Countering Deep fake Videos Using Blockchain

Siddhant Giri[1], Chaitanya Giram[1], Rajendra Kathale[1], Prof. Kshama S. Balbudhe[1]

[1] Information Technology,

[1] PVG's College of Engineering and Technology & G.K.Pate(Wani) Institute of Management , Pune, India

19113045@pvgcoet.ac.in[1], 19113037@pvgcoet.ac.in[1], 19113005@pvgcoet.ac.in[1], ksb_it@pvgcoet.ac.in[1]

Abstract - Fake digital content has become increasingly common in recent years due to the development of artificial intelligence (AI) and deep learning algorithms. Deep fakes, or fake footage, photos, audio, and videos, can be a frightening and dangerous phenomenon that has the capacity to distort the truth and erode confidence by creating a false world. To combat the plague of fake information, digital media must provide proof of authenticity (PoA). Available technologies do not yet support digital media provenance tracing and history tracking. The provenance and history of digital information may be tracked and traced back to their original source in this study using Ethereum smart contracts, even if the digital content has been duplicated numerous times.

Index Terms - Blockchain Technology, Digital Identity, Decentralization, Distributed Ledger, Cryptography, Authentication, Transparency, Consensus, Smart contracts, Immutable records.

I. INTRODUCTION

Deepfake videos have been made possible by the recent advances in AI, deep learning, and image processing. In April 2018, a minute-long clip of former US president Barack Obama went viral in which he was heard saying things he had never stated. Deep-fake videos can undermine the truth, perplex viewers, and accurately mimic reality. They are harmful. The spread of such content can become uncontrollable with the development of social networks, which could worsen issues with false information and conspiracy theories. When deep fake videos first started to target celebrities, they were not a major worry. Deep fake videos, according to Stover and Floridi, are a data calamity. They advocated encouraging people to use new technologies wisely and to post morally and responsibly on social media. Deeply fabricated digital stuff, which might include phony films, photographs, paintings, audio, and other types of content, must be identified, fought against, and countered using specific tactics. If there is a reliable, safe, and trusted mechanism to track the history of digital content, then achieving this goal is not difficult. Users should have access to trusted data about the digital content's provenance and be able to trace an item's origins back in time to confirm its legitimacy. Innumerable enterprises, industries, and domains are poised to be transformed by blockchain technology, including those in finance, the food industry supply chain management, health management, and IoT, to mention a few. In a way that is decentralized, highly trusted, and secure with tamper-proof records, logs, and transactions that are either openly accessible to all in the case of permissionless blockchain or restricted to certain participants in the case of permissioned blockchain, blockchain technology can provide key features that can be used to prove the authenticity and originality of digital assets The public or permissionless blockchain is best for deep fakes. The public Ethereum blockchain, which uses smart contracts to control and record the history of transactions involving digital material, serves as the foundation for our proposal in this paper. Using this approach, we suggest a generic framework and a blockchain-based method for proving the authenticity of digital assets, such as movies, audio, photos, etc.

II. LITERATURE SURVEY

A US-based start-up company called Truepic has developed a system involving mobile apps for typical users and freelancers for capturing images and saving them to the company's servers. The purpose of saving the images is to preserve their integrity. Hence, any forgery attempt can be easily discovered by comparing it with the image from the servers. They hope that in the future their technology will be used in collaboration with other social media parties that will verify any uploaded images with the images in Truepic's servers and any change would, therefore, be detected.

Li and Lyu proposed a method to detect deep fake videos using Artificial Intelligence (AI). The proposed method depends on an AI algorithm fighting another AI algorithm. Their technique relies on training convolutional neural networks (CNN) with manipulated and real figures. Gipp et al. focus on using blockchain to only ensure the integrity of video content. Their approach depends on hashing the video and securing the hash on the immutable blockchain. Any manipulations on the video will result in a mismatch in the hash.

III. METHODOLOGY

(1)Blockchain Technology

Blockchain is a decentralized, tamper-proof ledger that can be used to store and verify data. The technology is known for its immutability, transparency, and security features.

