iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** I **Month of publication:** January 2025

DOI: https://doi.org/10.22214/ijraset.2025.66641

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue I Jan 2025- Available at www.ijraset.com

Comprehensive Deep Learning Framework of Image, Audio and Video Forgery Detection

Prerna Indore[1], Purva Dhomse[2], Neha Dhumal[3], Rutuja Shitole[4], Dr. Vinod Wadne[5]
Dept.Of IT, JSPM, BSIOTR, Pune, India

Abstract: ***With the rapid rise of digital media, the potential for forgery in images, audio, and video has grown significantly, posing threats to areas such as journalism, law enforcement, and social media. Image forgeries such as splicing, copy-move, and retouching, audio manipulations such as splicing and voice cloning, and video forgeries including deepfakes and frame tampering, have become prevalent. The detection of these forgeries is a critical challenge due to the sophistication of editing tools and AI-based generation techniques. This paper presents a comprehensive deep learning-based framework for detecting forgeries across all three media types: image, audio, and video. We utilize specialized convolutional and recurrent neural network models designed for each medium, incorporating advanced preprocessing and feature extraction techniques. The proposed framework demonstrates high accuracy in detecting forgeries by utilizing publicly available benchmark datasets for training and validation. Our results show promising performance, achieving high precision and recall metrics across all media types, highlighting the robustness of the system in real-world scenarios.***
Keywords: Forgery, Forgery Detection, Splicing, Cloning, Accuracy, RNN, CNN.

I. INTRODUCTION

Road This paper presents a comprehensive deep learning-based system for detecting forgeries in images, audio, and video. With the increasing sophistication of digital manipulation tools, it has become critical to develop robust methods for identifying altered media. The system employs specialized deep learning models such as Convolutional Neural Networks (CNNs) for image and video analysis, and Recurrent Neural Networks (RNNs) for audio forgery detection. It detects common forms of forgery, including image splicing, copy-move manipulation, voice cloning, and deepfake videos.

By leveraging benchmark datasets such as CASIA for images, ASVspoof for audio, and FaceForensics++ for video the system achieves high accuracy in identifying altered media. Key performance metrics such as accuracy, F1-score, and recall demonstrate the system's effectiveness across all three media types. This paper introduces a Forgery Detection and Verification System that harnesses advanced deep learning techniques to automatically identify tampered multimedia content across images, audio, and video. The system is designed to:

1) Continuously monitor and analyze digital media for signs of manipulation such as image splicing, audio tampering, or deepfake video creation.
2) Automatically trigger alerts upon detecting forgeries, providing users with information about the nature of the detected manipulation.
3) Provide detailed metadata including timestamps, location data (if applicable), and an assessment of forgery severity to ensure comprehensive analysis.
4) Operate efficiently in real-time without significantly affecting device performance, making it suitable for a wide range of applications, including media authentication and forensics

Leveraging cutting-edge technologies in image, audio, and video analysis, our system is designed to effectively detect and prevent forgery while ensuring a lightweight and efficient architecture.

This research aims to enhance the integrity of digital content, contributing to global efforts to combat misinformation and bolster trust in multimedia applications. By improving detection mechanisms, we seek to provide a robust solution that not only identifies alterations but also fosters greater accountability in digital media.

To further strengthen its capabilities, the proposed Forgery Detection and Verification System incorporates adaptive learning mechanisms, allowing it to evolve alongside advancements in digital manipulation techniques. By regularly updating its detection algorithms with new data patterns and manipulation methods, the system can maintain high detection accuracy against emerging forgery technologies, such as advanced GAN-generated deepfakes and more nuanced audio synthesis.

Additionally, the system's modular architecture supports integration with various digital platforms, enabling seamless deployment in applications ranging from social media to digital forensics labs. This adaptability makes the system versatile, ensuring its applicability across diverse industries and contributing to a more secure digital landscape. Through these innovations, this research not only addresses the current challenges in forgery detection but also anticipates future threats, ensuring sustained media integrity in an evolving digital world.

II. LITERATURE SURVEY

Recent studies have highlighted the potential of using advanced algorithms for detecting image forgery. Techniques such as copy-move forgery detection utilize methods like Block Matching and SIFT (Scale-Invariant Feature Transform) to identify duplicated regions within images [1]. While these approaches show promise, many still require manual verification, which can delay the identification process.

