Justice Assistant: Real-Life Court Audio Tampering *Detection System*

Veilraj S[1], Sujith R[2], Sakthivel S[3]
[1,2,3] *Artificial Intelligence and Machine Learning Rajalakshmi Institute of Technology Chennai, India*

***Abstract*—Justice Assistant is basically an AI system we've built to catch when someone has forged the court audio recordings. Its very important to keep the legal evidence safer and unexposed to the other side. In that case this tool make sure nobody's been working with the audio files that could make or break a case in an unethical way. The system is very smart about spotting the different types of tampering. It knows whether some parts are cut, added new recordings to the existing, spliced multiple recordings together, speed up things or slowed them down or even just the compression of the audio file can also be detected. Justice Assistant can usually tell or indicate when something is not right. It works in such a way that the system takes audio files through a standard cleanup process, provides lot of detailed information from the sound, and then uses analysis and algorithms from signal processing to figure out what the original is and what might have been edited. It also gives analysis and score for each section of tampering detected and the differences made in it so that it is easier for the person to analyse and know the real issue in the tampered audio. It can work in two different ways either all received recordings can be analysed and asked to generate report in bulk or the necessary audio particular for the case can be analysed while court is actually happening without disrupting anything. It detects unusual gaps which are more frequent and keeps track of all the technical details and generate reports which are more important for court. So with this tool assisting audio recordings and the tampering the experts in forensics can work on something else which need more time and crucial. The Justice Assistant is mainly used for the purpose of fairness and transparency in the court and the people.**

***Index Terms*—Court audio, audio forensics, tampering detection, machine learning, Judicial evidence safeguarding.**

I INTRODUCTION

Courts nowadays use evidences as their key part of the judgement. These evidences involve Video recordings, Audio recordings and other documentations for their proceedings. So these evidences play a major role in the judiciary. Since these recordings play a huge role and as they are digital they can easily be tampered and manipulated. As there are lot of tools available today it is easier for people to tamper any sort of digital evidence. So the traditional way is less efficient compared to the modern tools available. So Courts expect these tampering to be captured soon and need to manage evidence efficiently.

The recent review progress states the move of audio procedures from old statistical methods to the new modern systems based on speak embeddings. This new system have improved accuracy in detecting speech in recordings. This deals with the overlapping of speeches and short speech segments but has limited real forensic training data[1]. There are some old methods where both face and voice recognition systems are

used. The voice recognition part in this helps when the audio

recordings are available. They show that using this could improve accuracy better and can be used for strong forensic verification[2]. Some deep learning models are used for audio recordings to detect and analyze the part of the audio. These helps in identifying pattern involved from the audio by using the spectral features. These can track the problems present such as editing, splicing, or any sort of manipulation of recordings. This improves detection accuracy in different tampering scenarios[3]. The common types of audio tampering involve splicing, cut copy paste and and deletion or silencing the audio recordings. In this comparison of features, datasets and methods takes place for improvising. This deals with compression and noisy backgrounds[4]. There are also

methods for detecting audio splicing using deep learning models. This uses convolutional neural networks with spectral audio features. The method can detect the area of splicing and shows improvised performance[5]. There are also simple methodologies that can detect small changes in the audio that can affect the entire result of the recording analysis. Even these small changes can overcome deepfake detectors and can involve tampering. So different attacks of tampering must be studied to ensure there is everything to fix it. So these are done in this deep fake audio detection system[6]. So by combining all the above approaches and methodologies this efficient audio forensic and audio tampering Justice assistant is developed.

II RELATED WORKS

As courts have been going digital really fast, there's been a lot more interest in creating tools that can check if audio evidence is actually real and hasn't been tampered with. The work done by us in this is a combination of both traditional audio forensics methods and AI based systems that can spot fake recordings, and also provides what sort of tampering is done in the audio provide as evidence. The traditional way involves analyzing sound patterns by experts whereas this new method combines analysis and algorithms to automatically find suspicious parts of recordings. But the thing is, both methods are good in their own ways and can be efficient, but the time taken is less in the new method. All of this previous work is what Justice Assistant builds on, but it also shows where there are still big gaps. Together, by combining all these, we have built the foundation of the Justice Assistant and made it look more time-efficient for this audio forensics. The usual time that the court and the audio forensic laboratory take to analyze the audio can vary from 45 days to 6 months. Even in sensitive cases, it can go up to 15 days. But with this, the whole detection and analysis is completed in minutes, and the final report can be submitted to the court in a few hours. Most tools that are present now cannot work or detect the audio recordings efficiently due to its noisy background and other external factors. Plus, there aren't many tools for audio tampering that can do both content analysis and identify the anomalies present in the audio.

