Watermarking-based Fake Audio NFT Detection in NFT Marketplace

Muhammad Rasyid Redha Ansori, Allwinnaldo, Revin Naufal Alief,
Ikechi Saviour Igboanusi, Jae Min Lee, Dong-Seong Kim
Department of IT Convergence, Kumoh National Institute of Technology, Gumi, South Korea
{rasyidred, winnaldo, revinnaufal, ikechisaviour, ljmpaul, dskim}@kumoh.ac.kr

***Abstract*—This research presents an audio watermarking-based technique to prevent users from minting fake Non-fungible Tokens (NFTs) by reusing the same audio uploaded into an NFT marketplace. The robustness of the proposed model is evaluated by calculating the bit error rate (BER) of the embedded watermark. Furthermore, the quality of the watermarked audio is evaluated using Objective Difference Grade (ODG) and Signal-to-Noise Ratio (SNR). The results demonstrate that the embedded watermark can sustain low pass filtering, resampling, and MP3 compression attacks with a BER of, at worst, 0.075. The proposed approach generates watermarked audio of high quality with the best ODG of -0.21 and SNR of 34.36 dB. Lastly, the average execution time for similarity is 1.3 seconds.**

***Index Terms*—Audio, Fake NFT, NFT Marketplace, Watermarking**

I. Introduction

As blockchain technology advances, its applications become more widespread, notably in the information and security sectors [1], [2]. In contrast to physical art, in which each work is one-of-a-kind, the originality of digital art is frequently disrupted by its ability to create identical copies. With blockchain technology, digital art can be tokenized, making each of it unique by using NFT technology [3].

NFTs, on the other hand, continue to have problems with their legitimacy because a fraudster is able to recreate the same data and resell them as their original work. The smart contract handling collection validates the genuineness of an NFT. Therefore, before making a purchase, users are recommended to validate the contract address of the collection using official sources such as the collection's website to guarantee that the token they are purchasing is authentic. However, buyers are only sometimes aware of the possibility of copycat collections or how to check the legitimacy of NFT collections. In addition, no NFT marketplace can perform a similarity check to see if a media file has been used previously by other NFTs, allowing malevolent individuals to create fake NFTs [4].

There is a study about protecting audio copyright using watermarking and blockchain. The authors of [5] proposed the implementation of the Ethereum blockchain and smart contracts to guarantee audio rights via BMCProtector. Vector quantization watermarking, advanced encryption standard encryption, and a digital rights management to safeguard the music copyright are the three methods used by BMCProtector. However, the authors did not present the reliability of the proposed audio watermarking algorithms.

This paper presents a framework for detecting fake audio NFTs using the audio watermarking technique. The Discrete Wavelet Transform (DWT) and Statistical Mean Manipulation (SMM) techniques are combined to achieve the watermarking approach, which embeds a watermark into audio files. Then, the robustness, imperceptibility, capacity, and processing time for watermarking are evaluated.

This study is divided into subsequent sections. Section II discusses the model being presented. The experimental setup and results are described in Section III, while the conclusion and future work are discussed in Section IV.

II. Proposed Model

This section covers the proposed model for detecting if audio has been used previously by some other NFT in the marketplace. The authors' prior work [6] is implemented as the watermarking methods in this paper. The framework of the proposed model is presented in Fig. 1. The steps are described below.

1) A user attempts to submit an audio file to mint NFT on the marketplace. The user could be a legitimate musician or a fraudulent one attempting to upload previously uploaded audio from another NFT.
2) Next, extract the watermark in the submitted audio.
3) The output of the previous step should be compared to the watermark of the NFT market.
4) Suppose a watermark is identified as similar to the one used in the marketplace. In that case, as represented by the similarity being more than the threshold, the audio already exists in the marketplace, and the model will reject it. But if the similarity value is less than the threshold, it signifies that no matching audio was located in the marketplace and the audio will be listed in the marketplace.
5) The model then embeds a watermark into the audio and upload both watermarked audio and the watermark into a hosting service, which returns URI for each.
6) Mint an audio NFT using both watermarked audio and watermark URIs.

III. Experimental Setup and Result

This study was conducted on a system with an Intel core i5-8500 @ 3 GHz, a GeForce GTX 1050 GPU, and 24 GB of RAM. The simulation for audio watermarking was carried

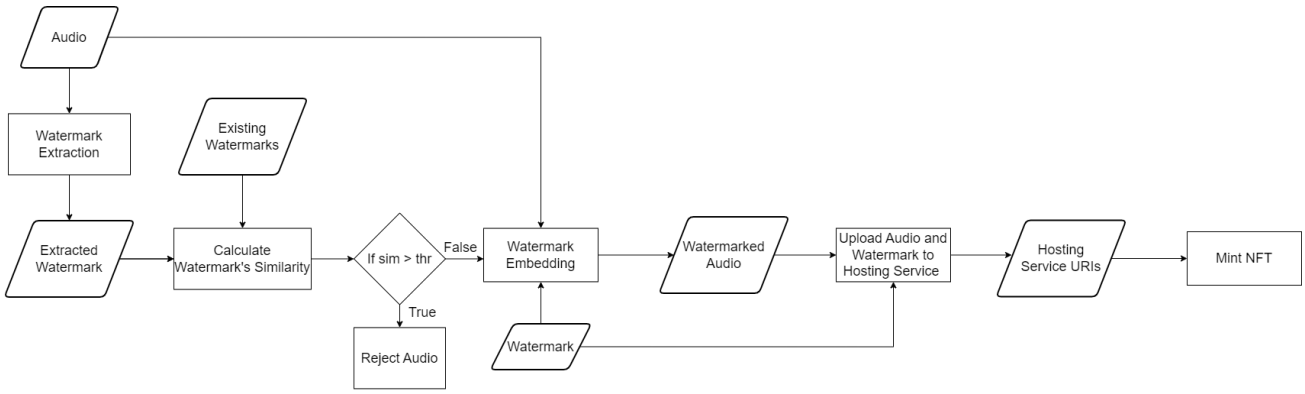

Fig. 1. Propose Model Flow Chart

out using Matlab 2021b, three audios with a sampling rate of 44,100 Hz, a quantization rate of 16 bits per sample, and a binary image with a size of 8×15 pixels as the watermark. The optimal audio watermarking parameters are as follows: DWT decomposition level 4, the first subband is chosen for embedding the watermark, 4096 samples per frame, and watermarking gain of 0.025. After embedding the watermark, LPF, resampling, and MP3 compression were applied to the watermarked audio to test the watermark's durability. The bit error rate (BER) is the metric for determining the similarity. Since the watermark becomes unreadable at BER values higher than 0.1, this value was chosen as the similarity threshold. The average processing time for embedding the watermark is 2.2 seconds, whereas for extraction and similarity calculation is 1.3 seconds.

TABLE I
IMPERCEPTIBILITY AND CAPACITY OF WATERMARKED AUDIO

Audio	ODG	SNR	C
audio-1	-1.23	25.81	0.67
audio-2	-0.21	27.76	0.67
audio-3	-0.21	34.36	0.67

Table I shows the imperceptibility and capability of the watermarked audio. The watermarking method resulted in good watermarked audio, with an ODG value ranging from -1.23 to -0.21 and an SNR that ranged from 25.81 to 34.36 dB. However, watermarked audio has a capacity of only 0.67 bps.

TABLE II
ROBUSTNESS OF AUDIO WATERMARKING METHOD

Attacks	Parameter	BER		
		audio-1	audio-2	audio-3
LPF	6 KHz	0.05	0	0
	9 KHz	0	0	0
Resampling	22.05 K	0.05	0	0
	16 K	0	0	0
MP3 Compression	32 K	0.075	0	0
	64 K	0	0	0

Table II shows the robustness of the watermark. The watermark can withstand attacks with a maximum BER value of 0.075. It implies that the approach of watermarking is resistant to signal processing manipulations.

IV. CONCLUSION AND FUTURE WORK

This paper presented a watermarking-based approach to prevent malicious users to reuse the same audio that has been used for minting audio NFT. The watermarking methods combine DWT and SMM, produces a good robustness and quality of watermarked audio, and fast similarity check process. For future improvement, it is necessary to detect fake NFT collection names, as there are many users deceived by similar name of NFT projects.

V. ACKNOWLEDGEMENT

This research work was supported by Priority Research Centers Program through NRF funded by MEST (2018R1A6A1A03024003), NRF (NRF-2022R1I1A3071844), and the Grand Information Technology Research Center support program (IITP-2023-2020-0-01612) supervised by the IITP by MSIT, Korea.

REFERENCES

[1] H. Tran-Dang and D.-S. Kim, "The Physical Internet in the Era of Digital Transformation: Perspectives and Open Issues," *IEEE Access*, vol. 9, pp. 164 613–164 631, 2021.

[2] I. S. Igboanusi, K. P. Dirgantoro, J.-M. Lee, and D.-S. Kim, "Blockchain Side Implementation of Pure Wallet (PW): An Offline Transaction Architecture," *ICT Express*, vol. 7, no. 3, pp. 327–334, 2021.

[3] F. Temmermans, D. Bhowmik, F. Pereira, and T. Ebrahimi, "Media security framework inspired by emerging challenges in fake media and NFT," in *Optics, Photonics and Digital Technologies for Imaging Applications VII*, P. Schelkens and T. Kozacki, Eds., vol. 12138, International Society for Optics and Photonics. SPIE, 2022, p. 121380P.

[4] D. Das, P. Bose, N. Ruaro, C. Kruegel, and G. Vigna, "Understanding Security Issues in the NFT Ecosystem," *CoRR*, vol. abs/2111.08893, 2021. [Online]. Available: https://arxiv.org/abs/2111.08893

[5] S. Zhao and D. O'Mahony, "BMCProtector: A Blockchain and Smart Contract Based Application for Music Copyright Protection," in *Proceedings of the 2018 International Conference on Blockchain Technology and Application*, ser. ICBTA 2018. New York, NY, USA: Association for Computing Machinery, 2018, p. 1–5.

[6] M. R. R. Ansori, S. Ajakwe, J. M. Lee, and D.-S. Kim, "A Robust MP3 Compression-Resistant Audio Watermarking Algorithm Based on DWT-SMM," vol. 2021, no. 11. Proceedings of the 2021 Conference of The Korean Institute of Communications and Information Sciences, 2021, pp. 393–394.