CHAPTER - 1

INTRODUCTION

1. Overview:

With the accelerated development of digital technology over the past years, the utilization of computer networks has altered the way of communication all over the world. Networks such as wireless sensor networks are growing exponentially with the increased demand of users. Additionally, underwater network is a most useful communication networks that utilizes frequent sound waves [22, 24–26, 49]. Generally, the utilization of communication networks and multimedia information in digital format, provides enormous opportunity for development and innovation, meanwhile creates multiple challenges to the researchers. The enhanced internet access and data transfer with less error have enabled the users to distribute multimedia files across multiple individuals [23,32].However, the open access of Internet creates the problems of illegal distribution of privately owned audio and multimedia products [1], which insists on the need for privacy preserving arising as an important concern in the multimedia networks. Several approaches are innovated for preserving the private data and other multimedia products. Initially, the cryptographic approaches are introduced for preserving the multimedia files of the legitimate users. Encryption techniques attempts to encode data into a format that allows only the legal users to receive the files. But the method does not ensure the ownership of the data, after the data are decoded into the new format. Hence, Digital watermarking is formulated which the process of embedding information into a signal including audio, video or pictures in a way that is complex to remove the ownership information of the data. If the signal is copied, then the ownership information is also carried in the copy of the file. Through this information, the security of digital product can be enabled against unauthorized copying, false property claims, non-authorized distribution, and so on [31, 33, 34, 40][4].

Digital audio watermarking is the process that embeds a series of hidden data in host audio files, as it cannot be audible and has a good robustness to signal processing attacks. Digital watermarking can be effectively employed to limit the scope of counterfeiting and piracy of audio files. The capability is due to crucial features such as imperceptibility, inseparability of the content from the watermark, and intrinsic ability to undergo same transformation experienced by the host audio signal [6]. However, watermark detection in audio files is simpler than image files that make audio watermarking to be more complicated than other areas of watermarking. Since, the human hearing system has more sensitivity than the human vision system [2], protection of imperceptibility in audio watermarking is more difficult than image watermarking [4,5][3]. Effective audio watermarking techniques must satisfies the imperceptibility, security, robustness, and high payload size requirements [22, 48]. The robustness of watermarked audio signal refers to the ability to extract the embedded watermark from the audio signal after applying different types of attacks [38][8]. The payload size of watermarked audio refers to the amount of data that can be embedded per unit of time into the audio signal without losing imperceptibility, which must be more than 20 bps [21]. However, maintaining these requirements in an audio watermarking scheme is a major challenge as the methods require trade-off between the required specifications [12]. Additionally, increasing the number of embedded in the audio signals would reduce the quality of the watermarked signal and robustness of the technique [5]. Deep Learning (DL) plays an significant role in numerous critical applications such as classification, autonomous vehicles, voice recognition and so on. Additionally, DL is employed mostly for security [59] as well as can be utilized within several types of data, like text [36], images [25], audio [10] and video [23]. Notably, Deep Convolutional Neural Networks (DCNN) such as AlexNet, GoogleNet, VGGNet, and ResNet etc, [7.8,37,47,86,87]exhibited a remarkable performance for solving computer vision and other applications problems [7].Traditional digital audio watermark techniques embeds watermarks in the time domain, frequency domain, and coded domains to insert information where the human auditory system (HAS) does not recognize [7]. Some audio watermark techniques insert the information in the frequency domain, and other techniques utilize the magnitude component of the time-frequency representation as input features for DL models [8, 9]. The recent techniques hide information resiliently as the audio file is deformed via MP3 conversion, air propagation, noise insertion, and so on but must be represented as a certain form of time-frequency representation in the process of extracting embedded data from the hidden watermark [2].

