Dean PG Studies, Pacific University, Udaipur for their valuable comment and continuous support during my research work. I am also thankful to **Dr. Suresh Ukarande**, Principal, KJSCOE for being the source of inspiration and motivation.

I would like to thank my family: my father **Late. Mr. Jayprakash S. Patil** and my loving mother **Mrs. Sunita J. Patil**, for their encouragement and always supporting me spiritually throughout my life. I am heartily thankful to my father who was the first person to motivate me to undertake research work. My special thanks to my wife **Mrs. Pallavi** and my daughters for their support and help in all ways I needed. I take this opportunity to thank all my friends and colleagues for their continuous motivation and well wishes.

(**Mr. Patil Abhijit Jayprakash**)

## ABSTRACT

This research introduces an advanced digital audio watermarking system designed to bolster the security of digital data, specifically focusing on ownership and copyright protection. Conventional audio watermarking methods often encounter limitations and reliability issues against various attacks. To address these challenges, a novel approach is proposed, leveraging deep learning and integrating the Discrete Wavelet Transform (DWT) with an optimized deep Convolutional neural network (DCNN). The primary contribution lies in the DCNN's adeptness in selecting optimal embedding locations, a critical factor for robust watermarking. Through hyper parameter tuning and search location optimization, classifier errors are minimized. Experimental results showcase superior performance, with the proposed model achieving a Bit Error Rate (BER) of 0.082, Mean Square Error (MSE) of 0.099, and Signal-to-noise ratio (SNR) of 45.363. This model surpasses the existing watermarking models and underscores the effectiveness of neural network architectures, particularly the DCNN, in optimizing watermark embedding and extraction with minimal bit error. The research also explores various hybrid and innovative techniques in digital audio watermarking, highlighting the utility of neural networks in advancing the field. The discussion encompasses classical and modern approaches incorporating machine learning, deep learning, bio-inspired algorithms, and cryptographic methods to enhance watermarking efficiency and security. The research provides insights into the challenges and opportunities in achieving robustness, imperceptibility, and security in digital audio watermarking, ultimately showcasing the applicability of neural network architectures in this domain.

# TABLE OF CONTENTS

Chapter No.	Section No.	TITLE	PAGE NO
		<i>Title Page</i>	<i>i</i>
		<i>Declaration</i>	<i>ii</i>
		<i>Certificate</i>	<i>iii</i>
		<i>Copyright</i>	<i>V</i>
		<i>Acknowledgement</i>	<i>vi</i>
		<i>Contents</i>	<i>viii</i>
		<i>List of Tables</i>	<i>xi</i>
		<i>List of Figures</i>	<i>xii</i>
		<i>Abbreviation</i>	<i>xiii</i>
<b>1</b>	<b>INTRODUCTION</b>		<b>01 - 16</b>
	1.1	Watermarking Process	03
	1.2	Audio Watermarking Techniques	04
	1.3	Requirements of watermarking techniques	05
	1.4	Background and Motivation	06
	1.5	Types of Watermarks (Spatial, Frequency, Temporal)	08
	1.6	Overview of Embedding and Extraction Processes	09
	1.7	Deep Learning Approaches	10
	1.8	Machine Learning for Improved Robustness	13
	1.9	Challenges	13
	1.10	Research Objectives	14
	1.11	Problem Formulation	15
	1.12	Contributions of the thesis	15
	1.13	Thesis Organization	16

<b>2</b>	<b>LITERATURE REVIEW</b>		<b>17 - 34</b>
	2.1	Introduction	17
	2.2	Review of Literatures	17
	2.3	Analysis and Discussion	21
	2.4	Research gaps	33
	2.5	Summary	34
<b>3</b>	<b>AUDIO WATERMARKING USING BASIC NEURAL NETWORK AND DEEP LEARNING TECHNIQUES</b>		<b>35 - 79</b>
	3.1	Chapter Introduction	35
	3.2	Basic Neural Network Model for audio watermarking	37
	3.3	DCNN with SLOA: Deep learning model for Audio Watermarking	44
	3.4	Existing Deep learning and optimization Models: Discussion and comparison.	52
	3.5	Results and Discussion	65
	3.6	Summary	79
<b>4</b>	<b>DATA ANALYSIS</b>		<b>80 - 93</b>
	4.1	Chapter introduction	80
	4.2	Performance evaluation metrics	80
	4.3	Comparative methods	82
	4.4	Comparative discussion	87
	4.5	Achievements of the research	92
	4.6	Chapter conclusion	93
<b>5</b>	<b>APPLICATIONS</b>		<b>94 - 97</b>
	5.1	Chapter introduction	94

