

Forensic Anatomy in Cyber Crime: Bridging Biological Forensics and Digital Investigations for Deepfake Authentication

Payal Panda ¹, Sai Aditya ², Krishna Sravya Chirla ³

¹Department of Anatomy C U Shah Medical College Surendranagar, Gujarat, India

²Department of Computer Science Engineering Anil Neerukonda Institute of Technology and Sciences Visakhapatnam, India

³GITAM Institute of Medical Sciences and Research Visakhapatnam, India

Abstract- With the rise of Generative AI models that are achieving near-pixel perfection, the traditional digital forensic methods, which rely on noise analysis and metadata, are becoming increasingly inefficient. Cyber-criminals and scammers are now using Generative AI models to generate synthetic media, which can be used for identity theft, spreading false information, and financial fraud. This advancement in the digital world requires us to shift from digital-only detection to anatomical or biological validation. This research introduces a safety framework that bridges the gap between clinical forensics and digital investigations. Real human faces exhibit muscle movements and heart pulses, unlike deepfake faces, which enable us to use the human body as a biological watermark. We use remote Photoplethysmography (rPPG) to observe and extract the pulse signals from the forehead and the cheeks, while a clinical audit led by anatomical experts evaluates musculoskeletal synergy to identify the biological impossibilities present in the synthetic media. This safety framework proposes a physiology-informed forensic framework for legal and investigative usage, providing a chance to combat cyber-criminals and scammers.

Keywords— Bio-Digital Forensics, Deepfake Detection, rPPG (remote Photoplethysmography), Clinical Anatomy, Safety Framework, Musculoskeletal Synergy.

I. INTRODUCTION

Generative Artificial Intelligence, also known as GenAI, is a branch of Artificial Intelligence that takes advantage of neural networks, specifically generative adversarial networks, to create synthetic data. GAN, also known as Generative Adversarial Networks, is a deep learning framework that was developed by Ian Goodfellow in 2014 to generate synthetic data by training two neural networks [1]. In this, the neural networks, which are the generator and the discriminator, play a zero-sum game in which the generator creates data (images, audio, videos, etc.) and the discriminator will evaluate that data and will let the generator know if the data is real or synthetic. This forces the generator to improve the quality of the data. The generator's aim is to minimize the probability of identifying fakes, while the discriminator's aim is to maximize the probability of

identifying fakes. This continues till the generator creates data that cannot be distinguished from real data. This is called the Nash Equilibrium.

Although Ian Goodfellow developed GANs for the development of generative activities, they have since been misused for malicious activities, such as synthetic data like deepfake videos. The cyber-criminals are taking advantage of the stage called "Nash Equilibrium". When the synthetic data (images, videos, audios, etc.) reach this stage, it is inefficient to detect them using the traditional digital forensics tools. Using such "near pixel-perfect" data can lead to identity theft, financial fraud, and social engineering [2]– [4]. By using such tactics, the cyber-criminals are able to steal personally identifiable information (PII), which can be used for account takeovers, by-passing multi-factor authentication (MFA), credential stuffing, etc. The cyber-criminals are using these deepfake videos and audios to bypass the "human-in-the-loop" security, which is

trusted in today's world. This creates a trust vacuum where the digital evidence (video, audio, image, etc.) cannot be verified by digital verification alone.

This is where we propose our "Anatomical Safety Framework", which focuses on Video-Based Deepfakes. This research specifically targets high-fidelity video deepfakes. It detects the involuntary muscle movements and pulse signals in a human body and concludes whether the human in the video is real or fake. This is achieved through two pillars—rPPG (Remote Pulse Detection) and anatomical audits of muscle detection. This research work provides a "high-confidence verification" in the digital world, where deepfake videos can be identified more efficiently.

II. LITERATURE SURVEY

In the past, deepfake videos used to have "glitches" like blurred edges, inconsistent lighting, or the wrong number of eyelashes, etc. [4]. These synthetic media can be detected using the traditional digital detection tools. But due to Generative Adversarial Networks (GANs), these deepfake videos have improved and thus have no "glitches", which renders the ability of the traditional detection tools to detect such synthetic media. The traditional detection tools are not reliable anymore as they only detect the media data. This asks for a change from the digital detection solutions to anatomical verification. This

kind of shift is actually supported by the traditional forensic doctrines like Locard's Exchange Principle [5], which states that whenever two objects come into contact, a transfer of material takes place between them. It also states that it is impossible to leave the crime scene without leaving evidence like hair, blood, tool marks, cosmetics, etc., behind, which ultimately helps the forensic investigators to link the crime to a person who is responsible for it. We can link this principle, by Dr. Edmond Locard, to digital evidence too. Here, the digital transfer refers to the transformation of physiological signals to visual representations during the generative AI render. Usually, information like pulse signals and the skin induced by it, and also micro-muscle dynamics, is included. These are difficult to replicate consistently

due to reasons like model smoothing and compression effects.

Recent studies done in 2024 utilized the rPPG (Remote Photoplethysmography) to find the heartbeat by detecting the micro-changes that happen in the human body in the provided media [6], [7]. These changes in the skin happen due to blood flow. While the recent studies have shown great advancements in rPPG, it is only a "math problem" for now, as it was used to look for signals, but it doesn't understand the anatomy. The tools, although very well made, can be easily confused by light changes or the subject moving their head in the video (only reliable for lab-quality data). Looking at the current rPPG models, they are prone to "false-positives" as they lack clinical grounding [8]– [10]. The current models can be fooled by confusing a pulse with just video noise or getting fooled by a deepfake video that simulates a pulse.

The biggest problem in deepfake detection research is the isolation. Computer Science researchers are trying to solve the problem of deepfake data without talking to medical experts who understand the pulse signals or muscular movements better. Our research, on the other hand, takes the human face as a living body—not just looking at the surface but also looking at the muscle movements and pulse signals which are beneath the surface. The musculoskeletal system and the vascular system have to work in a certain way to actually perform their functions; we can take advantage of this to distinguish between real videos and fabricated videos. AI cannot perfectly mimic the uneven blood flow through the anatomical zones of the face, giving us an edge in detecting deepfake videos.

III. PROPOSED METHODOLOGY

A. Image Acquisition and ROI Selection

The first step is to acquire high-quality data for forensic analysis. Following the data acquisition, the analysis is not performed on the entire face, as there will be obstacles like prescription glasses or hair, which might hinder the ability to analyze it properly. Instead, the framework will select ROI, also known as Regions of Interest, such as the frontal region

(Forehead) and malar regions (Cheeks). These ROIs, especially the malar region, are chosen by medical experts in our team. The reason to select these ROIs is due to the high density of blood vessels, which are present near the surface of the skin.

Therefore, the Regions of Interest are critical for the “blood flow authentication” which is required for this research.

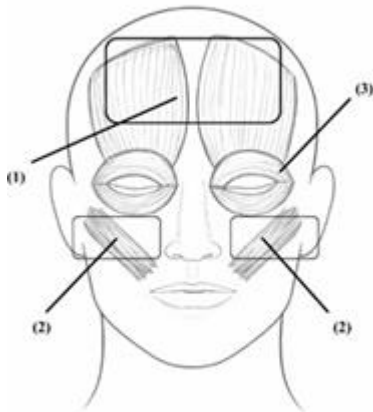


Figure. 1. Frontal view of the human face. (1) Frontalis, (2) Zygomaticus Major, (3) Orbicularis Oculi. ROIs: (1) & (2).

B. Vascular Pulse Extraction via rPPG

To capture a video, the camera uses the RGB color space, which is used to visualize the data. RGB (Red, Green, Blue) is one of two primary color spaces (the other one is YUV). For this research, the framework utilizes the green channel as it is the most effective channel to detect blood pulses in the human body. Once the video is captured, the green channel is extracted from the isolated ROIs (Cheeks and Forehead). The processing method used for this framework is CHROME (Chrominance-Based), due to CHROME’s ability to eliminate noise caused by movement. This is critical for this research as it allows the framework to track the pulse even if the subject is moving in the video (smiling, laughing, talking, etc.). The framework utilizes the green channel due

to the ability of hemoglobin in the blood to absorb green light more effectively than red or blue light. This results in the green channel having fluctuations in light intensity every time the heart pushes blood. These fluctuations help the framework determine if the subject in the video is real or fake, serving as a “biological watermark.” The BVP wave that is extracted in the output serves as a checkpoint providing an analysis of the pulsatile signal. Any deviation or lack of such signals can lead to the data being concluded as non-authentic.

C. Biomechanical Synergy and Anatomical Consistency Audit

The framework specifically evaluates the movement coordination between the Frontalis (forehead) and the Zygomaticus Major (Cheeks). As mentioned before, when a human face performs actions like smiling or talking, a group of muscles has to perform coordinated contraction to do the task. For