(2)Smart Contracts

Smart contracts are self-executing contracts with the terms of the agreement between buyer and seller being directly written into lines of code. They can be used to automatically enforce rules and conditions to ensure compliance and prevent fraud.

(3)Decentralized applications (DApps)

DApps are applications that run on a blockchain network, and they can be used to enable secure, decentralized processes for a variety of use cases.

(4) Cryptographic hashing

DApps are applications that run on a blockchain network, and they can be used to enable secure, decentralized processes for a variety of use cases.

(5)Digital signatures

Digital signatures are cryptographic techniques used to verify the authenticity of a digital message or document. They can be used to ensure that the person or entity sending a message or document is who they claim to be.

By using these concepts, blockchain technology can be used to combat deepfake videos by creating a decentralized, tamper-proof system for verifying the authenticity of video content.

For example, a blockchain-based system could use smart contracts to automatically verify the authenticity of a video and ensure that it has not been tampered with. The system could also use cryptographic hashing and digital signatures to ensure that the video's origin is verified and that it has not been altered in any way.

- The following diagram shows how each block corresponds to a fresh addition of video data to the blockchain. A unique hash value found in the genesis block, the very first block in the blockchain, serves as the chain's starting point. Because each block in the chain has a hash value dependent on the hash of the one before it, the chain of blocks is resistant to manipulation.
- It is conceivable to develop a decentralized, tamperproof system for certifying the authenticity of video material by employing a blockchain-based system to store video content. Deep fake videos and other internet misinformation can be stopped by using decentralized access controls and cryptographic hashes to further strengthen the system's security and resilience.

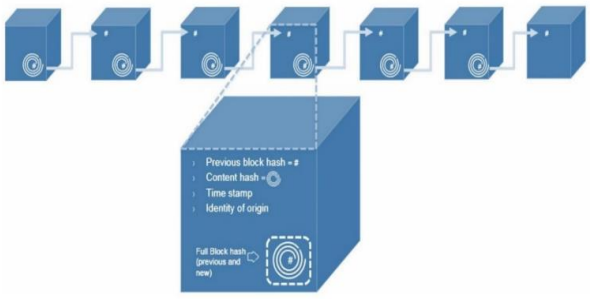

Fig.1 Ethereum ledger

IV. SYSTEM OVERVIEW

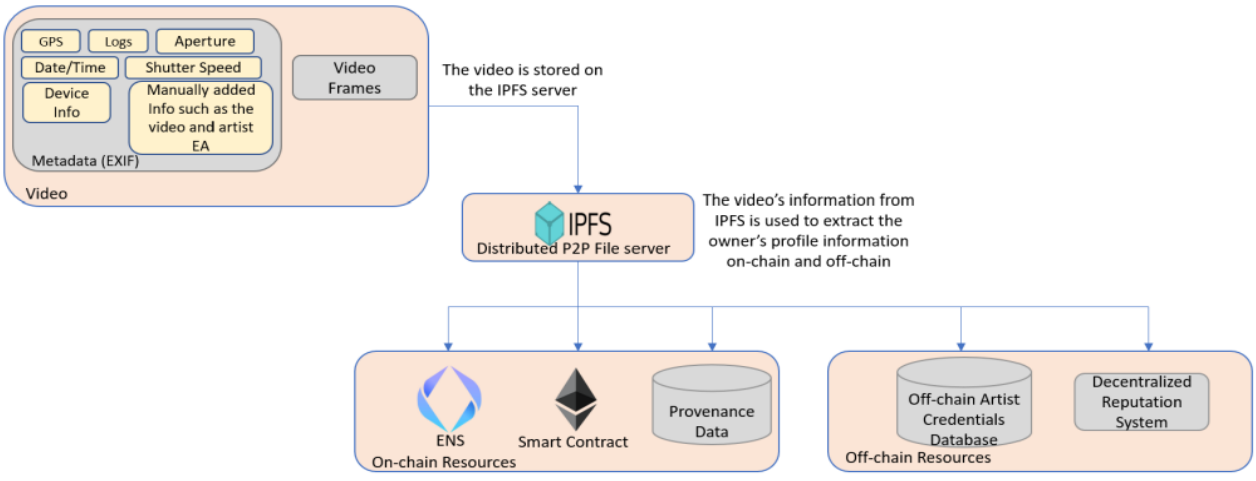

Fig.2 Proposed Solution

All the participants in our system, including the primary artist and any supplementary artists, have Ethereum addresses. We go on to define important system parts as seen in the figure.