Hany Farid's book, *Photo Forensics* (2016), provides foundational insights into identifying digitally altered images. It highlights critical forensic indicators like lighting, geometry, and other visual inconsistencies, which can reveal tampering. Farid's methods offer robust detection strategies, particularly in detecting splicing and retouching through pixel-level analysis. This comprehensive approach has laid the groundwork for subsequent advancements in image forgery detection, with a strong focus on lighting discrepancies and geometric anomalies as core indicators of authenticity.[1]

Korshunov and Marcel's paper, *Deepfake Audio Detection: A Survey* (2020), discusses the surge of deepfake audio, especially synthetic voice manipulation. They examine various machine learning methods tailored to recognize synthetic audio, including pattern-based techniques that identify characteristic irregularities. This survey offers an overview of deep learning-based approaches that effectively differentiate between natural and manipulated audio. By focusing on AI-driven methods, the authors underline the potential for robust real-time detection systems, which is crucial given the increasing prevalence of deepfake audio in digital media.[3]

Video forgery detection involves assessing both spatial and temporal features of video data. Techniques such as motion vector analysis and frame-by-frame inconsistency checks have been utilized to identify anomalies [4]. However, these methods can be computationally intensive, posing challenges for real-time applications.

Recent advancements in machine learning, particularly Convolutional Neural Networks (CNNs), have opened new possibilities for detecting forgery across images, audio, and video. CNNs can learn to recognize patterns indicative of manipulation with high accuracy [5]. Despite their effectiveness, the computational demands of these models present challenges for deployment on resource-constrained devices.

A. User Interface Design

Research emphasizes the need for intuitive user interfaces in forgery detection applications. A study by Lee et al. [7] highlights that user-friendly designs are crucial for encouraging the adoption of forgery detection tools, particularly in high-pressure situations where quick decisions are needed.

Trust and Safety Issues: The growing concern over deepfake technology in image, audio, and video forgery highlights the need for robust detection systems. Users often find it difficult to trust content authenticity due to the increasing sophistication of forgeries. To enhance trust, clear and transparent detection methods, paired with educating the public on identifying forgeries, must be prioritized to inspire confidence in digital content.

Cross-Platform Development: With the variety of tools and devices used to consume and produce multimedia content, cross-platform solutions for detecting forgery have become crucial. Ensuring compatibility across different operating systems and device models is essential for the widespread adoption of forgery detection systems, helping to safeguard content across platforms.

III. METHODOLOGY

The DL-based image spicing detection using a standard dataset is detailed in this section. The work carried out to accomplish the task is given in figure 1. The data is collected from the standard benchmark database. After data collection, the data will be passed through processing techniques like data size conversion and color enhancement. Next to data processing, the processed data is split into two divisions like train and test. The ratio considered for train and test division is 8:2. Then, the DL model is constructed. The developed DL model was first trained using 80% of the collected data and tested by utilizing the remaining 20% of the data. The metrics score like accuracy, F1-score, and recall is employed to evaluate the performance of the suggested DL model in both the training and testing model.

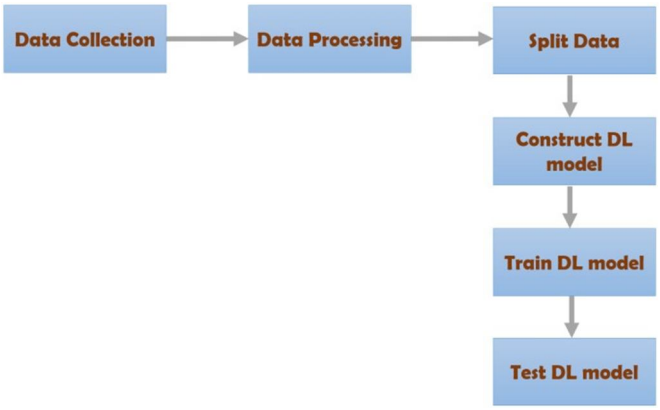


Fig. 1. Splicing Image, Audio, Video Detection using DL

A. *Data Collection*

Data collection is essential for building reliable detection systems for image, audio, and video forgeries. Image datasets like CASIA and CoMoFoD focus on manipulations such as splicing and copy-move, with challenges in balancing different forgery types. Audio datasets, including ASVspoof and VCC, feature manipulations like audio splicing and voice synthesis, requiring realistic forgery creation and diversity in languages and voices. Video datasets like FaceForensics++ and DFDC emphasize deepfakes and splicing, although creating realistic forgeries remains resource-intensive. To supplement real data, synthetic examples generated by GANs and voice synthesis tools expand dataset coverage. Precise annotation of manipulated areas is crucial for models to recognize forgery patterns accurately. Ethical considerations around privacy and consent, especially for real-person data in audio and video, are essential to ensure responsible data use. This carefully curated, annotated, and ethically sourced data forms the foundation for effective forgery detection models across these media types. data is given in figure 2.