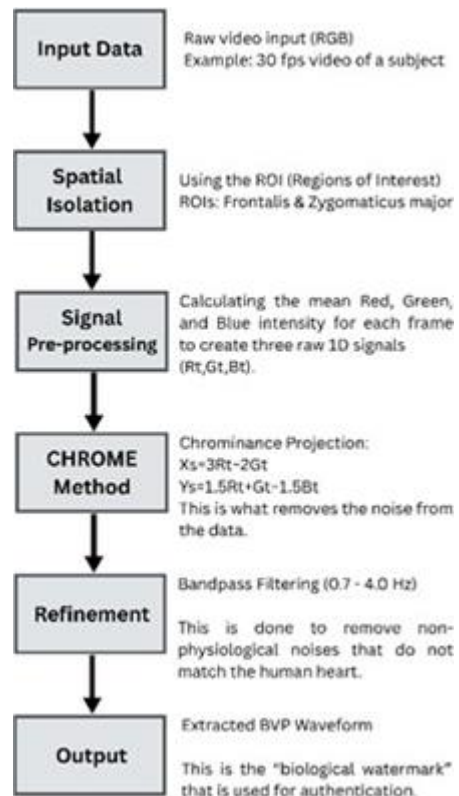


Figure. 2. The rPPG Processing Pipeline.

example, when a human smiles, their Zygomaticus major muscle and Orbicularis Oculi muscle coordinate with each other (Fig. 1), but Generative AI cannot perfectly replicate this in its generative task. The Generative AI might create a video where the subject's Zygomaticus Major is elevated, but the Orbicularis Oculi stays still. This will get flagged, and the framework can conclude that the given data is a deepfake.

D. Multi-Modal Integration and Decision Logic

The final stage integrates the biological pillars (Circulatory System and Anatomical Mechanics). The system follows a "veto" logic—if the video fails in either of the two tests, it will be concluded to be fake. The video has to pass both checks (rPPG check and musculoskeletal synergy audit) to be classified as authentic. This two-stage veto logic ensures that if a generative AI wants to bypass this framework, it has to mimic an authentic pulse and also the complex coordination of the biological human.

IV. RESULTS & OBSERVATIONS

The above table is a summary of the physiological signal quality observed among the randomly selected subjects from the UBFC-Dataset. Most videos had consistent lighting with proper display of the ROIs, with some exceptions where the subjects had one of their ROI covered with hair. Some

Table I
Summary of Physiological Signal Quality in Authentic Videos

Subject	Dataset	FPS	Signal Quality
1	2	29.26	Moderate
3	2	29.55	Poor
4	2	29.61	Moderate
5	2	29.66	Poor
8	2	29.89	Good
9	2	29.54	Poor
10	2	29.88	Good
11	2	29.70	Good
12	2	29.03	Good
13	2	29.79	Moderate
14	2	29.37	Good
15	2	29.59	Moderate
16	2	29.93	Moderate
17	2	29.90	Good
18	2	29.75	Moderate

videos even had minor to significant head movements, which decreased the signal quality.

Table II
Summary Of Physiological Signal Quality In Synthetic Videos

Video ID	Dataset	FPS	Signal Quality
DF-1	FaceForensics ++	24.00	Moderate
DF-2	FaceForensics ++	24.00	Poor
DF-3	FaceForensics ++	30.00	Moderate
DF-4	FaceForensics ++	30.00	Moderate
DF-5	CelebDF	30.00	Moderate
DF-6	CelebDF	30.00	Poor
DF-7	CelebDF	30.00	Moderate
DF-8	CelebDF	30.00	Moderate

The quality of physiological signals in synthetic facial videos is summarized in Table II. Deepfake samples from the FaceForensics++ and CelebDF datasets were chosen to evaluate whether the proposed framework can be generalized to a variety of deepfake generation pipelines. In all the synthetic videos analyzed, the rPPG signal extracted had moderate or poor signal quality. Even though recent deepfake videos can render very high visual quality and provide temporal smoothing, the monitored green-channel signals lacked constrained dynamics and temporal coherence of biologically plausible blood volume pulse properties. These results show that persuasive synthetic videos are not always able to replicate natural physiological indicators.

V. DISCUSSION

The findings have demonstrated the evident difference between genuine and counterfeited videos based on the behavior of physiological signs, regardless of the absence of visual distinction.

The evaluation of original videos demonstrated poor to good signal quality, which is anticipated in real-life situations due to head movement, uneven lighting, and partial face coverage by hair. The multi-ROI design ensured that when one ROI was blocked, the secondary ROI maintained signal extraction. This is possible because of the sufficient existence

of biologically correct blood pulses in the ROIs. Conversely, synthetic video analysis performed on FaceForensics++ and CelebDF demonstrated no instance of good signal quality. Whereas other videos showed consistent green light signals, they could not replicate the anatomical consistency of blood pulses as observed in actual videos.

This study made an important observation regarding visual differentiation. A good portion of the synthetic videos appeared unnatural to the human viewer, while other videos reflected high-tier visuals indistinguishable as deepfakes. Nevertheless, visual distinction was not the sole criterion, since such videos demonstrated the absence of biologically correct rPPG behavior. The physical cues act as a biological watermark.