(1)Video

As seen in Figure 2, a video contains significant information in addition to the video frames. In a (Exchangeable Image File) EXIF format, "Metadata" refers to the key characteristics of a video. The metadata of a video includes details about the camera used to record it, its settings, the time and date it was recorded, as well as any logs and manually added data that the producer of the video may have added. Each video will have a corresponding Ethereum smart contract, which can be made by a news organization or an artist. The artist's Ethereum address and the smart contract's address are both essential components of the metadata.

(2)IPFS storage

The movie and the metadata it contains are kept on a decentralized, peer-to-peer, content-addressable file system called the InterPlanetary File System (IPFS). An address for a collection of files including the video content and associated metadata is generated by IPFS as a special hash. To find and retrieve the collection of files kept on the IPFS network, utilize the hash address. A file with the terms and conditions agreement for copying and modifying may also be included in the IPFS package in the event that the video is to be duplicated and used by other authors or artists to produce new content. The smart contract would also make use of the IPFS hash produced from the saved form.

(3) On-chain Resources

Upon the creation of the video's IPFS hash, the original creator (owner) creates a smart contract on the Ethereum network. The contract includes variables and attributes for recording the information about the owner and the video specifics. Other secondary artists can use these functions to ask for permission to share, edit, and distribute their work in accordance with the terms and circumstances of the agreement form.

The smart contract also includes modifiers that limit access to the procedures according to roles or contract states. Events are also used to create notifications and maintain logs of significant requests and results. Moreover, static data like video-related data and the contract state are stored in variables.

A connection to the original video will be included in each altered video created by a second artist, which will have its own smart contract. Hence, any altered videos of any original video are considered "child" videos and are listed in the smart contract for the original video. Hence, utilizing on-chain resources like the smart contract that has a list of all the children's videos' smart contracts as well as a link to their parent's smart contract, a user may easily track a video back to its source. Because of this, the user can track with great transparency using this data combined with the logs and notifications made on the ledger and editing in the event that the film is stolen and used by other writers or artists to generate new material. The smart contract would also make use of the IPFS hash produced from the saved form.

(4) Off-chain Resources

Users can view the off-chain resources that are a component of our suggested solution when they trace data. As shown in Figure 2, it is also possible to connect a video's smart contract Ethereum address and the owner's Ethereum address to an off-chain credentials database. The owner's information and a link to their ENS profile are both included in this database. In order to give a complete profile and examples of the artist's work, it will also include information about other videos the video owner owns.

V. PROPOSED ARCHITECTURE

In order to accomplish our goals for video content security while operating on the blockchain as clear, System suggested a Smart Contract-based solution that can serve as a heart in the fight against deep fake. In these networks, blocks are created by consensus techniques, and validation of a block necessitates a majority vote. Furthermore, due to many signature transactions in the original data, when using blockchain technology, numerous identity stages are needed for ultimate certification. Metadata and video frames are the main focus of a smart contract. Metadata can include information such as the time and date of recording, speed, pertinent logs, GPS, device information, and other things in addition to the frames and information in the video. After a hash for a video is generated and it contains all the crucial information as well as video metadata, a smart contract is launched on the blockchain. Others have the right to alter and edit this video in accordance with the terms and conditions. For that video, the smart contract can also set roles or restricted access.

(1)User

Users could upload their own videos to the platform using the video uploader. Users would be prompted by the uploader to enter pertinent details about the video, such as the title, description, and genre. Additionally, the uploader might look for any known deepfake indicators in the video file and ask the user to validate the legitimacy of the video

(2)Ethereum Blockchain

The development of decentralized digital identification systems is one way Ethereum may be used to counteract deep fake videos. It is conceivable to build a safe and impenetrable identification system that might be used to confirm the legitimacy of video material using Ethereum's smart contract features. This may include giving each user a distinct identity that is connected to their Ethereum wallet address. This identity could then be used to authenticate video content that they produce or show that they have given permission for the use of their image in a deep fake film. As a final possibility, a system for tracking and regulating the distribution of video material may be developed using Ethereum's capacity for smart contract execution. To make it simpler to spot and deal with deepfake films, this may involve tracking a video's initial source as well as its distribution and consumption over time.

(3)IPFS

By developing a decentralized framework for storing and distributing video material, IPFS may be used to battle deepfake films. Creating a platform that is impervious to censorship and manipulation using IPFS makes it harder for criminals to produce and disseminate deep fake videos. By dividing files into smaller pieces and storing them across a network of nodes, IPFS makes it more challenging for one party to maintain control over the entire system. Another conceivable use case for IPFS in countering deepfakes is through the construction of a system for tracking and monitoring the dissemination of video material. To make it simpler to spot and deal with deep fake videos, this may involve tracking a video's initial source as well as its distribution and consumption over time

(4)Smart Contracts

By developing a tamper-proof digital identification system that may be used to confirm the authenticity of video material, smart contracts may be utilized to battle deep fake videos in the future. It is feasible to authenticate video content users produce or confirm that they have given their consent for the use of their likeness in a deep fake video by giving each user a distinctive digital identity that is linked to their blockchain wallet address. Smart contracts might be used to enforce these constraints, ensuring that only permitted material is disseminated on the blockchain. The development of a platform for content verification is another potential use of smart contracts in the fight against deepfakes. Before video footage is published to the blockchain, it may be validated by a network of validators that work as part of this platform. Validators might be motivated to engage in the platform through the usage of tokens, which could be used to compensate them for their efforts.

VI. RESULTS

The specifics of testing the smart contract utilising Remix IDE's in-browser development and testing environment are presented in this section. The section discusses important function testing while displaying the relevant outputs and logs.We consider two parties engaging with the smart contract in our testing scenarios. The Ethereum address (EA) for the primary artist is "0x5B38Da6a701c568545dCfcB03FcB875f56beddC4," while the EA for the secondary artist is "0xAb8483F64d9C6d1EcF9b849Ae677dD3315835cb2."

Every function has a condition that has been satisfactorily tested or a state need. If any of the requirements are not satisfied, the smart contract's state returns to its initial state. In our test scenario, a smart contract with the address "0x1439818dd11823c45fff01af0cd6c50934e27ac0" was produced by the secondary artist. The newly generated contract's parent information was also updated using the smart contract's features. This contains crucial details regarding the primary original video as well as the parent's smart contract address, which in this case is "0xd9145CCE52D386f254917e481eB44e9943F39138."

(1)Requesting permission


```
[
{
    "from": "0xd9145CCE52D386f254917e481eB44e9943F39138",
    "topic": "0x85ace08f038481adc2bfd0f855a5103eed84d76d1889e2837c00a5f0b64dd50b",
    "event": "ArtistRequestingPermission",
    "args": {
            "0": "0xAb8483F64d9C6d1EcF9b849Ae677dD3315835cb2",
            "artist": "0xAb8483F64d9C6d1EcF9b849Ae677dD3315835cb2"
    }
},
{
    "from": "0xd9145CCE52D386f254917e481eB44e9943F39138",
    "topic": "0xeb56067adf2919f9de03c1d1b99d640dc3897790d2af7784e06c9299b66b27f0",
    "event": "ArtistRequestRegistered",
    "args": {
            "0": "0xAb8483F64d9C6d1EcF9b849Ae677dD3315835cb2",
            "1": "0x6b6b6f7766776a6600000000000000000000000000000000000000000000000000",
            "artist": "0xAb8483F64d9C6d1EcF9b849Ae677dD3315835cb2",
            "IPFS_Hash": "0x6b6b6f7766776a6600000000000000000000000000000000000000000000000000"
    }
}
]
```