Classical audio forensics tools work manually by identifying the sound patterns, noisy backgrounds, and other differences. The experts should look in a proper and in a careful way at the suspicious areas that tampering can occur and compare the adjacent segments of the audio and spend more time in checking a single recording to identify splices, re-recordings and other factors that can cause a real problem to the audio in an unethical way. Such tools are generally deployed by experienced forensic specialists and using the outputs of those software tools to provide expert opinions to the court. These tools help the experts to analyze and work on inquiries, but they are very difficult to analyze when courts produce high volumes of audio day after day. These are not automatic analysis or identifying systems and require a huge time for complete analysis by the experts[9]. Justice Assistant builds on these principles and is used in the process of automatic analysis with the tools and functions provided in that and also requires manual work only for the most difficult or tedious process of audio analysis.

By using DL models along with the audio tampering detectors they can easily analyze the waveform patterns and the frequencies present. These models become familiar with features that can distinguish and that can accompany actions such as cut and paste, recompression, synthetic modifications and can even perform various audio segment level tampering. These detectors could provide better results in a large number of research models that are relevant to hand composed features on datasets. Most systems are developed and evaluated on controlled conditions, such as using clean speech and laboratory recorded datasets. The real-life courtroom recordings add extra complexity, like recording or speeches that overlap with each other, differences in the communication levels, different microphones and background noises[11]. Justice Assistant further refines the deep learning and addresses these difficulties and methodologies by using preprocessing and domain specific training algorithms to stay robust under realistic judicial conditions.

Evidence chain management systems purpose is to ensure privacy and keeping the evidence safely without any audio or digital recordings getting tampered. They generally use cryptographic hashes, secure logs and shows the key details of the document or evidence such as the creation, modification or any changes made in it after the stipulated time. Some

necessary solutions include watermarks or a legal stamp of the particular owning to hinder tampering with metadata or file history. Although they can document effectively who did what with the evidence of any particular case and they probably do not analyze the audio content. So a recording will have a neat custody record but still it can contain different manipulations and these can be introduced before it gets into its system[12]. Justice Assistant is designed in a way that it could analyze and find all these in a stipulated time, making it valuable evidence handling system in court.

Table. 1. Comparison of existing detection systems

Features	Classical Audio Forensics Instruments	Deep Learning–Based Tampering Detectors	Evidence Chain Management Systems
Primary Focus	Manual identification of edits using spectral, noise, and device artifacts	Automated detection of manipulations using learned time–frequency patterns	Preservation of file authenticity and custody history
Analysis Method	Expert-driven inspection of spectrograms, phase discontinuities, and segment comparisons	CNN/RNN models analyzing spectrograms and audio r	ryptographic hashes, secure logs, access control
Level of Automation	Low (fully manual, expert-dependent)	High (automatic or semi-automatic scoring)	High for tracking, none for content analysis)
Scalability	imited; time-intensive per file	High; suitable for large audio volumes	High for file management, not analysis
Output	Expert opinion and qualitative findings	Segment-level tampering probabilities	Audit trails and chain-of-custody reports
Strengths	High interpretability; legally trusted expert review	Strong detection performance on benchmark datasets	Strong integrity assurance and accountability

Table 1 is an overview of how the existing audio tampering detection and existing evidence handling systems perform and the differences between them are shown above. So even when many advances and upgrades have been made it still continues to be a challenge especially in the process where the data is a necessity and is very crucial for court cases to be dealt with and it is needed to perform and provide quick insights about the process taking place for efficiency in the analysis of the audio tampering detection.

III. PROPOSED METHODOLOGY

The digital audio recordings are widely used in the courtrooms to detect the tampered audio from the original audio. The recordings in the court audio are facing major problem that are to be altered or changed. With more AI software's the original audio can be edited, altered & modified completely. The small change in the audio will be the turnover point in the court regarding the justice provided by the judges. Because of the AI software's the court faces major issues to detect the original & tampered audio as evidence. To check whether the audio is original or tampered requires a lot of time & process, meanwhile the justice will be denied in the court room. If the tampered audio is not detected, then the justice will be in favor of the criminal, and it leads to unfair justice in the court room.