The deepfake generators currently strive to achieve Nash Equilibrium, in which the discriminator cannot distinguish between authentic and fake videos. However, the models did not consider the Exchange Principle of Locard [5]. The evidence left is the existence of the pulse, which is an aspect of human anatomy. It may be applied as forensic evidence even in the event that the video is near pixel-perfect, but does not contain the biological pulse of a human being.

VI. LIMITATIONS

While the framework provides a robust alternative for pixel detection, many limitations have been observed during the clinical audit:

A. Sensitive to Motion

As observed in subjects 5 and 13 in the real data analysis, any movements like head wobbles will disrupt the ROI alignment, which hinders the ability to obtain better results.

B. Inconsistent Lighting

The rPPG method relies on the absorption of light by hemoglobin. In the case of subject 9, changes in light intensity caused the green channel readings to become less accurate, resulting in the signal quality being classified as "poor." This framework requires

videos with consistent lighting to deliver better results.

C. Physical Covering of ROIs

As observed in subject 15, hair covered the ROIs, which hindered accurate results. The presence of makeup or any other coverage, like masks, can also be responsible for delivering inaccurate results.

B. Introduction of Triple-ROI Triangulation

In future iterations, the Nasal Bridge will be introduced as a new ROI, as it has a high concentration of capillaries and a stable position during facial movements. This will be a part of the "Triple-ROI Triangulation," where the ROIs form a triangle, improving green channel analysis and deepfake classification.

C. Multi-Modal Fusion

Beyond rPPG, the framework can be extended to include eye-tracking (saccadic movement analysis), creating a true multi-modal framework. This will be extremely challenging for current Generative Adversarial Networks to replicate in real-time.

VIII. CONCLUSION

Our research has demonstrated that the implementation of two-phase testing (rPPG and Musculoskeletal synergy) is crucial against current deepfake attacks. Facial landmark tracking and utilizing rPPG to differentiate between real and synthetic videos is essential. Despite challenges and limitations like head movements by subjects during analysis, the multi-ROI design has demonstrated itself to be efficient where traditional tools were futile.

Rather than serving as a binary detector, this framework offers evidence-based findings of physiological authenticity paired with a verifiable, evidence-based audit where synthetic media can efficiently be found.

VII. FUTURE SCOPE

This framework is currently on its foundational shift towards hybrid forensics. To transform this into an

operational countermeasure, the following improvements are proposed:

A. Integration of Thermal and Infrared

Due to the framework's heavy dependence on consistent lighting for the green channel to operate, integration of thermal and infrared imaging is required. Near-Infrared (NIR) imaging will be used as a mitigation for heavy makeup and low-light videos, heavily improving forensic reliability for nighttime or inadequate light scenarios.

REFERENCES

1. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 27, pp. 2672–2680, 2014.
2. H. T. Nguyen, J. Yamagishi, and I. Echizen, "Use of a capsule network to detect fake images and videos," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 2019, pp. 2307–2311.
3. Y. Li, X. Yang, P. Sun, H. Qi, and S. Lyu, "Celeb-DF: A large-scale challenging dataset for deepfake forensics," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2020, pp. 3207–3216.
4. D. Guëra and E. J. Delp, "Deepfake video detection using recurrent neural networks," in *Proc. IEEE Int. Conf. Adv. Video Signal-Based Surveill. (AVSS)*, 2018, pp. 1–6.
5. P. L. Kirk, *Crime Investigation: Physical Evidence and the Police Laboratory*, 2nd ed. New York, NY, USA: Interscience Publishers, 1974.
6. G. de Haan and V. Jeanne, "Robust pulse rate from chrominance-based rPPG," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 10, pp. 2878–2886, Oct. 2013.
7. W. Wang, A. C. den Brinker, S. Stuijk, and G. de Haan, "Algorithmic principles of remote photoplethysmography," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 7, pp. 1479–1491, Jul. 2017.
8. X. Niu, H. Han, S. Shan, and X. Chen, "SynRhythm: Learning a deep heart rate estimator from general to specific," *IEEE Trans. Image Process.*, vol. 29, pp. 3790–3803, 2020.
9. J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiol. Meas.*, vol. 28, no. 3, pp. R1–R39, 2007.
10. Z. Sun, H. Li, and Y. Wang, "Remote photoplethysmography: A review of recent advances," *IEEE Access*, vol. 9, pp. 97965–97985, 2021.
11. S. Bobbia, R. Macwan, Y. Benezeth, A. Mansouri, and J. Dubois, "Un-supervised skin tissue segmentation for remote photoplethysmography," *Pattern Recognit. Lett.*, vol. 100, pp. 107–112, 2017.
12. A. Rossler et al., "FaceForensics++: Learning to detect manipulated facial images," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, 2019.
13. S. Standring, *Gray's Anatomy: The Anatomical Basis of Clinical Practice*, 41st ed. Elsevier, 2016.