1.1. Watermarking Process:

Watermarking process consists of two blocks where the watermarking information is inserted into the original signal that should not degrade the quality of the original signal. The Watermarking process involves two blocks involving the watermark embedding as well as watermarking extraction [12].

i) Watermark Embedding

The watermark embedding process utilizes the watermarking key that protects the watermarking information as well as data from unauthorized users from manipulating the data. The water embedding process, in which the embedded object is referred to as watermark and the original signal acts as the watermark embedding medium and the modified object is termed as embedded signal or watermarked data. The watermark embedding block is depicted in Figure 1.1 which comprises watermark, original signal or cover object, and watermarking key as the inputs that generates the watermarked data or the embedded signal as the output of the process.

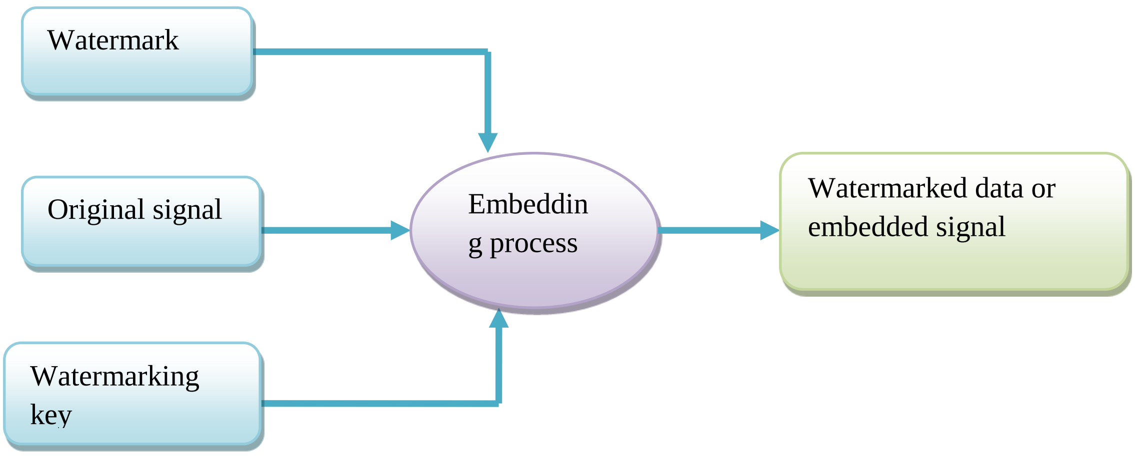

Figure 1.1: Watermark embedding process

ii) Watermark Extraction

Watermark extraction is the reverse process of the watermarking embedding, in which the embedded signal with the encrypted key is decomposed utilizing the wavelet transform employed in the embedding process. The watermark extraction process is depicted in Figure 1.2.

