Audio Watermarking: A Comprehensive Review

Mohammad Shorif Uddin[1], Ohidujjaman[2], Mahmudul Hasan[3], Tetsuya Shimamura[4]
Computer Science and Engineering, Jahangirnagar University, Savar 1342, Bangladesh[1]
Graduate School of Science and Engineering, Saitama University, Saitama 338-8570, Japan[1,2,4]
Computer Science and Engineering, Daffodil International University, Savar, Dhaka 1216, Bangladesh[2]
Computer Science and Engineering, Comilla University, Comilla 3506, Bangladesh[3]

***Abstract*—Audio watermarking has emerged as a potent technology for copyright protection, content authentication, content monitoring, and tracking in the digital age. This paper offers a comprehensive exploration of audio watermarking principles, techniques, applications, and challenges. Initially, it presents the fundamental concepts of digital watermarking, elucidating its key characteristics and functionalities. After that, different audio watermarking methods in both the time and transform domains are explained, such as feature-based, parametric, and spread-spectrum methods, along with how they work, and their pros and cons. The paper further addresses critical challenges in maintaining key criteria such as imperceptibility, robustness, and payload capacity associated with audio watermarking. Additionally, it examines watermarking evaluation metrics, datasets, and performance findings under diverse signal-processing attacks. Finally, the review concludes by discussing future directions in audio watermarking research, emphasizing advancements in deep learning-based approaches and emerging applications.**

***Keywords*—Audio watermarking; deep learning approach; spread-spectrum method; signal-processing attacks; time domain; transform domain**

I. Introduction

The proliferation of digital audio content has revolutionized the way we consume music, cinema, podcasts, and audiobooks. However, this ease of access has also fueled copyright infringement and unauthorized distribution. Audio watermarking [1-6] has emerged as a robust solution to address these concerns. The first works on digital audio watermarking were reported in references [7,8]. It involves imperceptibly embedding a unique audio identifier, called a watermark, into a host audio signal. This watermark can be extracted later to verify the content's authenticity, identify ownership, and track or monitor its distribution in digital rights management.

A huge amount of research work has been carried out on digital audio watermarking techniques in the last three decades, hence, the field has matured enough. Sophisticated signal processing techniques were widely utilized to develop numerous audio watermarking techniques in both time and transform domains [9-52], each having its own distinct benefits and boundaries. Fig. 1 shows a generic digital audio watermarking system where signal manipulations are carried out in the watermark embedding (encoding) and extraction (decoding) process. Watermarked signals frequently face various attacks [48-52] aimed at destroying or removing the watermarks by intentional attackers with bad motives. Besides, some users are treated as unintentional users since they may distort the image during signal compression, equalization, and effects addition without any bad motive. For this reason, the effectiveness of an audio watermarking technique is very important and its effectiveness is mainly evaluated based on five criteria: (i) imperceptibility, which indicates that the watermarked signal should be the same as the host audio concerning auditable quality (ii) robustness that upholds the unalteredness of the watermark after experiencing any attack by the unauthorized users, (iii) security that confirms the watermark signal should be secured from tampering, distortion and forging, (iv) capacity that ensures the increased number of watermarks embedded in the audio signal per unit time, and (v) computational complexity confirms the computational simplicity of the watermarking algorithm. Among these five criteria, the first two, imperceptibility and robustness, are the fundamental issues in evaluating the performance of a watermarking algorithm.

Some review works [53-62] have also been reported on audio watermarking. However, these are not sufficient, as many things, such as benchmarks, methodologies, datasets, and evaluation metrics, are not sufficiently described for a comprehensive knowledge of this domain. This review article provides a comprehensive overview of audio watermarking, encompassing its principles, techniques, applications, and challenges. We aim to equip readers with a thorough understanding of this vital technology and its role in safeguarding digital audio content. The major contributions of this research are as follows:

- Summarizes the existing audio watermarking methods in different categories
- Explains the datasets and evaluation metrics
- Compares and investigates the performance of various audio watermarking algorithms to find out the state-of-the-art
- Point out the challenges that must be addressed by future researchers.

The remainder of the paper is organized as follows. Section 2 presents the basic concept of the audio watermarking method. Section 3 highlights the requirement issues of audio watermarking along with performance evaluation metrics. Section 4 describes a survey of methodologies of different audio watermarking algorithms along with the state-of-the-art audio watermarking approaches. Section 5 shows the directions for future research for further improvements. Finally, Section 6 concludes the paper.

II. Concept of the Audio Watermarking

An audio watermark is a unique identifier embedded in an audio signal that is used to prove the ownership or copyright of the audio document. Therefore, audio watermarking is a

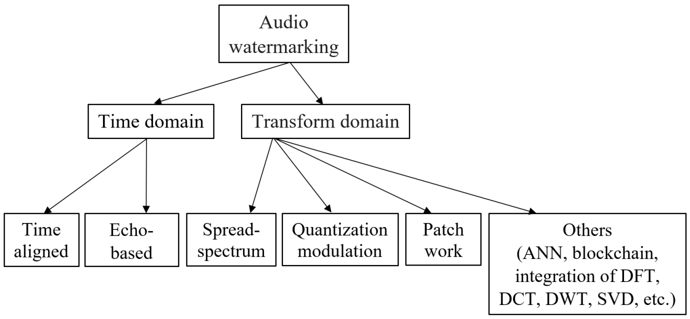

Fig. 1. Audio watermarking categorization.

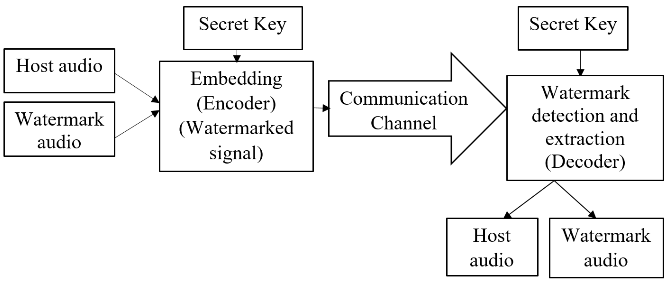

Fig. 2. A generic digital audio watermarking system.

process of embedding information into an audio signal in a way that is difficult to remove or tamper. Hence, watermarking has become increasingly important to enable copyright protection and ownership verification. In the last 30 years, many different watermarking methods have been created. These methods can be put into two groups: time domains and transform domains. Fig. 1 shows the watermarking categorization and detailed techniques of each category. Time domain methodologies are further divided into time-aligned [18, 28, 31, 34, 40, 41] and nontime-aligned (echo-based) [63–71] methods. Similarly, transform domain methodologies are further divided into spread spectrum (SS)-based [8,14], patchwork-based [19, 23, 37, 38, 46, 52], quantization index modulation (QIM) based [72–74], and other [20, 21, 29, 30, 33, 36, 43, 75–79] methods. The other methods include ANN (artificial neural network), blockchain, and the integration or hybridization of multiple transformation techniques such as DFT (discrete Fourier transform), DCT (discrete cosine transform), DWT (discrete wavelet transform), SVD (singular value decomposition), etc. to embed watermarks into audio signals. Fig. 2 shows a generic digital audio watermarking system where signal manipulations are carried out in the watermark embedding (encoding) and extraction (decoding) process. Let $x(n)$ be the host signal in the time domain. Hence, the generic model for embedding the watermark $w(n)$ into the $x(n)$ by which the watermarked signal $y(n)$ can be generated in the time domain as

$$y[n] = x[n] + \alpha w[n] \quad (1)$$

where α is the watermark strength – a controlling parameter and n is the time variable. In the transform domain, at first Eq. (1) is transformed and it becomes,

$$Y[k] = X[k] + \alpha W[k] \quad (2)$$

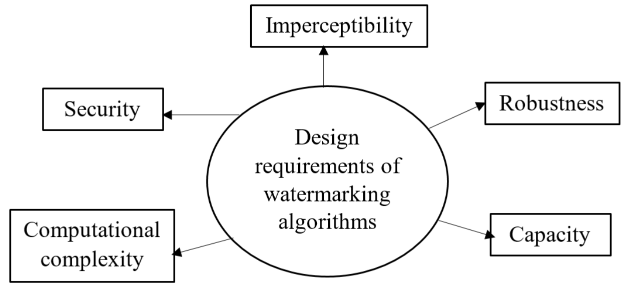

Fig. 3. Design requirements of watermarking techniques.

where X, Y, and W are the transformed representations of x, y, and w, respectively, and k is the transform domain variable. In Eq. (1) the watermark signal is additive with the host signal. However, sometimes, it can be multiplicative. In the multiplicative environment, Eq. (1) can be written as

$$y[n] = x[n] + (1 + \alpha w[n]) \quad (3)$$

III. Design Requirements Issues and Evaluation Metrics

Audio watermarking algorithms embed a watermark into the host signal to uphold the authenticity and copyright from the unauthorized use of the host signal [80]. Hence, it is necessary to define the requirements of an effective watermarking algorithm. Fig. 3 illustrates the design requirements of watermarking techniques. However, for effective watermarking, there is a trade-off among these issues.

A. Imperceptibility

Imperceptibility plays a crucial role in assessing the efficacy of a watermarking algorithm, akin to preserving audio fidelity. In this context, the watermarked image should maintain an identical appearance to the original audio, remaining imperceptible to human observers despite minor alterations. Thus, any impact on audio quality must be minimal. Various subjective and objective methods exist for evaluating the imperceptibility of a watermarking system. Objective measurements consist of evaluating parameters such as SNR (Signal-to-Noise Ratio), fwsSNR (Frequency-Weighted Segmental Signal-to-Noise Ratio) [81, 82], and ODG (Objective Difference Grade) scores [83].

$$\mathrm{SNR} = 10\log_{10}\frac{\sum_1^n x^2[n]}{\sum_1^n (y[n]-x[n])^2}\mathrm{dB} \quad (4)$$

$$\mathrm{fwsSNR} = \frac{10}{N_{seg}}\sum_{i=1}^{N_{seg}}\frac{\sum_1^k |X[k]|^2 \log_{10}\frac{|X[k]|^2}{(|Y_i[k]|-|X_i[k]|)^2}}{\sum_1^k |X[k]|^2}\mathrm{dB} \quad (5)$$

where N_{seg} is the number of non-overlapped frames of the original and watermarked signals and i is the frame index. Other symbols are mentioned in Eq. (1) and (2). Eq. (1) to (5) are explained in detail in [53].

TABLE I. IMPERCEPTIBILITY GRADES AND EXPLANATION

SDG (Subjective Difference Grade)	ODG (Objective Difference Grade)	Explanation/Quality
5.0	0.0	Imperceptible/Excellent
4.0	-1.0	Perceptible, but not annoying/Very Good
3.0	-2.0	Slightly annoying/Good
2.0	-3.0	Annoying/Fair
-1	-4.0	Very annoying/Bad

Performance metrics usually assess imperceptibility by leveraging human auditory capabilities. For this, many tests are used, including the 2AFC (Two-alternative forced choice), AXB paradigm, post hoc test with ANOVA (Analysis of Variance), and SDG (Subjective Difference Grade) [84–86]. In the 2AFC test, the masking curve is determined based on listeners' responses to the original audio compared to watermarked versions at different embedding levels. The AXB test involves three versions of audio clips labeled A, B, and X. A and B, selected randomly from original and watermarked signals (ensuring they are not the same), are presented along with X, randomly chosen from A and B. Listeners then identify which of A or B matches X. The post hoc test requires listeners to grade two audio clips using a scale from 0 to 3, where 0 signifies identical and 3 signifies completely different. These clips consist of an original clip paired with another randomly chosen from original and watermarked versions. The scores provided by all listeners are then subject to ANOVA. In the SDG test, three audio clips A, B, and C are presented. Listeners are tasked with identifying which of B and C closely resembles the original audio A. A grade within {0, -1, -2, -3, -4} is assigned to the selected piece, with 0 indicating imperceptibility and 4 representing significant annoyance. The ODG (Objective Difference Grade) produces scores identical to the SDG but is an automated test without peer listeners. A description of imperceptible grading based on based on ITU-R BS.1387 [83] is shown in Table I.

B. Robustness

Robustness denotes the ability of the original watermark to remain intact despite common signal processing manipulations and attacks by unauthorized users. These manipulations encompass filtering, lossy compression, scaling, translation, rotation, analog-to-digital (A/D) conversion, digital-to-analog (D/A) conversion, and more. Attacks may involve geometric or non-geometric alterations such as filtering, cropping, time shifting, time and pitch scaling, closed-loop attacks, jittering, additive Gaussian noise, echoes, mask, and replacement attacks among others. The robustness of audio watermarking stands as a paramount design criterion, safeguarding against diverse noisy and intentional attacks while preserving the integrity of the watermark data. Robust watermarks find utility in domains like copyright protection, broadcast monitoring, and copy control [87, 88]. Robustness against different types of attacks is measured using the similarity between the watermark signal w and extracted watermark signal w' using normalized correlation (NC) and bit error rate (BER) metrics.

$$\mathrm{NC}(\mathrm{w}, \mathrm{w}') = \frac{\sum_1^n w[n] \sum_1^n w'[n]}{\sqrt{\sum_1^n (w[n])^2 \sum_1^n (w'[n])^2}} \tag{6}$$

$$\mathrm{BER}(\mathrm{w}, \mathrm{w}') = \frac{\sum_1^n w[n] \oplus w'[n]}{N} \tag{7}$$

where $\oplus$ indicates the exclusive OR (XOR) operator between w and w'.

C. Security

Watermarking algorithms lacking security cannot be effectively utilized in copyright protection, data authentication, or audio content tracking. Security assurance is established through various encryption methods, where the encryption key dictates the level of security. Techniques such as chaos-based encryption, Discrete Cosine Transform (DCT), logistic map-based encryption, and binary pseudo-random sequences have been employed to fortify the security and confidentiality of embedded audio watermarks [89, 90]. The watermark key serves as the pivotal secret element ensuring security, and determining specific parameters of the embedding function [91]. This key encompasses aspects like the subset of signal coefficients, embedding direction, and/or embedding domain, comprising a private key, detection key, and public key. The private key remains exclusive to the user, the detection key holds legal acknowledgment, and the public key is accessible to the general populace.

D. Capacity

The watermarking data payload capacity measures how many bits of the watermark are embedded or inserted covertly into the audio signal per unit of time. Therefore, it is quantified in bits per second (bps). The following equation can represent it mathematically:

$$C = \frac{B}{T} \tag{8}$$

where C and B correspond to the data payload capacity and the number of bits embedded in the original audio signal, respectively, and T is the duration of the embedding in seconds. Increasing the amount of watermark information by embedding additional bits presents a challenging endeavor. The insertion of more watermark data into the host audio inevitably leads to heightened distortion becoming perceptible. Consequently, the capacity of the watermarking system delineates the boundaries for watermarking information, all while ensuring robustness and imperceptibility. To this end, watermarking techniques must be adept at minimizing distortion despite having a lesser data embedding capacity. Conventionally, the data payload for audio watermarking should exceed 20 bits per second (bps) to meet the standards set forth by the International Federation of the Phonographic Industry (IFPI).

E. Computational Complexity

The computational expense associated with embedding and extracting a watermark from an audio signal should be kept minimal. This encompasses two primary concerns: the overall time necessary for both embedding and extracting the watermark. Striking a balance between robustness and computational complexity is essential to ensure an optimal trade-off.

Based on the preceding discussions, it is evident that achieving imperceptibility, robustness, and capacity simultaneously poses a challenge due to their inherent conflicts [80]. In any watermarking system, efforts to enhance robustness and capacity may compromise imperceptibility, and vice versa [92]. Conversely, increasing payload capacity can potentially weaken robustness. Hence, finding a delicate balance among these requirements is crucial. For instance, when aiming to render a watermark imperceptible, reducing the energy of the signal seems intuitive. However, a signal's robustness is typically linked to its energy level, as stronger signals are less susceptible to disruption by noise or malicious manipulations. So, finding the right balance between not being able to be detected and being strong is very important. This requires carefully adjusting the energy in the watermark signal so that it does not go too high or too low. It is important to note that there is no universally applicable set of properties that all watermarking systems adhere to.

IV. Audio Watermarking Methods

In the last three decades, diverse methods have been developed for digital audio watermarking, which are categorized in Fig. 2. In this section, we will explain them briefly.

A. Time-domain Techniques

Digital audio watermarking systems that conduct watermark embedding in the time domain offer straightforward solutions by directly modifying the audio samples [18, 41]. In a simple time-domain watermarking system, the least significant bits (LSB) of the audio signal are replaced with watermark bits. Although easy to implement, this method is susceptible to noise manipulation.

Echo-based audio watermarking [63–71] is another method in the time domain that embeds a watermark by adding weak echoes to the host signal. The watermark is then extracted using cepstral analysis. To bolster the security of the audio watermarking system against unauthorized watermark detection, it is recommended to integrate a secret key during both the embedding and extraction phases. Time-spread echo-based techniques have been proposed to meet this security requirement [67]. Echo-based audio watermarking strikes a balance between imperceptibility and robustness, making it suitable for embedding copyright information or other concealed data in audio signals. While it is a well-established technique, for stronger protection against sophisticated audio processing attacks, more advanced watermarking methods in the transform domain might be necessary.

B. Transform-domain Techniques

Transform domain audio watermarking techniques are typically preferred by researchers and designers over time domain methods due to their inherent resilience against various signal processing operations and attacks. In this approach, audio signals undergo initial conversion from the time domain to a transformed domain utilizing mathematical transformations such as DFT, DWT, DCT, or SVD [20, 21, 29, 30, 32, 36, 43, 44, 76, 77]. Following transformation, watermark bits are embedded into specific coefficients within the transformed domain. These coefficients are meticulously selected to ensure imperceptibility to human ears while maintaining robustness against common signal processing operations and attacks. Upon reception, to extract the watermark from the watermarked audio signal, the recipient employs the inverse process. The audio signal is transformed back into the original domain utilized during embedding, after which the watermark extraction algorithm is applied to retrieve the embedded watermark bits. As depicted in Fig. 1, transform domain audio watermarking techniques are broadly categorized into four groups as follows:

1) Spread spectrum (SS)-based method: This audio watermarking technique [8, 14, 93] functions on the principle of dispersing the watermark signal across a broad frequency range within the audio spectrum. Initially, the watermark data undergoes modulation with a pseudo-random sequence, typically generated using algorithms such as pseudo-random noise sequences or pseudo-random phase modulation. These sequences possess specific properties that render them suitable for spreading the watermark across the audio spectrum. The host audio signal is then transformed from the time domain to the frequency domain using DFT or DWT, thereby decomposing the audio signal into its constituent frequency components. Within the frequency domain, the modulated watermark is embedded into selected frequency coefficients of the audio signal. This embedding process entails adding or modulating the watermark information onto the frequency coefficients in a manner that disperses the watermark signal across a wide range of frequencies. The spread spectrum modulation ensures that the embedded watermark remains imperceptible to human ears while demonstrating resilience against common signal processing operations and attacks. Given that the watermark is distributed across multiple frequencies and embedded using pseudo-random sequences, it becomes resistant to localized distortions or attempts to remove it. To extract the watermark from the watermarked audio signal, the recipient employs the same spreading sequence utilized during embedding. By correlating the received signal with the spreading sequence, the embedded watermark can be accurately extracted. This process facilitates the retrieval of the embedded data without significantly compromising the quality of the original audio signal.

Spread spectrum-based audio watermarking finds applications in copyright protection, content authentication, and digital rights management, as it empowers content owners to embed invisible identifiers into their audio content, thereby facilitating the tracking and safeguarding of intellectual property rights.

2) Patchwork-based method: In this technique [19, 23, 37, 38, 46, 52, 94], the audio signal undergoes division into smaller segments or patches, which can vary in length depending on the specific implementation, but typically encompass a few milliseconds of audio data each. Within each patch, watermark data is embedded using various techniques, such as adjusting the amplitude or phase of the audio samples, introducing minor noise alterations, or manipulating frequency components to ensure imperceptibility and resilience against diverse signal processing operations and attacks. Patchwork-based methods often entail analyzing the frequency content of audio patches to identify suitable embedding locations or to modify spectral characteristics for watermarking purposes.

3) Quantization index modulation (QIM)-based method: In digital audio processing, quantization involves mapping continuous amplitude values to discrete levels, thereby reducing the bit depth of the audio signal while preserving perceptual fidelity. Each sample of the audio signal is quantized to a specific level based on its amplitude. QIM-based watermarking modifies the quantization indices of the audio signal to embed the watermark data. Rather than directly altering the amplitude of the samples, it adjusts the indices representing the quantized levels. This adjustment is typically achieved by adding or subtracting a small value from the quantization index, introducing subtle changes in the encoded signal. The QIM technique [73–79] entails modulating the watermarks within the indices of a sequence of quantizers, which are subsequently applied to the host signal. The foundational concept is detailed in [72], where the authors thoroughly explore this technique from an information-theoretic standpoint to practical realization examples. To extract the watermark from the watermarked audio signal, the recipient analyzes the quantization indices of the signal. By comparing the modified indices with the original ones, the embedded watermark data can be extracted. This process necessitates knowledge of the embedding parameters, such as the quantization step size and the location of the watermark within the signal.

Transform domain audio watermarking is utilized in copyright protection, content authentication, and tamper detection in audio signals. It enables content owners to embed invisible identifiers into their audio content, aiding in tracking and safeguarding their intellectual property rights.

4) Other techniques: Other techniques in audio watermarking encompass artificial neural networks (ANN), blockchain technology, and the integration or hybridization of various transformation methods such as DFT, DCT, DWT, SVD, and Schur transform to embed watermarks into audio signals. This hybridization strategy capitalizes on the complementary strengths of different transforms to bolster robustness, imperceptibility, and security, making it a highly sought-after approach in the audio watermarking domain.

Charfeddine et al. [5, 95] introduced an audio watermarking technique rooted in the DCT transform and a neural network (NN) architecture. In this method, the watermark is inserted into middle-frequency bands following the DCT transformation, with the NN model establishing relationships between frequency samples around a central sample during embedding and extraction processes.

Natgunanathan et al. [33] proposed a pioneering privacy protection mechanism for multimedia distribution networks (MDN) by amalgamating the advantages of both blockchain and watermarking technologies. Their approach involves utilizing a specifically designed watermarking algorithm to link copyright information with audio files, alongside a novel blockchain-based smart contract mechanism to ensure the proper functioning of entities within the distribution network. This method demonstrates computational efficiency, with its validity substantiated by simulation results.

Numerous researchers have explored the integration or hybridization of multiple transformation techniques to embed watermarks into audio signals, leveraging the strengths of different methods to bolster robustness, imperceptibility, and security. For instance, Dhar and Shimamura [96-100] combined FFT or DWT with SVD, Aniruddha, and Gnanasekaran [29] integrated DCT with SVD, and Wang and Zhao [77] merged DWT with DCT. These hybridization approaches, often coupled with neural networks, have gained significant traction in the watermarking domain, emerging as state-of-the-art methodologies for achieving heightened robustness, imperceptibility, and security.

In audio watermarking, the choice of watermark signal, whether it be an audio or an image, depends on various factors, including the specific application scenario, the desired level of watermark robustness, and perceptual requirements. Audio watermarks, being in the same format as the original audio, can be seamlessly integrated without noticeable alteration to audio quality. However, the capacity for embedding information within an audio watermark without significantly degrading audio quality may be limited. Conversely, image watermarks typically consist of a binary logo or signature (often 32×32 pixels), allowing for visual verification without specialized equipment, making them ideal for scenarios requiring quick verification. Additionally, images offer greater information capacity compared to audio signals, enabling larger payloads to be embedded within the watermark. However, image watermarks may introduce visible artifacts and be more susceptible to common image processing operations like scaling, cropping, or color adjustments, potentially affecting the visibility or recoverability of the watermark.

Numerous comparative studies have been conducted using simulations, employing standard music signals such as ”Tunisia.wav” for rhythmic music and ”Svega.wav” for a female audio song, as well as Quranic audio files spanning Tracks 1 to 52 [5]. These studies also utilized 16-bit mono audio signals including Pop, Folk, Classical, and Speech, among others. The majority of signals were sampled at a frequency of 44.1 kHz and had durations ranging from approximately 5 to 20 seconds.

Various authors employ various metrics to assess their proposed digital audio watermarking schemes [101, 102]. For instance, imperceptibility analysis results often lack straightforwardness, posing challenges in comparison. Subjective listening tests play a vital role in evaluating the perceptual quality of watermarked audio, though results may vary among listeners. However, the most widely used methods demonstrate imperceptibility through SDG/ODG scores indicating non-annoying and good quality, with a payload capacity exceeding 20 bps to meet IFPI and ITU-R BS.1387 requirements [83]. Robustness evaluation involves subjecting audio watermarking approaches to diverse attacks such as noise addition, filtering, cropping, time shifting, pitch scaling, and masking, etc. Some attacks affect the audio signal more than other attacks. Evaluation in this survey focuses on comparing the performance of widely used schemes using SNR, NC, and BER scores to provide insights into imperceptibility and robustness, particularly under MP3 Compression and Stirmark attacks [103]. Objective comparison results are presented in Table II showcasing benchmark audio watermarking methods. Among the existing methods, this review identifies the technique developed by Charfeddine et al. (2022), [5], as highlighted in Table II, as the state-of-

TABLE II. IMPERCEPTIBILITY AND ROBUSTNESS COMPARISON

SDG (References)	Algorithms	SNR	NC	BER
Charfeddine et al., 2022, [5]	DCT-NN-Human Psychoacoustic Model	47.62	1.00	0.01
Charfeddine et al., 2014, [95]	DCT-NN	43.52	1.00	0.00
Wu and Wu, 2018, [32]	Modifying the average amplitude in the transform domain	23.49	0.98	0.14
Wu and Wu, 2018, [104]	Chaotic encryption in hybrid domain	24.58	0.98	1.92
Lanxun et al., 2007, [105]	DWT-coefficients mean-quantization	37.97	0.98	0.29

the-art in terms of fundamental watermarking requirements. This method conceals the signature within the narrow middle-frequency band of an audio frame, utilizing a neural network architecture for insertion and detection processes to enhance security and robustness, even with high watermark capacity. Additionally, it incorporates aspects of the human psychoacoustic model, aiming to determine the masking threshold curve and align it with the estimated power spectrum density envelope for precise signature insertion. Experimental results underscore the superiority of this masking technique in copyright protection for both standard audio files and sensitive data such as Quranic files, facilitating content integrity verification, proof of authenticity, and tamper detection.

V. RECOMMENDATION FOR FUTURE RESEARCH

In the preceding section, we have highlighted a cutting-edge method for audio watermarking, applicable to real-world scenarios such as copyright protection, content integrity verification, authenticity proof, and tamper detection. Real-time implementation of this technique is paramount. Notably, there exists a discernible contrast between academic and industrial audio watermarking solutions. Industrial solutions, for instance, prioritize imperceptibility over robustness. This prioritization stems from the specific applications defined by each industry solution, necessitating the efficient implementation of audio watermarking systems wherein exhaustive attacks may not be a concern.

Through a comprehensive review of widely employed methods, we have identified the DCT-NN-Human Psychoacoustic Model [5] as the current state-of-the-art. However, the recent integration of blockchain technology holds promise for enhancing the robustness and security of audio watermarking, particularly in the context of copyright protection, tampering detection, and authenticity preservation in the MDN (multimedia distribution networks) environment. A major challenge is the limited availability of standardized databases for evaluating audio watermarking algorithms, underscoring the need for researchers to prioritize this area of focus. Given the superior accuracy observed in image watermarking with deep learning techniques [106], there is potential for leveraging such methodologies in the development of more effective audio watermarking algorithms. Researchers are encouraged to address these issues and explore novel approaches in their endeavors even in speech signals also [107-112].

VI. CONCLUSION

The widespread availability and use of the internet have made audio watermarking an essential technique for safeguarding copyright, preserving ownership, preventing tampering, verifying authenticity, and monitoring audio signal broadcasts. This paper presents a detailed survey of audio watermarking techniques. After outlining the fundamental concepts of audio watermarking, we describe the design criteria and performance metrics. There exists a trade-off among design criteria, including imperceptibility, robustness, and payload capacity. Subsequently, we explore various methods to identify the state-of-the-art technique through performance analysis using evaluation metrics. Furthermore, we discuss remaining challenges and potential avenues for enhancing audio watermarking systems. We also examine the disparities between academic and industrial solutions in audio watermarking. This paper aims to assist researchers in identifying and developing optimal algorithms tailored to audio watermarking.

Credit Authorship Contribution Statement Mohammad Shorif Uddin: Writing – original draft, Writing –review & editing, Methodology. Ohidujjaman : Writing –review & editing, Conceptualization, Formal analysis, Data curation. Mahmudul Hasan: Writing –review & editing, Data curation. Tetsuya Shimamura: Supervision.

REFERENCES

[1] V. Atti, T. Painter, and A. Spanias, “Audio Signal Processing and Coding,” John Wiley & Sons, 2007.
V. Atti, T. Painter, A. Spanias

[2] M. Swanson, M. Kobayashi, and A. Tewfik, “Multimedia data-embedding and watermarking technologies,” Proc. IEEE 86 (6), 1064–1087, 1998.

[3] C. Xu, J. Wu, Q. Sun and K. Xin, “Applications of digital watermarking technology in audio signals,” Journal of the Audio Engineering Society, vol. 47, No. 10, pp. 805-812, 1999.

[4] M. Steinebach, and J. Dittmann, “Watermarking-based digital audio data authentication,” EURASIP Journal on Advances in Signal Processing, pp. 1-15, 2023.

[5] M. Charfeddine, E. Mezghani, S. Masmoudi, C. B. Amar, and H. Alhumyani, “Audio watermarking for security and non-security applications,” IEEE Access, vol. 10, pp. 12654-12677, 2022.

[6] S. Masmoudi, M. Charfeddine, S. Alsharif, and C. B. Amar, “A New Blind IoT-Based MP3 Audio Watermarking Scheme for Content Integrity Checking and Copyright Protection,” Wireless Communications and Mobile Computing, 2022.

[7] L. Boney, A. Tewfik, K. Hamdy, “Digital watermarks for audio signals,” Third IEEE International Conference on Multimedia Computing and Systems, 1996, pp. 473–480.

[8] I. J. Cox, J. Kilian, F.T. Leighton, and T. Shamoon, “Secure spread spectrum water-marking for multimedia,” IEEE Trans. Image Process. vol. 6, No. 12, pp. 1673–1687, 1997.

[9] F. Hartung, and M. Kutter, “Multimedia watermarking techniques,” Proc. IEEE 87(7), 1999, pp. 1079–1107.

[10] Y. Xiang, G. Hua, and B. Yan, “Digital audio watermarking: fundamentals, techniques and challenges,” Singapore: Springer, 2017.

[11] A. Patil and R. Shelke, “Digital audio watermarking: techniques, applications, and challenges,” Intelligent Sustainable Systems: Selected Papers of WorldS4 2021, vol. 2, pp. 679-689, 2022.

[12] Xing He, “Watermarking in audio: key techniques and technologies,” Cambria Press, 2008.

[13] R. M. Thanki, “Advanced Techniques for Audio Watermarking,” Springer, 2020.

[14] D. Kirovski and H. S. Malvar, “Spread-spectrum watermarking of audio signals,” IEEE Transactions on Signal Processing vol. 51, No. 4, pp. 1020-1033, 2003.

[15] A. G. Acevedo, “Audio watermarking: properties, techniques and evaluation,” Multimedia security: Steganography and Digital Watermarking Techniques for Protection of Intellectual Property. IGI Global, 2005, pp. 75-125.

[16] Chung-Ping Wu, Po-Chyi Su, and C-C. Jay Kuo, “Robust and efficient digital audio watermarking using audio content analysis,” Security and Watermarking of Multimedia Contents, vol. 3971. SPIE, 2000.

[17] S. D. Larbi, and M. Jaïdane-Saïdane, “Audio watermarking: a way to stationnarize audio signals.” IEEE Transactions on Signal Processing, vol. 53, No. 2, pp. 816-823, 2005.

[18] P. Bassia, I. Pitas, and N. Nikolaidis, ”Robust audio watermarking in the time domain,” IEEE Transactions on Multimedia, vol. 3, No. 2, pp. 232-241, 2001.

[19] Y. Chincholkar and S. Ganorkar, “Audio watermarking algorithm implementation using patchwork technique,” 5th International Conference for Convergence in Technology (I2CT), IEEE, 2019.

[20] P. K. Dhar, and T. Shimamura, “DWT-DCT-Based Audio Watermarking Using SVD,” Advances in Audio Watermarking Based on Singular Value Decomposition, Springer Briefs in Electrical and Computer Engineering, Springer, 2015, pp. 17-35.

[21] H. Karajeh, T. Khatib, L. Rajab, and M. Maqableh, “A robust digital audio watermarking scheme based on DWT and Schur decomposition,” Multimedia Tools and Applications, vol. 78, pp. 18395-18418, 2019.

[22] A. K. Chowdhury and M. I. Khan, “A tutorial for audio watermarking in the cepstrum domain,” Smart Computing Review, vol. 3, No. 5, pp. 323-335, 2013.

[23] In-Kwon Yeo and H. J. Kim, “Modified patchwork algorithm: A novel audio watermarking scheme,” IEEE Transactions on speech and audio processing, vol. 11, No. 4, pp. 381-386, 2003.

[24] Y. Lin, and W. H. Abdulla, “Audio watermark,” Springer, 2015. https://doi.org/10.1007/978-3-319-07974-5

[25] Hyoung-Joong Kim, S. Katzenbeisser, and Anthony T. S. Ho, “Digital watermarking,” 7th International Workshop, IWDW 2008, Busan, Korea, November 10-12, 2008.

[26] X. Wang, W. Qi, and P. Niu, “A New Adaptive Digital Audio Watermarking Based on Support Vector Regression,” in IEEE Transactions on Audio, Speech, and Language Processing, vol. 15, no. 8, pp. 2270-2277, Nov. 2007.

[27] M. D. Swanson, B. Zhu, A. H. Tewfik, and L. Boney, “Robust audio watermarking using perceptual masking,” Signal Processing, vol. 66, No. 3, pp. 337-355, 1998.

[28] S. Esmaili, S. Krishnan and K. Raahemifar, “Audio watermarking time-frequency characteristics,” in Canadian Journal of Electrical and Computer Engineering, vol. 28, no. 2, pp. 57-61, April 2003.

[29] K. Aniruddha and A. Gnanasekaran, “Robust image-in-audio watermarking technique based on DCT-SVD transform,” EURASIP Journal on Audio, Speech, and Music Processing, Article 16, 2018.

[30] S. M. Pourhashemi, M. Mosleh, and Y. Erfani, “A novel audio watermarking scheme using ensemble-based watermark detector and discrete wavelet transform,” Neural Computing and Applications, vol. 33, No. 11, pp. 6161-6181, 2021.

[31] W. Li, X. Xue and P. Lu, “Localized audio watermarking technique robust against time-scale modification,” in IEEE Transactions on Multimedia, vol. 8, no. 1, pp. 60-69, Feb. 2006.

[32] Q. Wu and M. Wu, “A novel robust audio watermarking algorithm by modifying the average amplitude in transform domain,” Appl. Sci., vol. 8, no. 5, p. 723, May 2018.

[33] I. Natgunanathan, P. Praitheeshan, L. Gao, Y. Xiang, and L. Pan, “Blockchain-based audio watermarking technique for multimedia copyright protection in distribution networks,” ACM Transactions on Multimedia Computing, Communications, and Applications, vol. 18, No. 3, pp. 1-23, 2022.

[34] Wen-Nung Lie and Li-Chun Chang, “Robust and high-quality time-domain audio watermarking based on low-frequency amplitude modification,” in IEEE Transactions on Multimedia, vol. 8, no. 1, pp. 46-59, Feb. 2006.

[35] K. M. Abdelwahab, S. M. Abd El-atty, W. El-Shafai, S. El-Rabaie, and F. E. Abd El-Samie, ”Efficient SVD-based audio watermarking technique in FRT domain,” Multimedia Tools and Applications, vol. 79, pp. 5617-5648, 2020.

[36] K. A. Darabkh, “Imperceptible and robust DWT-SVD-based digital audio watermarking algorithm,” Journal of Software Engineering and Applications, vol. 7, pp. 859-871, 2014.

[37] N. K. Kalantari, M. A. Akhaee, S. M. Ahadi and H. Amindavar, “Robust Multiplicative Patchwork Method for Audio Watermarking,” in IEEE Transactions on Audio, Speech, and Language Processing, vol. 17, no. 6, pp. 1133-1141, Aug. 2009.

[38] I. Natgunanathan, Y. Xiang, G. Hua, G. Beliakov and J. Yearwood, “Patchwork-Based Multilayer Audio Watermarking,” in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 25, no. 11, pp. 2176-2187, Nov. 2017.

[39] M. Fallahpour and D. Megías, “Audio Watermarking Based on Fibonacci Numbers,” in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 23, no. 8, pp. 1273-1282, Aug. 2015.

[40] S. Xiang and J. Huang, “Histogram-Based Audio Watermarking Against Time-Scale Modification and Cropping Attacks,” in IEEE Transactions on Multimedia, vol. 9, no. 7, pp. 1357-1372, Nov. 2007.

[41] A. N. Lemma, J. Aprea, W. Oomen and L. van de Kerkhof, “A temporal domain audio watermarking technique,” in IEEE Transactions on Signal Processing, vol. 51, no. 4, pp. 1088-1097, April 2003.

[42] X. Tang, Z. Ma, X. Niu, and Y. Yang, “Robust audio watermarking algorithm based on empirical mode decomposition,” Chinese Journal of Electronics, vol. 25, no. 6, pp. 1005-1010, 2016.

[43] S. Gulivindala, N. V. Lalitha, D. P. Gangwar, and A. K. Sahu, “False-positive-free SVD based audio watermarking with integer wavelet transform,” Circuits, Systems, and Signal Processing, vol. 41, no. 9, pp. 5108-5133, 2022.

[44] M. Mosleh, S. Setayeshi, B. Barekatain, and M. Mosleh, “A novel audio watermarking scheme based on fuzzy inference system in DCT domain,” Multimedia Tools and Applications, vol. 80, no. 13, pp. 20423-20447, 2021.

[45] B. A. F. Agradriya, F. K. Perdana, I. Safitri, L. Novamizanti, “Audio watermarking technique based on Arnold transform,” 2017 2nd International Conference on Automation, Cognitive Science, Optics, Micro Electro-Mechanical System, and Information Technology (ICACOMIT), Jakarta, Indonesia, pp. 17-21, 2017.

[46] I. Natgunanathan, Y. Xiang, Y. Rong, W. Zhou and S. Guo, “Robust Patchwork-Based Embedding and Decoding Scheme for Digital Audio Watermarking,” in IEEE Transactions on Audio, Speech, and Language Processing, vol. 20, no. 8, pp. 2232-2239, Oct. 2012.

[47] M. V. D. Veen, L. A. van, and B. Fons, “Reversible audio watermarking,” In Audio Engineering Society Convention 114. Audio Engineering Society, 2003.

[48] S. Voloshynovskiy, S. Pereira, T. Pun, J. J. Eggers, and J. K. Su, “Attacks on digital watermarks: classification, estimation based attacks, and benchmarks,” IEEE Communications Magazine, vol. 39, no. 8, pp. 118-126, Aug. 2001.

[49] J. Cichowski, A. Czyżewski, and B. Kostek, “Analysis of the impact of audio modifications on the robustness of watermark for non-blind architecture,” Multimedia Tools and Applications, vol. 74, pp. 4415-4435, 2013.

[50] M O Agbaje and Adebayo A. O., “Robustness and Security Issues in Digital Audio Watermarking,” International Journal of Engineering and Information Systems (IJEAIS), pp. 1–10, 2017.

[51] M. Tanha, D. Sajjadi, M. T. Abdullah, and F. Hashim, “An overview of attacks against digital watermarking and their respective countermeasures,” International Conference on Cyber Security, Cyber Warfare and Digital Forensic (CyberSec), pp. 265-270. IEEE, 2012.

[52] Z. Liu, Y. Huang and J. Huang, “Patchwork-Based Audio Watermarking Robust Against De-Synchronization and Recapturing Attacks,” IEEE Transactions on Information Forensics and Security, vol. 14, no. 5, pp. 1171-1180, May 2019.

[53] G. Hua, J. Huang, Y. Q. Shi, J. Goh, and Vrizlynn L. L. Thing, “Twenty years of digital audio watermarking—a comprehensive review,” Signal Processing, vol. 128, pp. 222-242, 2016.

[54] R. D. Shelke, and M. U. Nemade, “Audio watermarking techniques for copyright protection: A review,” International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC), Jalgaon, India, pp. 634-640, 2016.

[55] R. Jain, M. C. Trivedi, and S. Tiwari, “Digital audio watermarking: A survey,” Advances in Intelligent Systems and Computing, vol. 554, 2018, Springer, Singapore.

[56] S. Kumar, B. K. Singh, and M. Yadav, “A recent survey on multimedia and database watermarking,” Multimedia Tools and Applications, vol. 79, pp. 20149-20197, 2020.

[57] Sin-Joo Lee and Sung-Hwan Jung, ”A survey of watermarking techniques applied to multimedia,” IEEE International Symposium on Industrial Electronics, Pusan, Korea (South), pp. 272-277, 2001.

[58] P. Garg and R. R. Kishore, “Performance comparison of various watermarking techniques,” Multimedia Tools and Applications, vol. 79, pp. 25921-25967, 2020.

[59] F. A. P. Petitcolas, R. Anderson, and M. G. Kuhn, “Information hiding - a survey,” Proc. IEEE 87(7), pp. 1062–1078, 1999.

[60] M. V. Patil and J. S. Chitode, ”A Concise Review of Digital Audio Watermarking Techniques,” International Journal of Digital Signal Processing, vol. 4, no. 5, May 2012.

[61] Y. Li, H. Wang, and M. Barni, ”A survey of deep neural network watermarking techniques,” Neurocomputing vol. 461, pp. 171-193, 2021.

[62] A. Jain, D. Somwanshi, C. Khurana, and K. Joshi “Review on Digital Watermarking Techniques and Its Retrieval,” In 2022 International Conference on Fourth Industrial Revolution Based Technology and Practices (ICFIRTP), pp. 274-278. IEEE, 2022.

[63] D. Gruhl, W. Bender, “Echo hiding,” In Proceedings of the Information Hiding Workshop, Cambridge, U.K., 1996, pp. 295–315.

[64] H. O. Oh, J. W. Seok, J. W. Hong and D. H. Youn, “New echo embedding technique for robust and imperceptible audio watermarking,” IEEE International Conference on Acoustics, Speech, and Signal Processing, Salt Lake City, UT, USA, 2001, pp. 1341-1344 vol. 3.

[65] H. J. Kim and Y. H. Choi, ”A novel echo-hiding scheme with backward and forward kernels,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 8, pp. 885-889, Aug. 2003.

[66] O. T. C. Chen and W. C. Wu, ”Highly Robust, Secure, and Perceptual-Quality Echo Hiding Scheme,” IEEE Transactions on Audio, Speech, and Language Processing, vol. 16, no. 3, pp. 629-638, March 2008.

[67] Byeong-Seob Ko, R. Nishimura and Y. Suzuki, ”Time-spread echo method for digital audio watermarking,” IEEE Transactions on Multimedia, vol. 7, no. 2, pp. 212-221, April 2005.

[68] Y. Xiang, D. Peng, I. Natgunanathan and W. Zhou, ”Effective Pseudonoise Sequence and Decoding Function for Imperceptibility and Robustness Enhancement in Time-Spread Echo-Based Audio Watermarking,” in IEEE Transactions on Multimedia, vol. 13, no. 1, pp. 2-13, Feb. 2011.

[69] Y. Xiang, I. Natgunanathan, D. Peng, W. Zhou and S. Yu, ”A Dual-Channel Time-Spread Echo Method for Audio Watermarking,” in IEEE Transactions on Information Forensics and Security, vol. 7, no. 2, pp. 383-392, April 2012.

[70] G. Hua, J. Goh, and V. L. L. Thing, “Time-spread echo-based audio watermarking with optimized imperceptibility and robustness,” IEEE/ACM Transactions on Audio, Speech and Language Processing, vol. 23, no. 2, pp 227–239, 2015.

[71] P. Hu, D. Peng, Z. Yi, and Y. Xiang, “Robust time-spread echo watermarking using characteristics of host signals,” Electron. Lett. vol. 52, no. 1, 2016.

[72] B. Chen and G. W. Wornell, ”Quantization index modulation: a class of provably good methods for digital watermarking and information embedding,” IEEE Transactions on Information Theory, vol. 47, no. 4, pp. 1423-1443, May 2001.

[73] I. D. Shterev and R. L. Lagendijk, ”Amplitude Scale Estimation for Quantization-Based Watermarking,” IEEE Transactions on Signal Processing, vol. 54, no. 11, pp. 4146-4155, Nov. 2006.

[74] S. Wu, J. Huang, D. Huang and Y. Q. Shi, ”Efficiently self-synchronized audio watermarking for assured audio data transmission,” in IEEE Transactions on Broadcasting, vol. 51, no. 1, pp. 69-76, March 2005.

[75] K. Khaldi and A. O. Boudraa, ”Audio Watermarking Via EMD,” in IEEE Transactions on Audio, Speech, and Language Processing, vol. 21, no. 3, pp. 675-680, March 2013.

[76] B. Lei, I. Y. Soon and E. -L. Tan, ”Robust SVD-Based Audio Watermarking Scheme With Differential Evolution Optimization,” in IEEE Transactions on Audio, Speech, and Language Processing, vol. 21, no. 11, pp. 2368-2378, Nov. 2013.

[77] X. Y. Wang and H. Zhao, ”A Novel Synchronization Invariant Audio Watermarking Scheme Based on DWT and DCT,” in IEEE Transactions on Signal Processing, vol. 54, no. 12, pp. 4835-4840, Dec. 2006.

[78] Xiang-Yang Wang, Pan-Pan Niu, and Hong-Ying Yang, ”A Robust, Digital-Audio Watermarking Method,” in IEEE MultiMedia, vol. 16, no. 3, pp. 60-69, July-Sept. 2009.

[79] X. Wang, W. Qi and P. Niu, ”A New Adaptive Digital Audio Watermarking Based on Support Vector Regression,” IEEE Transactions on Audio, Speech, and Language Processing, vol. 15, no. 8, pp. 2270-2277, Nov. 2007.

[80] H. Tao, L. Chongmin, J. M. Zain, A. N. Abdalla, “Robust Image Watermarking Theories and Techniques: A Review,” J. Appl. Res. Technol. vol. 12, no. 1, pp. 122–138, 2014.

[81] Y. Hu, and P. C. Loizou, “Evaluation of objective quality measures for speech enhancement,” IEEE Trans. Audio, Speech, Lang. Process. vol. 16, no. 1, 2008.

[82] A. Nishimura, M. Unoki, K. Kondo, and A. Ogihara, “Objective evaluation of sound quality for attacks on robust audio watermarking,” Proc. Meet. Acoust., pp. 1–9, 2013.

[83] Recommendation ITU-R BS.1387-1: Method for objective measurements of perceived audio quality, (1998–2001).

[84] S. A. Gelfand, “Hearing: An Introduction to Psychological and Physiological Acoustics,” 3rd Edition, Marcel Dekker, Basel, Switzerland, 1998.

[85] B. C. J. Moore, “An Introduction to the Psychology of Hearing,” 4th Edition, Academic, New York, 1997.