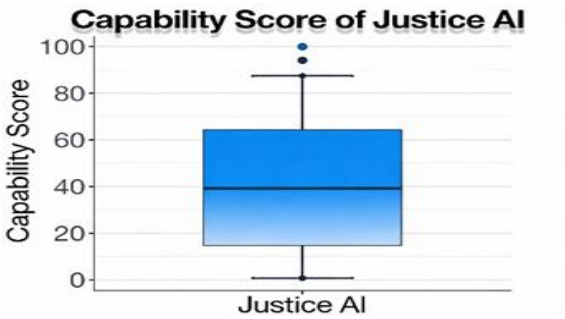

Fig. 1. Capability Score of Justice AI

The fig. 1 shows that the audio methods from the traditional techniques are not suitable for the modern courts or in the AI generation. The method scores only 30 which symbolizes that the traditional method will not handle the audio recordings with more durations. The traditional method only gives limited support and it takes more time to process the given audio. While audio that are in courtroom may contain background noise, in traditional methods it doesn't have the noise cancellation and advance software requirements in the

modern era. At max, it only scores up to 40, and by manual checking of the audio it only gives the score of 25 because it depends on the human and it takes more time.

The New Justice AI Model will work very Fastly & performs more score compared to the traditional audio tampering methods. The Model will handle longer duration recordings without more time consuming which means it performs without difficulty. It scores up to 90, which shows that it can be used in the live court proceedings and detect the tampered audio instantly. The new model will be suitable for the modern environment even though the audio is not so perfect. The model has the high score of 85 in the court room management. This shows that the model can easily be used in the courtrooms and it will make easier for the judges to identify the audio is tampered or not. Overall, the traditional model has low score and it will be not suitable in the court room recordings. So, the new AI Model improves the score and accuracy of the audio recordings. It will also increase the performance and reduces the time consuming.

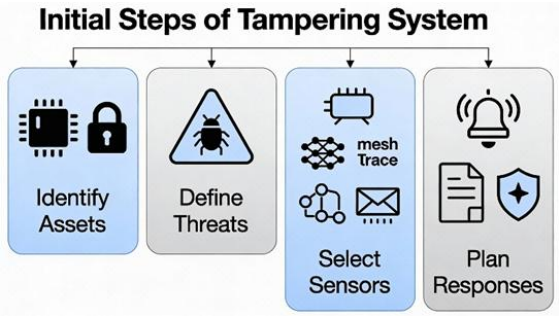

Fig. 2. Initial Steps of Tampering System

The fig. 2 explains the initial process of the audio tampering system from the initial to the final stage. The process starts from identifying the audio and check the frequency of the audio, at the initial stage, the model only listens to the audio recordings. It begins from hearing in the microphones in the courtroom. The next stage is the transmission stage, in which the recordings are transferred to the selected sensors without no loss in the quality of the audio. After the audio reached in the selected sensors stage, the model will check the audio carefully. In this process, the sensors will clean the audio, remove the noise in the background and it extracts the important features that are in the audio. The final stage is the decision-making process, the model shows the final process whether the audio is original or tampered.

In the processing of the data stage, the model will clean the audio, extract the different speech sounds with frequency, remove the noise in the background by noise cancellation, also it gives the important parts that are present in the audio. In the decision making process, the audio will be tested to the maximum potential, the model checks the sudden sound cut, repeated audio used, different waveforms in the same audio. Based on this testing, the model will provide and show whether the audio is original or altered. By working with this we can ensure safety of the audio recordings and can easily identify tampering related to the cases. These can provide the results in a very efficient way that everyone can understand instead of a report submitted to the court. This reduces time, efforts and a large manual work done by the experts and provide more efficient results with the indication of tampering.

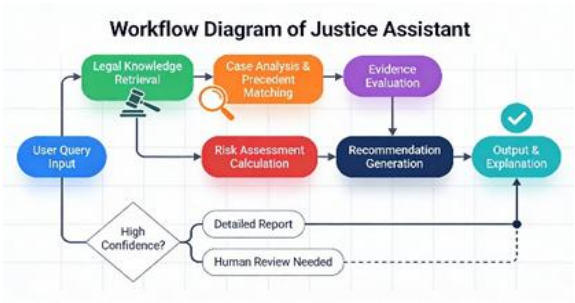

Fig. 3. Workflow diagram of the Justice Assistant

In Fig. 3, It shows the workflow diagram of the AI model, The first step is to get the input from the user query, then it is divided into two parts. The first part is the legal knowledge retrieval. From that process it goes deeper into the case analysis and audio analysis that are given by the user. Also, it checks the evidence whether it's is original or tampered audio. There's a part called risk assessment calculation, in which it shows the risk of audio that is tampered in the original audio. The final process is the output and explanation, in which it shows the detailed output with valid explanation and shows the audio is tampered or not. If the recording has the high duration and more voices involved, the noise filtering is used to mute the background noise and the model also has noise cancellation. Then the audio is converted into many parts such as frequency, waveforms and other components. By this conversion we can identify whether the audio has unusual modifications in it. The

conversion parts are then processed to the audio tampering model. The model will analyse the data as an audio and it focus on the unusual cut in the audio, and different waveforms that are produced in the recordings.