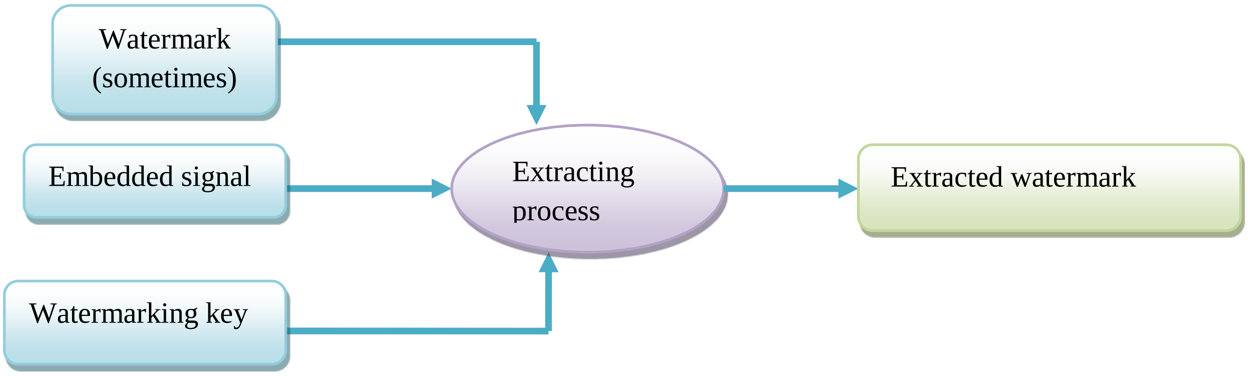

Figure 1.2: Watermark extracting process

1.2. Audio Watermarking Techniques:

The main purpose for digital audio watermarking is based on the human auditory system (HAS) as the human ear can perceive the power range more than 109:1 and the range frequencies 103:1. Additionally, the human ear can hear the low Gaussian noise of 70 dB and the useful features such as the louder sounds mask the corresponding slow sounds. Hence, the features are employed to embed the watermark information to the original signal. The watermarking algorithms based on the properties of frequency masking and temporal masking is utilized for protecting the original data [12]. Audio watermarking algorithms are categorized into time domain techniques, frequency domain techniques and hybrid techniques.

1.2.1 Time domain Techniques:

Watermarking is utilized over the original samples of the audio signal. Most commonly utilized time domain watermarking technique is the least significant bit (LSB) method. The watermark information is embedded within the least significant bits of the original signal [9][12]. Additionally, the robustness of the method relies on the number of bits that are replaced in the host signal. However the time domain techniques based watermarking will not offer security against the attacks.

1.2.2. Frequency Domain Technique

Frequency domain watermarking embeds the watermarks over the spectral coefficients of the original signal. Most commonly utilized transforms includes the Discrete Cosine Transform (DCT), Discrete Fourier Transform (DFT), and Discrete Wavelet Transform (DWT), Discrete Cosine form transform (DCT), Discrete Fourier transform (DFT)[10]. These techniques offers the merits of masking of different tones of HAS for better watermarking. The original audio signal is converted into frequency domain by utilizing the different transform and watermark is replaced in the magnitude or phase response of the signal. Then the audio watermark is applied with inverse transform for acquiring the watermarked domain [12].

1.2.3. Hybrid techniques:

The hybrid digital audio watermarking technique applies two or more transform methods such as DCT, DFT, DWT, and singular value decomposition (SVD)[9] over the host signal for acquiring the watermarked or embedded signal.

1.3. Requirements of watermarking techniques:

The watermarking techniques applied over the original signal should satisfy the requirements of imperceptibility, robustness, security as well as capacity. However, maintaining all the requirements in audio watermarking is a major challenge as the methods require trade-off between them.

- ***Imperceptibility***: The quality of the original signal should not get degraded with the watermark[13]
- ***Robustness:*** The watermarked data should remain resilient and not get eliminated by the illegal distributors.
- ***Security:*** The watermark should be detected only by the authorized user of the data.
- ***Capacity:*** The amount of bits that can be embedded or inserted with the original signal.

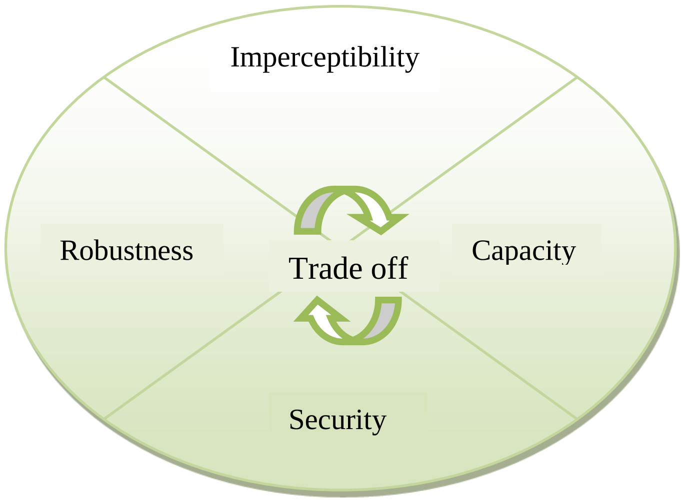

Figure 1.3: Requirements of digital audio watermarking techniques

1.4 Background and Motivation:

The rapid proliferation of digital multimedia content in the modern age has given rise to unprecedented challenges in protecting intellectual property and ensuring the integrity of digital assets [50]. With the ease of content distribution over digital channels, concerns about unauthorized use, piracy, and data tampering have become pressing issues [51-54]. This backdrop underscores the need for innovative and robust solutions, such as digital audio watermarking, to address these challenges. The motivation behind exploring and developing audio watermarking technologies lies in their potential to embed imperceptible yet identifiable signatures into audio files, providing a means to authenticate, track, and safeguard digital content in an interconnected and content-rich digital landscape.

1.4.1 Significance in the Digital Era:

Digital audio watermarking holds immense significance in the contemporary digital landscape. As a technology designed to embed invisible markers within audio files, it offers a powerful tool for various applications. Its primary significance lies in the protection of intellectual property, enabling content creators and distributors to assert ownership and control over their digital assets [55-57]. Additionally, audio watermarking plays a pivotal role in content authentication, broadcast monitoring, forensic analysis, and ensuring the integrity of audio data in diverse domains. Its adaptability and versatility make it a critical component in the broader effort to mitigate the risks associated with the digital distribution of multimedia content [58].

1.4.2 Key Challenges in Multimedia Content Protection:

The surge in multimedia content creation and consumption has brought forth a host of challenges in safeguarding digital assets. Challenges include unauthorized duplication, distribution, and manipulation of digital content. Ensuring the authenticity of digital files in the face of evolving technological threats and the prevalence of sophisticated attacks poses a significant hurdle. These challenges necessitate advanced techniques, such as digital audio watermarking, to provide a robust layer of security and authentication [59]. Addressing these challenges is crucial for preserving the rights of content creators and maintaining the integrity of digital content in an increasingly interconnected and digital-centric environment [60].

1.4.3 Fundamentals of Audio Watermarking:

Audio watermarking involves the process of embedding imperceptible and unique identifiers, or watermarks, into audio signals without compromising the quality of the original content. The fundamentals revolve around understanding the types of watermarks, which can be embedded in the spatial, frequency, or temporal domains [61-63]. The choice of domain depends on the specific requirements of the application. Techniques employed in audio watermarking include modifying the amplitude, phase, or frequency of the audio signal in a way that is subtle to the human ear but detectable by algorithms during extraction. The goal is to strike a delicate balance between robustness, transparency, and resistance to various signal processing attacks. The fundamentals of audio watermarking form the basis for exploring its applications, challenges, and advancements in the digital realm.

1.4.4 Definition and Purpose:

Digital audio watermarking involves the incorporation of imperceptible and unique identifiers, known as watermarks, into audio signals. The primary purpose is to provide a means for authentication, tracking, and protection of intellectual property in the digital domain [64]. These watermarks serve as digital signatures, allowing content creators and distributors to assert ownership, trace the origin of the content, and verify its integrity. The embedding process is designed to be subtle to the human ear while remaining detectable through specialized algorithms during extraction, ensuring that the watermark does not compromise the quality of the original audio.

1.4.5 Characteristics of a Robust Watermark:

A robust watermark exhibits certain key characteristics to effectively fulfill its purpose in various applications. Robust watermarks are imperceptible to the human senses, ensuring that they do not degrade the quality of the audio signal, they are resistant to common signal processing operations and attacks making them durable in the face of potential alterations or malicious tampering. Additionally, a robust watermark should be able to withstand compression and decompression processes without significant loss of information [65]. Achieving a balance between robustness, imperceptibility, and resistance to attacks is crucial in designing an effective watermark for digital audio [66].

1.5 Types of Watermarks (Spatial, Frequency, Temporal):

The choice of domain for embedding watermarks depends on the specific requirements of the application. Spatial domain watermarking involves modifying the amplitude or phase of the audio signal directly [67]. Frequency domain watermarking alters the characteristics of the signal in the frequency domain, often by manipulating the Fourier transform coefficients. Temporal domain watermarking involves modifications in the time domain, such as slight shifts in timing or amplitude variations over time. Each type has its advantages and trade-offs, and the selection depends on factors like the desired level of robustness, transparency, and resistance to different types of attacks.

1.5.1 Techniques and Algorithms:

Various techniques and algorithms are employed in the process of digital audio watermarking. These can include [72]:

i) Frequency Domain Techniques: Modifying frequency components of the audio signal, often by manipulating Fourier transform coefficients.

ii) Time Domain Techniques: Introducing subtle changes in the time domain, such as amplitude variations or temporal shifts.

iii) Spread Spectrum Techniques: Spreading the watermark across the entire frequency spectrum to enhance robustness.

iv) Quantization Index Modulation (DTCWT**):** Modulating quantization indices to embed information imperceptibly.

Additionally, advanced approaches may involve the integration of machine learning and deep learning algorithms for more sophisticated and adaptive watermarking. The choice of technique depends on factors such as the desired level of robustness, transparency, and resistance to various signal processing attacks.

1.6 Overview of Embedding and Extraction Processes:

The embedding and extraction processes in digital audio watermarking involve two fundamental steps. During embedding, imperceptible watermarks are introduced into the audio signal using specific techniques and algorithms [73-75]. This process is designed to be transparent to the human ear, ensuring minimal impact on the perceived audio quality. In contrast, the extraction process involves recovering the embedded watermark from the watermarked audio signal. Specialized algorithms are employed to detect and decipher the watermark without compromising the original audio content. These processes are critical for the successful integration and retrieval of watermarks in a variety of applications.

i) Frequency Domain Techniques: Frequency domain techniques in audio watermarking involve manipulating the frequency components of the audio signal. This may include altering the amplitude or phase of specific frequency bands, often achieved through techniques such as discrete Fourier transform (DFT) or discrete cosine transform (DCT). Frequency domain methods aim to embed watermarks in a manner that is robust against common signal processing operations while maintaining imperceptibility [76].

ii) Time Domain Techniques: Time domain techniques focus on introducing subtle changes in the amplitude or timing of the audio signal. This may involve slight variations in amplitude over time or temporal shifts in the signal, time domain watermarking aims to achieve imperceptibility while providing resistance to certain types of signal processing attacks. Common approaches include amplitude modulation or the addition of small delays [77].

iii) Transform Domain Techniques: Transform domain techniques leverage mathematical transformations, such as wavelet transforms, to embed watermarks into specific domains. By modifying coefficients in the transformed domain, these techniques enhance the robustness and imperceptibility of the watermark [78]. Transform domain methods are known for their adaptability to different types of audio signals and resistance to various attacks.

1.7 Deep Learning Approaches:

Deep learning approaches in audio watermarking involve the use of neural networks to learn and embed watermarks directly. These approaches leverage the power of deep neural networks to automatically discover features and patterns for embedding imperceptible watermarks. Deep learning models can adapt to complex relationships within the audio data, potentially enhancing both robustness and imperceptibility [79].

1.7.1 Robustness and Security Considerations:

Ensuring the robustness of digital audio watermarking involves making the embedded watermark resilient to common signal processing operations, compression, and potential attacks. Robust techniques are designed to withstand alterations to the watermarked signal while allowing for reliable extraction of the watermark. Security considerations encompass protecting the watermark against malicious attacks and unauthorized removal attempts. Achieving a balance between robustness and security is crucial to the effectiveness of audio watermarking in real-world scenarios, ensuring that the embedded information remains intact and trustworthy even in the face of intentional tampering or unintentional signal processing [68].

1.7.2 Resistance to Signal Processing Attacks:

The effectiveness of digital audio watermarking hinges on its resistance to various signal processing attacks that may be employed intentionally to manipulate or remove the embedded watermark. Robust watermarking techniques are designed to withstand common attacks such as compression, equalization, filtering, and other signal alterations. Techniques employed in resistant watermarking aim to ensure that the embedded information remains intact and recoverable even after exposure to these intentional attacks, maintaining the integrity of the watermark and its associated information [69].

1.7.3 Sensitivity to Non-Malicious Operations:

While resistance to signal processing attacks is crucial, a well-balanced digital audio watermarking system should also consider its sensitivity to non-malicious operations. Non-malicious operations include routine signal processing tasks, like compression for storage or transmission [70]. Sensitivity to these operations ensures that the watermark remains imperceptible to the human ear and does not degrade the audio quality during normal, everyday use. Striking the right balance between robustness against attacks and sensitivity to non-malicious operations is essential for the practical deployment of audio watermarking systems.

1.7.4 Trade-off between Robustness and Transparency:

A fundamental challenge in digital audio watermarking lies in achieving an optimal trade-off between robustness and transparency. Robust watermarks are often more detectable and resistant to attacks, but this may come at the expense of transparency, meaning they might introduce perceptible artifacts in the audio signal. Striking the right balance involves finding methods that embed watermarks robustly while minimizing their impact on the perceived audio quality [71]. This trade-off is crucial to ensure that the watermarking process does not compromise the user experience while providing effective protection and authentication.

1.7.5 Security Measures in Watermark Embedding:

To enhance the security of digital audio watermarking, various measures can be implemented during the embedding process. These measures include encryption techniques to secure the embedded watermark, ensuring that it is not easily decipherable by unauthorized parties. Additionally, secret key-based methods can be employed to add an extra layer of security, requiring a specific key for successful watermark extraction [20]. These security measures contribute to safeguarding the integrity of the embedded information and protecting against unauthorized attempts to tamper with or remove the watermark.

1.7.6 Optimizations and Enhancements:

Continual research and development efforts focus on optimizing and enhancing digital audio watermarking techniques. Optimizations may involve improving the efficiency of watermark embedding and extraction processes to reduce computational demands. Enhancements may include the integration of advanced algorithms, machine learning, or deep learning approaches to adaptively enhance the robustness and imperceptibility of watermarks. Ongoing optimizations and enhancements aim to address emerging challenges, enhance the overall performance of audio watermarking systems, and ensure their relevance in the ever-evolving landscape of digital content protection [21].

1.7.7 Role of Optimization Algorithms:

Optimization algorithms play a crucial role in the effectiveness of digital audio watermarking. These algorithms are utilized to optimize parameters during the embedding and extraction processes, ensuring that the imperceptibility and robustness of the watermark are enhanced. Optimization algorithms, such as genetic algorithms, particle swarm optimization, or simulated annealing, contribute to finding optimal solutions in the vast solution space, improving the overall performance of audio watermarking systems [14]. The use of these algorithms aids in fine-tuning parameters to achieve an optimal balance between robustness, imperceptibility, and resistance to attacks.

1.7.8 Time Complexity Considerations:

Time complexity is a critical aspect of digital audio watermarking systems, especially in real-time or near-real-time applications. The efficiency of embedding and extraction processes directly impacts the practicality and usability of the watermarking solution. Time complexity considerations involve optimizing algorithms and techniques to minimize computational demands while maintaining high performance. Advances in algorithmic efficiency and parallel processing contribute to reducing time complexity, ensuring that audio watermarking can be applied seamlessly in various applications without significant computational overhead.

1.7.9 Advances in Signal Processing:

Advances in signal processing techniques have a direct impact on the performance and capabilities of digital audio watermarking. Innovations in signal processing contribute to the development of more sophisticated and robust embedding and extraction methods. Techniques such as adaptive filtering, wavelet transforms, and statistical signal processing enhance the overall quality of watermarking processes. These advances enable audio watermarking systems to adapt to different types of audio signals, improving their resilience against signal processing attacks and ensuring effectiveness across diverse applications.

1.8 Machine Learning for Improved Robustness:

Machine learning techniques, including supervised and unsupervised learning, are increasingly integrated into digital audio watermarking for enhanced robustness. Machine learning models can learn patterns and characteristics of audio signals, enabling adaptive embedding strategies that optimize imperceptibility and resistance to attacks. By leveraging machine learning as well as deep learning, audio watermarking systems can become more adaptive and resilient, learning from data and improving their performance over time. This integration contributes to the development of intelligent watermarking solutions capable of addressing dynamic challenges in the digital landscape.

1.9. Challenges

- Robustness, imperceptibility and watermark embedding capacity are the major factor in the digital audio watermarking technique. However, attaining all at the same time is a major challenge as the methods require trade-off between the required specifications [12].

- Additionally, maintaining the payload size which is the amount of data that can be embedded per unit of time into the audio signal without losing imperceptibility, which must be more than 20 bps, is a complex task [21].

- Increasing the amount of embedded in the audio signals would minimize the quality of the watermarked signal and robustness of the technique [5].

- Audio watermarking found to be more complicated as the human hearing system was more sensitive than the human vision system that makes protection of imperceptibility in audio watermarking to be more complex than image watermarking [3].

- It was very difficult to achieve better robustness against challenging signal processing attacks in spatial domain [81].

- Although it is simple to embed the audio signal utilizing the spatial domain and is computationally reasonable but attaining good imperceptibility at acceptable payload is a challenging task [81].

- Most of the conventional techniques could not resist synchronization attacks, such as jittering and time scale modification (TSM), because it lacked synchronization mechanism and the attacks including cropping are still a big challenge under high payload [80].

1.10 Research Objectives

- To effectively perform the audio watermarking utilizing the optimized deep CNN and DWT for attaining the high efficiency.

- To enhance the performance via Search Location Optimization Algorithm (SLOA) that tunes the parameters to attain the optimal solution for attaining the robust watermarking.

- To employ the DWT as a standard preprocessing technique to extract the different frequency components and to enhance the security of watermarking.

- To reduce the computational complexity of the model utilizing the SLOA that enhanced the speed and the efficiency of the audio watermarking.

1.11 Problem Formulation:

1. What are the recent methods involved in the audio water marking?

2. What are the challenges that exist in the recent methods of audio watermarking?

3. What is the performance metrics involved in the audio watermarking?

The formulated audio watermarking technique overcame the challenges in the existing techniques utilizing the deep learning and combines the Discrete Wavelet Transform (DWT) with an optimized deep Convolution neural network (DCNN). The significance of the contribution relies on the DCNN's capability in selecting optimal embedding locations. The hyper parameter tuning is performed with the search location optimization, which minimizes the errors. The proposed model exceeded the existing watermarking techniques via optimizing the watermark embedding and extraction with minimal bit error and reduced the time complexity of watermarking.

1.12 Contributions of the thesis:

The contribution of this research involves integrating the Discrete Wavelet Transform (DWT) with an optimized deep Convolutional neural network (DCNN) for robust audio watermarking. DWT is employed for the effective preprocessing and feature extraction that decomposes audio signal into different frequency components through a series of wavelet transformations. The significance of the contribution lies in the DCNN's ability in selecting optimal embedding locations, utilizing the SLOA. The optimization is formulated by the searching behavior of creatures to enhance the speed of the search utilizing the locating characteristics. Further, the SLOA tunes the hyper parameter and minimizes the classifier errors of the deep CNN and offers the effective audio watermarking.

1.13 Thesis Organization:

The basic organization of the thesis is follow as, the introduction comprising the overview of the digital audio watermarking techniques, their challenges, research objectives, problem formulation, and thesis contribution is detailed explained in chapter 1. The literature survey of the recent watermarking techniques involving the taxonomy, research gaps, motivation and conclusion are briefly interpreted in chapter 2. The optimized Deep Neural Network classifier for the effective audio watermarking is described in chapter 3. In chapter 4 the detailed comparative discussion of the contribution and the performance evaluation was elaborated. The chapter 5 elucidates the applications of the audio watermarking .Finally, the conclusion and the future work of the audio watermarking is discussed in the chapter 6.