[86] M. Unoki, K. Imabeppu, D. Hamada, A. Haniu, R. Miyauchi, “Embedding limitations with digital-audio watermarking method based on cochlear delay characteristics,” J. Inf. Hiding Multimed. Signal Process. vol. 2, no. 1, pp. 1-23, 2011.

[87] J. Liu, X. He, “A Review Study on Digital Watermarking,” 1st International Conference on Information and Communication Technologies, ICICT, Karachi, Pakistan, 27–28 August 2005, pp. 337–341.

[88] U. Yadav, J.P. Sharma, D. Sharma, P.K. Sharma, “Different Watermarking Techniques & its Applications: A Review,” Int. J. Sci. Eng. Res. vol. 5, 1288–1294, 2014.

[89] N. A. Loani, N.N. Hurrahi, S.A. Parah, J.W. Lee, J. A. Sheikhi, G. Mohiuddin Bhat, “Secure and Robust Digital Image Watermarking Using Coefficient Differencing and Chaotic Encryption,” IEEE Access, vol. 6, pp. 19876–19897, 2018.

[90] F. Zhang, H. Zhang, “Digital Watermarking Capacity and Reliability,” IEEE International Conference on e-Commerce Technology, San Diego, CA, USA, 9 July 2004, pp. 295–298.

[91] L. P’erez-Freire, P.C. Na, J. Ramon, J.R. Troncoso-Pastoriza, F.P. Gonzalez, “Watermarking Security: A Survey,” Transactions on Data Hiding and Multimedia Security; Lecture Notes in Computer Science: Berlin/Heidelberg, Germany, pp. 41–72, 2006.

[92] C. De Vleeschouwer, J. F. Delaigle, B. Macq, ”Invisibility and application functionalities in perceptual watermarking an overview,” in Proceedings of the IEEE, vol. 90, no. 1, pp. 64-77, Jan. 2002.

[93] W. Bender, D. Gruhl, N. Morimoto and A. Lu, ”Techniques for data hiding,” in IBM Systems Journal, vol. 35, no. 3.4, pp. 313-336, 1996.

[94] M. Arnold, “Audio watermarking: features, applications and algorithms,” IEEE International Conference on Multimedia and Expo, 2000, (ICME 2000), vol. 2, IEEE, New York, NY, 2000, pp. 1013–1016.

[95] M. Charfeddine, M. El’arbi, and C. Ben Amar, “A new DCT audio watermarking scheme based on preliminary MP3 study,” Multimedia Tools Appl., vol. 70, no. 3, pp. 1521–1557, Jun. 2014.

[96] P. K. Dhar, T. Shimamura, “FFT-Based Audio Watermarking Using SVD and CPT,” Advances in Audio Watermarking Based on Singular Value Decomposition. Springer Briefs in Electrical and Computer Engineering. Springer, Cham., 2015, pp. 37-52.

[97] P. K. Dhar, T. Shimamura, “LWT-Based Audio Watermarking Using FWHT and SVD,” Advances in Audio Watermarking Based on Matrix Decomposition. Springer Briefs in Speech Technology, Springer, Cham., 2019, pp. 11-22.

[98] P. K. Dhar, T. Shimamura, “Audio Watermarking Based on LWT and QRD,” Advances in Audio Watermarking Based on Matrix Decomposition. Springer Briefs in Speech Technology. Springer, Cham, 2019, pp. 23-32.

[99] P. K. Dhar, T. Shimamura, “Audio Watermarking Based on LWT and SD,” Advances in Audio Watermarking Based on Matrix Decomposition. Springer Briefs in Speech Technology. Springer, Cham., 2019, pp. 43-52.

[100] P. K. Dhar, T. Shimamura, “Audio Watermarking Based on FWHT and LUD,” Advances in Audio Watermarking Based on Matrix Decomposition. Springer Briefs in Speech Technology. Springer, Cham., 2019, pp. 33-42.

[101] Y. Kowalczuk, J. Holub, “Evaluation of digital watermarking on subjective speech quality,” Scientific Reports, vol. 11, No. 20185 (2021).

[102] B. B. Zaidan, A. A. Zaidan, H. Abdul. Karim, N. N. Ahmad, “A new digital watermarking evaluation and benchmarking methodology using an external group of evaluators and multi-criteria analysis based on large-scale data,” Software Practice Experience, Wiley, vol. 47, pp. 1365–1392, 2016.

[103] M. Steinebach, F. A. P. Petitcolas, F. Raynal, J. Dittmann, C. Fontaine, S. Seibel, N. Fates, and L. C. Ferri, ”StirMark benchmark: audio watermarking attacks,” International Conference on Information Technology: Coding and Computing, Las Vegas, NV, USA, 2001, pp. 49-54.

[104] Q. Wu and M. Wu, “Adaptive and blind audio watermarking algorithm based on chaotic encryption in hybrid domain,” Symmetry, vol. 10, no. 7, pp. 284, Jul. 2018.

[105] W. Lanxun, Y. Chao, and P. Jiao, “An audio watermark embedding algorithm based on mean-quantization in wavelet domain,” Int. Conf. Electron. Meas. Instrum., Xi’an, China, Aug. 2007, pp. 423–425.

[106] X. Zhong, A. Das, F. Alrasheedi, and A. Tanvir, “A Brief, In-Depth Survey of Deep Learning Based Image Watermarking,” Appl. Sci. vol. 13, no. 11852, 2023.

[107] Ohidujjaman, M. Hasan, and M. N. Huda, “Improving Speech Signal Intelligibility by Optimal Computation using Single-Channel Adaptive Filtering,” International Journal of Computer Applications, vol. 106, no. 9, Nov. 2014.

[108] Ohidujjaman, N. Yasui, Y. Sugiura, T. Shimamura and H. Makinae, “Packet Loss Concealment Estimating Residual Errors of Forward-Backward Linear Prediction for Bone-Conducted Speech,” International Journal of Advanced Computer Science and Applications (IJACSA), vol. 15, no. 4, April 2024.

[109] M. Hasan, Ohidujjaman, and M. N. Huda, ”Enhancement of Speech Signal by Originating Computational Iteration using SAF,” 14th IEEE – ISSPIT 2014, PID-12, 15th-17th December 2014, JIIT, Noida, India.

[110] Ohidujjaman, Y. Sugiura, N. Yasui, T. Shimamura and H. Makinae, “Regularized Modified Covariance Method for Spectral Analysis of Bone-Conducted Speech,” Journal of Signal Processing, Japan, vol. 28, no. 3, 2024.

[111] M. Hasan and M. E. Hamid, ”A Parametric Formulation to Detect Speech Activity of Noisy Speech using EDON,” ICCIT, IEEE, Paper ID-155, pp. 252-255, 23-25 December 2010, Dhaka, Bangladesh.

[112] Ohidujjaman, N. Yasui, Y. Sugiura, T. Shimamura and H. Makinae, “Packet Loss Compensation for VoIP through Bone-Conducted Speech Using Modified Linear Prediction,” IEEJ Trans. Electrical and Electronic Engineering, vol. 18, no. 11, 2023.