The Justice Assistant system for audio detection primarily starts with uploading the audio recording we want proceed with. Then the system with the developed and incorporated algorithms and techniques collects all the speech segments and data from the recording. The system then evaluates the audio and provides the insights such as unusual gaps, change in tempo, edits in between and any other compression. These detected problems can cause a huge difference in court hearings and the proceedings. Therefore the spectrogram and waveform frequency of the audio is analysed for any such tampering and the duration of the particular tampering is also mentioned. Then the tampering is flagged with a redline in the waveform and in the spectrogram. With combining all these the final result is shown. This score is based on the combination of all the other factors. Based on the scores and the conditions defined the final output is displayed. This indicates whether the audio is original or tampered. This also shows whether anomalies are present and could differentiate between minor anomalies and major anomalies present in the audio recording.

IV. RESULTS AND DISCUSSIONS

The main process and the objective of this paper is to reduce the time that are consumed in the traditional methods such as processing the tasks. The below graph illustrates the complex queries that are the problems of the audio recordings if it's an tampered audio.

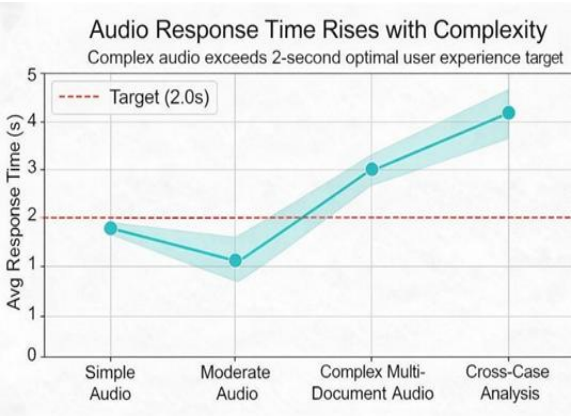

Fig. 4. Audio Responses Time Increases Diagram

In this figure, 4, The graph shows the audio response time rises with complexity in the different process in the audio tampering system.

The horizontal axis shows the different types of audio such as simple, moderate, complex multi document and cross case analysis and results. The vertical axis shows the average time response thar ranges from 0 – 5, the ideal response time for audio recordings is 2 seconds, which shows the good experience to the user. The results of the graph show the simple, moderate and complex audio that are answered by the model to different types of tampering. If the complexity increases in the model the response time to analyze the audio recordings is also increases and will reach the maximum limit as shown in the graph of the audio recordings.

The complexity mainly increases due to the larger duration, volume of the audio recordings and specific case details in the audio recordings. Overall, the graph shows the challenges that are faced by the AI model system across all the different audio recordings uploaded by the user. It also doesn't reduce the accuracy of the original audio and remove the time consuming by the traditional audio detection systems. The most important thing is that the audio file processing delays when the audio recording has got huge amount of tampering or it has got a longer duration. Therefore the time for that analysis is a little longer than we expected. But even after all this still the justice assistant for audio tampering is much efficient and less time consuming compared to all the other methods available.

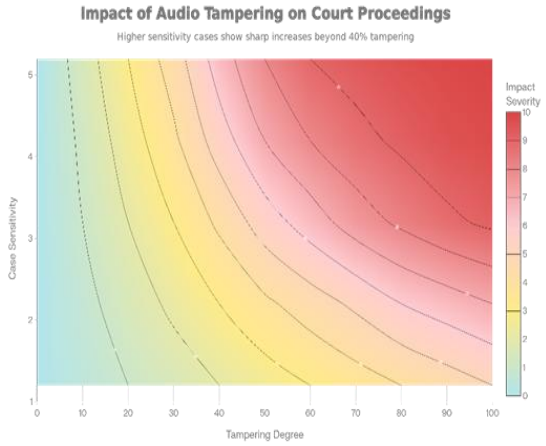

Fig. 5. Impact and severity of audio tampering on court proceedings.

This Fig.5 clearly shows us all about the impact made on the audio manipulation on the court proceedings with the help of contour-based heatmap based which gives us a clear view on the amount of audio tampering. This also tells us about the increasing levels of audio tampering and the increased sensitivity in legal cases which may risk in such manipulations

The color scale in the picture shows that how severe the impact is and also the cooler shades represents very little influence and red colors show more serious effects on judicial outcomes. Once the level of tampering crosses a certain amount of point which is around 40%, the damage starts to increase steadily. After this level the increase becomes much more sharper which makes the overall impact even severe.

This shows a that there is a certain limit beyond which the reliability of the audio evidence is seriously affected. When small changes are made it causes very little trouble in the case of low sensitivity, but in the highly sensitive cases even some moderate tampering can have a major legal effect.

The visualization tells that there is a need for an strong audio authentication and identification methods to safeguard the integrity of the evidence and to maintain trust in digital audio evidences used in court proceedings.

The sharp increase seen in sensitive cases shows that courts handling serious criminal or important civil cases are more at risk because of audio tampering. This tells us there is an need for real-time monitoring and simple forensic checks which also includes AI tools to find the issues at the early stage. By dealing the audio problems at the early stage the courts can reduce the evidence conflicts and also avoid the delays in the trials and make sure that the judgements are made using real audio records which builds more trust in the justice system.The impacts these evidence cause in the trail session is very important and these are the core system which influences the entire case. It not only serves for this particular case but also being an example for the other cases that could occur in the future.

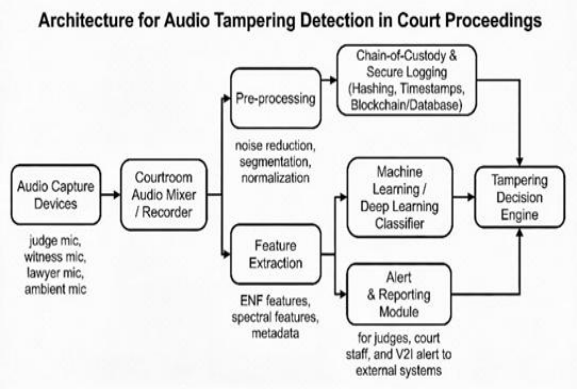

Fig. 6. Architecture of Audio Tampering Detection

This Fig. 6 shows that the framework given tells us about the structure of the Justice Assist AI system overall architecture, the diagram shows us that the workflow is divided into four main layers, which together handles the data and gives us useful information. The first part which is called as the evidence input layer is the part where different forms of information like documents or audio files are collected.

This layer acts as the starting point where all the evidence is properly collected and prepared before moving to the next step. once the input is being received it is been sent to the processing core to handle tasks like document parsing, transcription, and structuring of evidence.

It mainly changes the raw data into a form that is easier to analyse. The processing core makes sure that the raw data follows some of the standard pattern, so that the next module can follow and work on it a properly in a consistent manner.

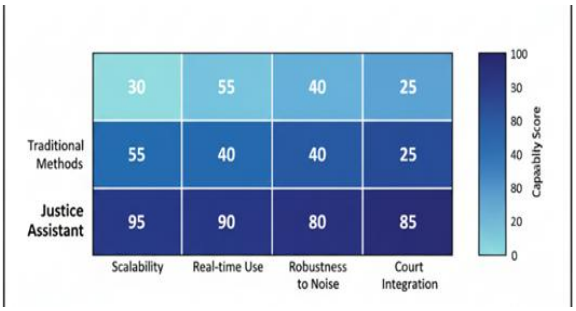

Fig. 7. Comparison of audio tampering detection

The comparison shown in the figure explains how traditional audio forensic techniques are not able to meet the increasing need of the modern court systems. These traditional methods still struggles

and face problems with scalability and get a low score of about 30 which shows they are not suitable for handling long audio recordings or large amounts of digital evidence.

V CONCLUSION

The checking of the courtroom evidence is now slowly being improved with the use of AI- based software which helps in checking audio integrity almost like real time. By connecting the recording systems early with the automated tampering detection tools this method helps in better reliability and creates safer and more trusted systems for the justice process. Justice Assistant is really great. It's great at spotting any weird stuff in audio files, so you can tell if they've been tampered with. This makes professional reviews a lot easier, and it can handle really lengthy recordings without any trouble. Basically, it builds trust in our digital justice system by keeping audio evidence secure and real.

Fig. 8. Oversight of Court room evidence

The Fig.8 shows the entire workflow of the Justice assistant that is easier for the non technical people or the stakeholders and other workers who need this in their working environment and moreover can be useful if interpreted visually. This also serves as a tutorial for the application of Justice assistant and explains how it works.

REFERENCES

[1] J.Villalba and N. Brümmer, “Towards speaker diarization for forensic applications: A review of recent advances,” in Proc. Odyssey Speaker and Language Recognition Workshop, 2020.

[2] P. Korshunov and T. Ebrahimi, “Using face and voice biometrics for identity verification in multimedia forensic applications,” in IEEE Int. Conf. Advanced Video and Signal Based Surveillance (AVSS), 2017.

[3] M. Ammar, R. Kinnunen and T. H. Nguyen, “Audio tampering detection using deep spectral features,” in Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP), 2021.

[4] A. de Souza Brito et al., “A survey on audio tampering detection for forensic applications,” IEEE Access, vol. 9, pp. 122345–122365, 2021.

[5] S. Wu, H. Wang and Q. Jin, “Audio splicing detection and localization with CNN-based spectral features,” in Proc. Interspeech, 2019.

[6] A. Neekhara, P. Perera, R. Kumar, S. et al., “Adversarial threats to deepfake detection: A case study in audio,” in Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP), 2020.

[7] A. Nautsch et al., “Speech forensic authentication with modern speaker recognition: A survey,” IEEE Signal Processing Magazine, vol. 37, no. 6, pp. 82–94, 2020.

[8] H. Malik, “Digital audio authentication using background noise,” Digital Investigation, vol. 8, no. 2, pp. 146–155, 2011.

[9] H. Fayyad-Kazan, M. Khalil, M. Azzi, and A. Sarkis, “Verifying the Audio Evidence to Assist Forensic Investigation,” *International Journal of Computing and Digital Systems*, vol. 10, no. 1, pp. 1–12, 2021.

[10] A. Kraetzer and J. Dittmann, “Digital audio forensics: Challenges and opportunities,” in Information Hiding and Multimedia Security,

Springer, 2016.

[11] S. T. Tisza and Z. H. Tan, "Convolutional neural networks for detecting MP3 double compression in audio forensics," in Proc. IEEE Int. Workshop Information Forensics and Security (WIFS), 2019.

[12] S. Nath, K. Summers, J. Baek, and G.-J. Ahn, "Digital Evidence Chain of Custody: Navigating New Realities of Digital Forensics," *in Proc. Digital Forensics Conf.*, 2024. This paper explores theoretical foundations and modern implementations for preserving evidence integrity throughout its lifecycle in digital forensics.

[13] S. Agarwal and H. Farid, "Detecting visual and audio deepfakes: A survey," ACM Computing Surveys, early access, 2023.

[14] C. N. Evans et al., "Speaker recognition in forensic science: From laboratory to court," Computer Speech & Language, vol. 71, 2022.

[15] M. Gunawan and T. Kinnunen, "Audio spoofing and countermeasures: A survey," in Proc. Odyssey Speaker and Language Recognition Workshop, 2018.

[16] M. Fontani, T. Bianchi, A. De Rosa, A. Piva and M. Barni, "Audio forensics: A comprehensive bibliography," EURASIP J. Information Security, vol. 2013, article 1, 2013.

[17] R. Yang, H. Zhong and X. Li, "Blockchain-based chain of custody for digital evidence," in IEEE Int. Conf. Trust, Security and Privacy in Computing and Communications, 2019.

[18] K. K. Patel and D. Das, "Secure logging mechanisms for digital evidence management," in Proc. IEEE Int. Conf. Cyber Security and Protection of Digital Services, 2018.

[19] C. Kraetzer, A. Oermann and J. Dittmann, "Digital audio forensics: A roadmap for future research," in Proc. ACM Workshop on Multimedia and Security, 2012.

[20] T. Gloe and R. Böhme, "The Dresden image database for benchmarking digital image forensics," J. Digital Forensics, Security and Law, vol. 2, no. 2, pp. 150–159, 2007 (referenced for general forensic benchmarking methodology).

[21] D. L. Sharman, "Standards and best practices for the forensic examination of digital audio," in Handbook of Digital Forensics of Multimedia Data and Devices, Wiley, 2015.

[22] Council of Europe, "Electronic evidence in civil and administrative proceedings: Guidelines," Strasbourg, France, 2019 (used for legal and procedural context of audio evidence).