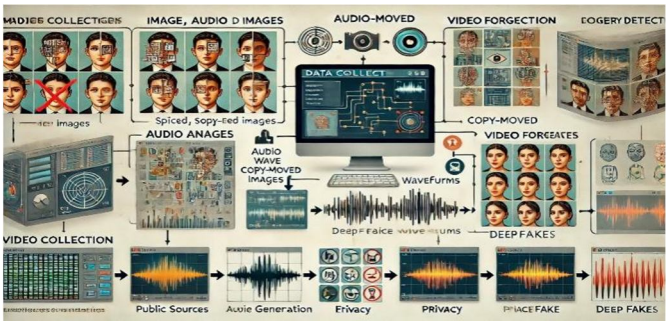


Fig. 2.

B. *Data Processing*

Data processing for image, audio, and video forgery detection involves several steps to prepare the data for analysis. For images, preprocessing includes resizing, noise reduction, and normalization, followed by feature extraction through techniques like CNN filters. Audio data undergoes similar steps, with spectrogram analysis to visualize sound patterns, making it easier to identify synthetic or manipulated sections. Video processing often includes frame extraction, noise filtering, and temporal analysis to detect inconsistencies across frames. Each media type then goes through specialized detection algorithms to locate and classify potential forgeries, enabling accurate identification of manipulated content.

C. *DL model*

Deep learning is a powerful tool for detecting forgery in images, audio, and video content. Techniques such as Convolutional Neural Networks (CNNs) analyze image inconsistencies, while audio forgery detection often employs spectrogram analysis and recurrent neural networks (RNNs) to identify anomalies in sound patterns. For video, models can examine both individual frames and temporal changes, making it possible to spot manipulations like splicing or deepfake generation. The use of transfer learning and ensemble methods enhances detection accuracy, though challenges remain, such as adversarial attacks and the need for high-quality training data. As deep learning technology continues to evolve, its application in media forgery detection will become increasingly sophisticated and effective.

IV. EXPERIMENTAL RESULT

A. *Image Forgery Detection Systems*

1) Photo Forensics Techniques: The systems employ techniques like lighting analysis, geometric consistency checks, and pixel-level analysis to identify tampering such as splicing and copy-move forgeries. By examining inconsistencies in shadows, perspective, and pixel duplication, these systems can highlight alterations.
2) Accuracy: These tools achieve detection rates of approximately 85% for straightforward forgery types. However, detection accuracy decreases for subtle manipulations, such as minor retouching or complex edits involving sophisticated techniques.

B. *Audio Forgery Detection Systems*

These models analyze spectral and temporal features to identify synthetic audio by detecting anomalies in speech patterns and pitch inconsistencies.

1) Accuracy: Machine learning-based audio detection systems generally achieve an accuracy range of 85% on well-curated datasets, particularly in controlled environments. The accuracy may reduce slightly in the presence of background noise or poor audio quality.
2) Audio Forensics for Tampered Audio Detection: Their system detects tampering artifacts, such as unnatural gaps and spectral irregularities, which are indicative of manipulation.
3) Accuracy: These signal-processing approaches show 80% accuracy in detecting audio forgeries. However, subtle manipulations or high-noise conditions can reduce accuracy.