(2)Granting permission


```
[
{
    "from": "0xd9145CCE52D386f254917e481eB44e9943F39138",
    "topic": "0x30f499dee5aca26bf88e6932509f2693b637e648ec36ac97034db8869f118639",
    "event": "PermissionGranted",
    "args": {
            "0": "Permission Granted to address ",
            "1": "0xAb8483F64d9C6d1EcF9b849Ae677dD3315835cb2",
            "info": "Permission Granted to address ",
            "artist": "0xAb8483F64d9C6d1EcF9b849Ae677dD3315835cb2"
    }
}
]
```

(3)Creating Child Contract


```
{
"string information": "newVideo",
"bytes32 hash": "0x6b6b6f7766776a6600000000000000000000000000000000000000000000001",
"address SC": "0xd9145CCE52D386f254917e481eB44e9943F39138",
        "address EA": "0x5B38Da6a701c568545dCfcB03FcB875f56beddC4"}
```


(4) Attestation request


```
[
{
    "from": "0xd9145CCE52D386f254917e481eB44e9943F39138",
    "topic": "0xb9b72052a89e6a2f938dcb2d2d2868ff8fbb9343f1b6a9573626813554cea783",
    "event": "AttestationRequest",
    "args": {
        "0": "Address of child: ",
        "1": "0x1439818DD11823c45fFF01aF0Cd6c50934e27Ac0",
        "info": "Address of child: ",
        "Scaddress": "0x1439818DD11823c45fFF01aF0Cd6c50934e27Ac0"
    }
}
]
```


(5) Attestation grant request


```
[
{
    "from": "0xd9145CCE52D386f254917e481eB44e9943F39138",
    "topic": "0xb9b72052a89e6a2f938dcb2d2d2868ff8fbb9343f1b6a9573626813554cea783",
    "event": "AttestationRequest",
    "args": {
        "0": "Address of child: ",
        "1": "0x1439818DD11823c45fFF01aF0Cd6c50934e27Ac0",
        "info": "Address of child: ",
        "Scaddress": "0x1439818DD11823c45fFF01aF0Cd6c50934e27Ac0"
    }
}
]
```


VII. CONCLUSION

The primary goal of the suggested approach is to ensure and build confidence in the capacity to establish traceability to the original video maker or source in a decentralized fashion. Decentralized reputation system, Ethereum name service, and IPFS. Our suggested framework for a solution, as well as the system architecture, algorithms, sequence diagrams, implementation, and testing information, are sufficiently general to be used with various kinds of digital material, such as audio, video, photographs, and manuscripts.

VIII. FUTURE SCOPE

This model can offer crucial elements that may be used to demonstrate the uniqueness and validity of digital assets in a decentralized manner that is extremely trustworthy and safe thanks to Blockchain technology. This blockchainbased system offers evidence of the validity of digital assets, such as audio, video, and photographs, among others. In the event of a permissioned blockchain, the model will offer tamper-proof logs, records, and transactions that are either confined to a small number of participants or made publicly available to everyone in the case of a permissionless blockchain.

IX. REFERENCES

[1] Y. Li and S. Lyu. (2018). ‘‘Exposing deepfake videos by detecting face warping artifacts.’’ [Online].
[2] Eth Gas Station. Accessed: Jan. 2, 2019. [Online].
[3] Authenticity Verification of User Generated Video Files. Accessed: Jan. 2, 2019. [Online]. Available: [4] L. X. Hu and D. Evans, “Using directional antennas to prevent wormhole attacks,” in *Proc. IEEE Symp. Network and Distributed System, Security (NDSS 04*), San Diego; February 2004.
[4] K. Toyoda, P. T. Mathiopoulos, I. Sasase, and T. Ohtsuki, ‘‘A novel blockchain-based product ownership management system (POMS) for anti-counterfeits in the post supply chain,’’ IEEE Access, vol. 5, pp. 17465–17477, 2017.
[5] IPFS is the Distributed Web. Accessed: Jan. 2, 2019. [Online].
[6] Ethereum Name Service. Accessed: Jan. 2, 2019. [Online].
[7] K. Biswas and V. Muthukkumarasamy, ‘‘Securing smart cities using blockchain technology,’’ in Proc. IEEE 18th Int. Conf. High Perform. Comput. Commun.; IEEE 14th Int. Conf. Smart City; IEEE 2nd Int. Conf. Data Sci. Syst.