	5.2	Applications of Audio Watermarking	94
	5.3	Summary	97
<b>6</b>	<b>CONCLUSION AND FUTURE WORK</b>		<b>98 - 99</b>
	6.1	Chapter Introduction	98
	6.2	Conclusion and research significance	98
	6.3	Future Scope	99
	<b>REFERENCES</b>		<b>100 - 109</b>
	<b>Appendix:</b>		
	Appendix-I	<b>Published Research Paper I</b>	-
		<b>Published Research Paper II</b>	
		<b>Published Research Paper III</b>	
		<b>Published Research Paper IV</b>	
	Appendix-II	<b>Conference Certificates</b>	-
	Appendix-III	<b>Plagiarism Checking Report</b>	-

## LIST OF TABLES

<b>Table No.</b>	<b>Title</b>	<b>Page No.</b>
2.1	Summary of the related works	22
2.2	Analysis based on different methods	30
3.1	Layer information of deep CNN	55
3.2	Analysis based of different signals performance for Image-1	72
3.3	Analysis based of different signals performance for Image-2	76
4.1	Based on Image-1 comparative analysis for different Signals	88
4.2	Based on Image-2 comparative analysis for different Signals	90
4.3	DCNN-SLOA model time complexity analysis	92

## LIST OF FIGURES

Figure No.	Title	Page No.
1.1	Watermark embedding process	03
1.2	Watermark extracting process	04
1.3	Requirements of digital audio watermarking techniques	06
2.1	Taxonomy of the Digital Audio Watermarking techniques	17
2.2	Analysis based on Performance metrics	33
3.1	Illustrative representation of the working of NN in watermarking	38
3.2	Layer details for the NN model	43
3.3	Flow architecture of the proposed digital audio watermarking model	46
3.4	The secret data embedding process	50
3.5	Layer details for the LSTM	54
3.6	Architecture of the deep CNN layer	55
3.7	Experimental result obtained using NN model a) Audio Signal before applying DWT, b) Audio Signal after applying DWT c) Sample Watermark Image and d) Extracted Watermark Image	66
3.8	Experimental result obtained using the deep CNN with SLOA.	70
3.9	Graph representing the model training for 100000 epochs and reduction in loss for Basic NN Model.	70
3.10	Performance analysis based on Image-1 a) MSE, b) BER, c) SNR	74
3.11	Performance analysis based on Image-2 a) MSE, b) BER, c) SNR	78
4.1	Comparative analysis for input Image-1 a) MSE b) BER c) SNR	85
4.2	Comparative analysis for input Image-2 a) MSE b) BER c) SNR	87

## LIST OF ABBREVIATIONS

DWT	:	Discrete Wavelet Transform
DCNN	:	Deep Convolutional Neural Network
BER	:	Error Rate
MSE	:	Mean Square Error
SNR	:	Signal-To-Noise Ratio
DL	:	Deep Learning
HAS	:	Human Auditory System
LSB	:	Least Significant Bit
DFT	:	Discrete Fourier Transform
DCT	:	Discrete Cosine Transform
SVD	:	Singular Value Decomposition
QIM	:	Quantization Index Modulation
SLOA	:	Search Location Optimization Algorithm
DWT	:	Discrete Wavelet Transform
FRT	:	Fractional Fourier Transform
IWT	:	Integer Wavelet Transform
FPP	:	False Positive Problem
KNN	:	K-Nearest Neighbor
SVM	:	Support Vector Machine
FFT	:	Fast Fourier Transform
DTCWT	:	Dual-Tree Complex Wavelet Transforms
STFT	:	Sort-Time Fourier Transform
RFPS	:	Robust Feature Point Scheme Algorithm

AES	:	Advanced Encryption Standard
AVNM	:	Adaptive Vector Norm Modulation
GA	:	Genetic Algorithm
MME	:	Meet-In-The-Middle Embedding
PSHT	:	Parametric Slant-Hadamard Transform
HD	:	Hessenberg Decomposition
PEAQ	:	Perceptual Evaluation Of The Audio Quality
BPNN	:	Back Propagation Neural Network
FCNN	:	Fully Connected Neural Network
LGG	:	Lower-Grade Glioma
HH	:	High-High
HL	:	High-Low
LH	:	Low-High
LL	:	Low-Low
RNNs	:	Recurrent Neural Networks
HMS	:	Human Mental Search
HSO	:	Hybrid Swarm Optimization
PAR	:	Pitch Adjusting Rate
ODG	:	Objective Difference Grade
MSE	:	Mean Square Error
BER	:	Bit Error Rate
SNR	:	Signal-To-Noise Ratio