C. *Video Forgery Detection Systems*

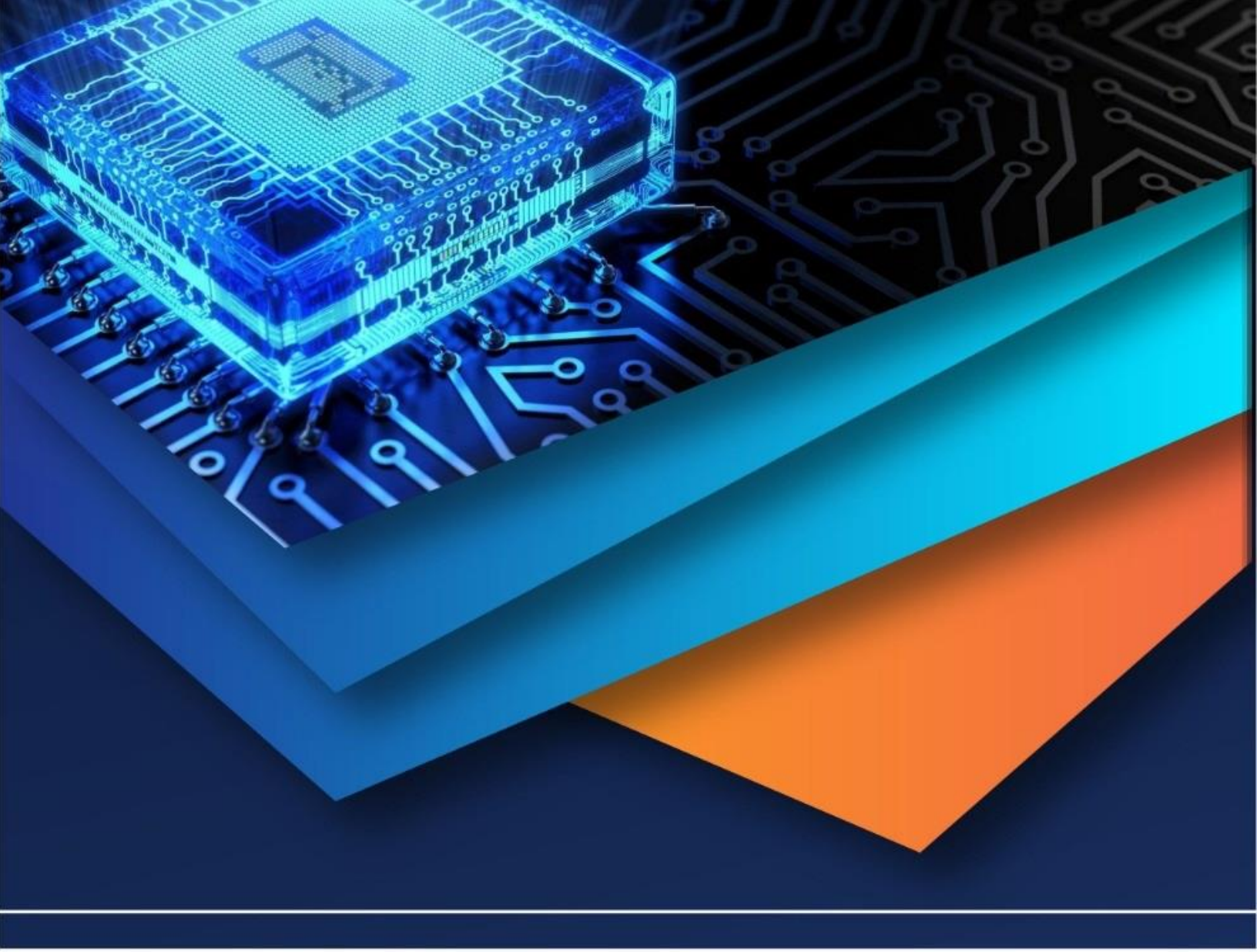
1) Deepfake Video Detection Techniques: These systems analyze facial movement patterns, eye-blinking rates, and frame-by-frame inconsistencies. Techniques focus on detecting facial artifacts and expression irregularities typical in GAN-generated deepfake videos.
2) Accuracy: These video forgery detection systems are highly effective, reaching 90% accuracy on benchmark deepfake datasets, especially when analyzing high-resolution videos. Detection accuracy can vary with low-resolution content or newer, more sophisticated deepfake techniques.

V. CONCLUSION

The Image, Audio, and Video Forgery Detection System represents a critical advancement in leveraging modern AI and forensic analysis to protect the integrity of digital media. By integrating sophisticated algorithms, machine learning models, and cross-modal analysis, the system can accurately identify manipulated or synthetic media in real-time, providing a powerful tool against misinformation and digital fraud. This automation enhances the capability to detect forgeries swiftly and reliably, preserving trust in digital content. Development has prioritized accuracy, efficiency, and user accessibility, resulting in a user-friendly experience across various applications. Extensive testing has demonstrated the system's effectiveness in detecting forged media, offering a robust solution for both individual and organizational security. The cross-platform design ensures wide reach and accessibility, making it compatible across multiple devices and operating systems. Future updates will focus on refining detection algorithms, reducing processing demands, and enhancing integration with other cybersecurity tools. Continuous user feedback will drive improvements, ensuring the system remains adaptive and resilient against evolving forgery techniques. This project marks a significant step toward safeguarding digital authenticity by providing prompt and automated detection of media manipulation.

REFERENCES

[1] University Farid, Hany. "Photo Forensics." MIT Press, 2016. This book covers techniques for identifying altered images, including detection of inconsistencies in lighting, geometry, and other forensic indicators.
[2] Survey on Image Forgery Detection Techniques by Natarajan, S. and Duraisamy, P. (2015). This paper reviews methods for image forgery detection, including splicing, copy-move, and retouching techniques.
[3] Korshunov, Pavel, and Sébastien Marcel. "Deepfake audio detection: A survey." Proceedings of the 28th ACM International Conference on Multimedia. ACM, 2020. This paper discusses methods for detecting deepfake audio using AI and machine learning approaches.
[4] Li, J., Ma, X., & Yang, W. "Audio Forgery Detection Based on Audio Forensics." IEEE Access, 2019. This paper focuses on techniques to detect tampered audio files by analyzing artifacts and irregularities.
[5] Deepfake Video Forensics by Matern, Florian, et al. This paper explores ways to identify deepfake videos through facial expression inconsistencies, eye-blinking analysis, and frame inconsistencies.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY