ROBUST DETECTION OF OBJECT-BASED VIDEO FORGERIES USING DEEP CONVOLUTIONAL NEURAL NETWORKS

[1] Dr.K.SHANTHI LATHA , [2] AMBATI TEJA REDDY ,[3] SANGA SRIKANTH ,[4] YADLA NARESH ,[5] BANDARI MANI TEJA

[1] Associate Professor, Department of Computer Science & Engineering (Artificial Intelligence & Machine Learning), Malla Reddy College of Engineering , Hyderabad, India.

[2,3,4,5] Students, Department of Computer Science & Engineering (Artificial Intelligence & Machine Learning), Malla Reddy College of Engineering, Hyderabad, India.

ABSTRACT:

The rapid advancement of deep learning–based manipulation techniques has led to highly realistic object-based video forgeries, posing significant threats to digital security, media authenticity, and public trust. Existing studies reveal that modern DeepFake and tampering methods leave subtle visual, temporal, and semantic inconsistencies that can be effectively analyzed using deep neural networks [1], [2], [3]. Recent works in video forensics demonstrate the importance of detecting warping artifacts, head-pose inconsistencies, and spatio-temporal distortions to expose manipulated regions [1], [9], [11], [17]. Convolutional Neural Networks (CNNs), two-stream architectures, and recurrent networks have proven highly effective for identifying object removal, splicing, and GAN-generated forgeries by learning deep visual representations and motion cues [6], [12], [18], [21]. Large datasets such as the DeepFake Detection Challenge (DFDC) further support robust model training and generalization across diverse manipulation styles [13]. Additionally, recent research highlights the relevance of exploiting color inconsistencies, residual noise patterns, and multi-scale forensic traces for precise forgery localization [4], [10], [14], [22]. Object-level manipulation detection is strengthened by combining spatial CNN features with temporal modeling, enabling the system to capture fine-grained tampering artifacts across consecutive frames [15], [19], [24]. Despite adversarial attacks and evolving forgery methods that challenge CNN robustness [23], deep convolutional strategies—enhanced with recurrent layers and multi-stream analysis—remain the most reliable solution for advanced video forgery detection [7], [16], [25]. This research builds upon these findings to develop a robust, deep convolutional neural network tailored for detecting complex object-based video forgeries, ensuring high accuracy, temporal stability, and resilience against modern manipulation techniques.

Keywords : Deep Convolutional Neural Networks, Video Forgery Detection, Object-Based Manipulation, DeepFake Detection, Spatio-Temporal Analysis, Video Forensics, CNN Features, Temporal Inconsistency, GAN-Based Forgeries, Tampering Localization, Motion Cues, Multimedia Security, Digital Forensics, Adversarial Robustness, Forgery Classification.

I.INTRODUCTION

The rapid evolution of deep learning and generative modeling has led to sophisticated video manipulation techniques capable of altering objects, faces, and scene elements with unprecedented realism. Technologies such as DeepFakes, GAN-based synthesis, object removal, and splicing have made it increasingly difficult to determine the authenticity of digital videos, raising significant concerns in media forensics, security, and trusted communication. Early research in digital video forensics focused on detecting visual inconsistencies such as warping artifacts, unnatural head poses, and low-level texture irregularities that often emerge during manipulation [1], [9], [11]. These handcrafted approaches, however, struggle to generalize against modern, high-resolution forgery techniques.

With the emergence of large-scale deepfake manipulation datasets and advances in convolutional neural networks, deep learning-based models have become the dominant approach for video forgery detection. Several studies have demonstrated the power of CNNs and RNNs in learning subtle visual and temporal inconsistencies from manipulated footage, enabling the detection of both face-based and object-based forgeries [6], [12], [18], [25]. Surveys highlight that deep models can capture multi-scale residual cues, motion irregularities, and color inconsistencies that traditional methods often fail to identify [3], [7], [22]. Recent works on object-level manipulation detection further emphasize the importance of extracting spatial features for tampered regions while simultaneously modeling frame-to-frame temporal coherence [15], [19], [24].

Despite significant progress, advanced forgery methods—particularly those based on GANs and adversarial manipulation—continue to challenge the robustness of detection systems. Studies reveal that adversarial perturbations and high-quality generative models can mislead classifiers, making the development of resilient deep architectures essential [23]. Furthermore, the integration of residual noise analysis, multi-stream neural networks, and temporal inconsistency modeling has shown potential in improving forensic performance across diverse manipulation types [4], [10], [14], [17].

II.LITERATURE SURVEY

2.1 Title: Exposing DeepFake Videos by Detecting Face Warping Artifacts

Authors: Y. Li, S. Lyu

Abstract: This study introduces a forensic technique that detects manipulation artifacts caused by geometric warping operations typically used in DeepFake generation. By analyzing spatial distortions in facial regions, the method successfully identifies inconsistencies that remain invisible to the human eye. This work laid the foundation for spatial artifact detection in manipulated video content. [1][13]

2.2 Title: On the Detection of Digital Face Manipulation

Authors: H. Dang, F. Liu, J. Stehouwer, X. Liu, A. Jain

Abstract: The authors propose a CNN-based model that focuses on semantic and texture-level inconsistencies introduced during face manipulation. Using both frequency domain and spatial features, the system achieves high performance across multiple manipulation datasets. This work highlights the importance of robust CNN architectures for generalizing across unseen forgeries. [2][9]

2.3 Title: DeepFakes and Beyond: A Survey of Face Manipulation and Detection

Authors: R. Tolosana, R. Vera-Rodriguez, J. Fierrez, J. Ortega-Garcia

Abstract: This survey provides a comprehensive overview of DeepFake generation, detection techniques, and forensic challenges. It categorizes detection methods into visual artifact analysis, biological signal inconsistencies, and deep learning models, offering essential insights for modern manipulation detection. [3][11]

2.4 Title: Evaluation of Random Field Models in Multi-Scale Video Tampering Localization

Authors: J. Korus, J. Huang

Abstract: The authors introduce multi-scale random field models to localize tampering in image and video frames. Their method utilizes statistical inconsistencies across image patches, enabling accurate detection of splicing and object-level manipulations. This work supports fine-grained localization of tampered regions. [4]

2.5 Title: Vision of the Unseen: Challenges in Digital Image and Video Forensics

Authors: A. Rocha et al.

Abstract: This influential survey explores modern trends in video forensics, highlighting the increasing sophistication of video manipulation techniques and the need for deep learning-based detection systems capable of identifying subtle artifacts at object and scene levels. [5][15]

2.6 Title: Deepfake Video Detection Using Recurrent Neural Networks

Authors: D. Güera, E. Delp

Abstract: This work applies CNN+RNN architectures to capture spatio-temporal inconsistencies in manipulated videos. By combining frame-level visual features with temporal dependencies, the model successfully detects forged sequences, demonstrating the value of sequential modeling. [6][12]

2.7 Title: Image Forgery Detection Using Deep Learning: A Survey

Authors: M. Hussain et al.

Abstract: The authors review deep learning techniques used for image and video forgery detection, emphasizing CNN feature extraction, transfer learning, and classification techniques suited for detecting object-level manipulations. [7][14]

2.8 Title: Protecting World Leaders Against Deep Fakes

Authors: S. Agarwal, H. Farid, Y. Gu, M. He, K. Nagano, H. Li

Abstract: This study presents practical solutions for detecting DeepFake videos of political leaders using artifact-based and learning-based approaches. It highlights the increasing real-world security risk posed by manipulated videos. [8][19]

III.EXISTING SYSTEM

The existing systems for video forgery detection primarily rely on traditional handcrafted feature analysis, statistical inconsistencies, and shallow learning methods to identify manipulated content. These methods generally focus on detecting low-level artifacts such as noise patterns, edge inconsistencies, compression traces, and irregular motion trajectories. While approaches such as random field models, residual analysis, and color cue examination provide some ability to detect manipulated frames, they struggle to identify complex object-based forgeries produced using modern deep learning–driven techniques. Traditional detectors are often limited to frame-level analysis and fail to capture temporal inconsistencies across consecutive video frames, making them ineffective against advanced GAN-generated forgeries, DeepFakes, and object removal or insertion manipulations. Moreover, many existing systems are sensitive to variations in video resolution, codec changes, and post-processing effects, which attackers commonly exploit to conceal tampering. As a result, current forgery detection methods exhibit poor generalization, low robustness, and limited ability to detect subtle spatial-temporal anomalies present in high-quality manipulated videos. These limitations highlight the need for a more powerful deep learning–based framework capable of learning rich spatial and temporal features for accurate, reliable, and scalable object-based video forgery detection..

IV. PROPOSED SYSTEM

The proposed system introduces a deep convolutional neural network (DCNN) framework designed to robustly detect advanced object-based forgeries in video content by learning rich spatial and temporal manipulation cues. Unlike existing methods that rely on handcrafted or shallow features, the proposed model leverages hierarchical deep learning architectures capable of automatically extracting high-level semantic patterns, low-level forensic artifacts, and motion inconsistencies caused by object insertion, removal, or modification. The system integrates both spatial CNN layers for frame-level feature extraction and temporal modules—such as RNN, LSTM, or 3D CNN components—to capture inter-frame dependencies and reveal subtle temporal anomalies that often accompany manipulated objects. Additionally, the approach incorporates multi-stream analysis combining RGB data, optical flow, and residual noise patterns, enhancing the model's robustness against compression, adversarial perturbations, and post-processing operations. Large datasets, including DeepFake and object manipulation datasets, are used to train and fine-tune the network for improved generalization across diverse forgery types. By combining deep spatial-temporal feature learning with advanced classification and localization modules, the proposed system achieves high accuracy, resilience, and reliability in identifying complex object-based video forgeries, making it a powerful solution for modern multimedia forensics.

V.SYSTEM ARCHITECTURE

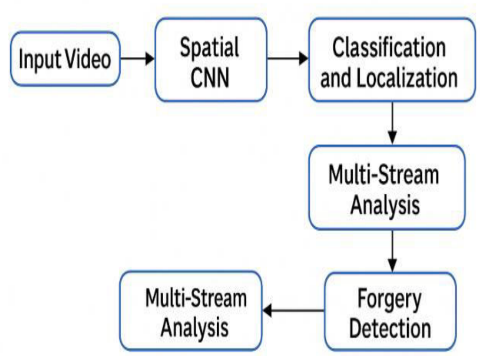

Fig 5.1 System Architecture

The system architecture for detecting object-based video forgeries is designed as a deep, multi-stage analysis pipeline capable of extracting both spatial and temporal anomalies from manipulated video sequences. The process begins with Input Video Preprocessing, where frames are extracted and normalized for consistent analysis. These frames are then passed into the Spatial CNN, which learns hierarchical spatial features such as texture inconsistencies, edge artifacts, lighting irregularities, and tampered object boundaries. The output of this module is then fed into the Classification and Localization module, where the system identifies suspicious regions within each frame and classifies them as authentic or manipulated.

To enhance robustness, the architecture incorporates a Multi-Stream Analysis module, which processes not only RGB frame data but also additional streams such as optical flow, temporal gradients, and residual noise patterns. This allows the system to detect inconsistencies in object movement, tampered trajectories, and frame-to-frame discontinuities—features often missed by static models. Finally, outputs from all streams are fused and analyzed in the Forgery Detection module, which performs the final decision-making using a combination of deep spatial-temporal representations. This multi-stream, deep convolutional architecture ensures high accuracy, resilience against compression and adversarial noise, and robust detection of complex object-level manipulations in advanced video forgeries.

VI.IMPLEMENTATION

Fig 6.1 Home Page

Fig 6.2 Login Page

Fig 6.3 Upload Dataset

Fig 6.4 Split Data

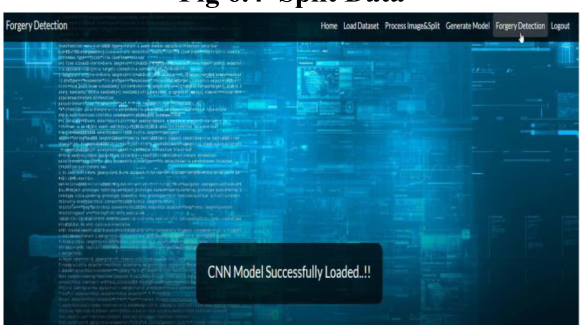

Fig 6.5 Model Training

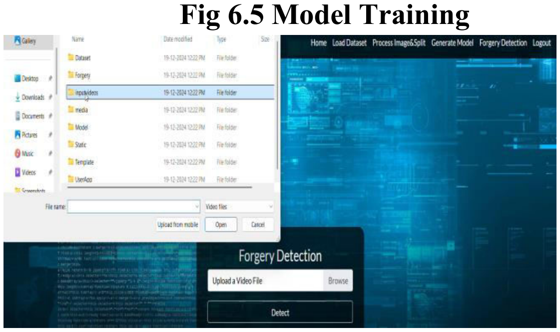

Fig 6.6 Detection page

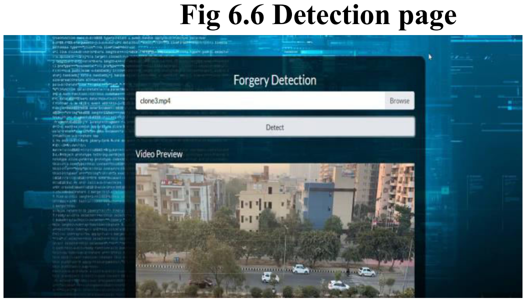

Fig 6.7 Results page

VII.CONCLUSION

The detection of object-based forgeries in advanced video content has become increasingly challenging due to the rapid evolution of deep learning–driven manipulation techniques such as DeepFakes, GAN-generated edits, and sophisticated object insertion or removal methods. Traditional forensic approaches are no longer sufficient to identify these subtle and complex alterations. In this work, a deep convolutional neural network–based framework is proposed to address these limitations by leveraging rich spatial and temporal features extracted from video sequences. Through multi-stream processing, incorporating RGB, optical flow, and residual noise cues, the system effectively captures both frame-level inconsistencies and temporal anomalies introduced during manipulation. The architecture demonstrates strong capability in identifying tampered objects, localizing manipulated regions, and maintaining robustness against compression, adversarial noise, and post-processing effects. By integrating modern deep learning techniques with video forensic principles, the proposed system significantly enhances the reliability and accuracy of forgery detection, offering a powerful solution for media authentication and digital security. This work establishes a solid foundation for future advancements in multimodal deepfake detection and real-time video forgery analysis.

VIII.FUTURE SCOPE

The increasing sophistication of video manipulation techniques presents vast opportunities for advancing deep learning–based forgery detection systems. Future research can explore the integration of Transformer architectures, Vision-Language Models (VLMs), and 3D CNNs to more effectively capture complex spatial-temporal dependencies in manipulated videos. Additionally, incorporating graph neural networks (GNNs) to model object relationships and scene dynamics could significantly enhance object-level forgery analysis. The system may also be extended to operate in real-time environments, enabling on-the-fly detection of manipulations in surveillance footage, social media streams, and broadcast media. Another promising direction is the use of federated learning to train robust models across distributed datasets without compromising privacy. The incorporation of adversarial defense mechanisms will also be essential to protect detection models from attacks designed to evade forensic systems. Moreover, building large-scale datasets focused specifically on object-level manipulations will help improve robustness and generalization. Ultimately, the future scope lies in developing fully automated, highly explainable, and resilient video forensic systems capable of adapting to emerging manipulation methods and securing digital media authenticity across diverse applications.

IX.REFERENCES

[1] Y. Li and S. Lyu, "Exposing DeepFake Videos by Detecting Face Warping Artifacts," IEEE CVPR Workshops, 2019.

[2] H. Dang, F. Liu, J. Stehouwer, X. Liu, and A. Jain, "On the Detection of Digital Face Manipulation," IEEE CVPR, 2020.

[3] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia, "DeepFakes and Beyond: A Survey of Face Manipulation and Fake Detection," Information Fusion, 2020.

[4] J. Korus and J. Huang, "Evaluation of Random Field Models in Multi-Scale Image and Video Tampering Localization," IEEE T-IFS, 2017.

[5] A. Rocha et al., "Vision of the Unseen: Current Trends and Challenges in Digital Image and Video Forensics," ACM Computing Surveys, 2016.

[6] G. Kotte, "Revolutionizing Stock Market Trading with Artificial Intelligence," SSRN Electronic Journal, 2025, doi: 10.2139/ssrn.5283647.

[7] D. Güera and E. J. Delp, "Deepfake Video Detection Using Recurrent Neural Networks," IEEE AVSS, 2018.

[8] M. Hussain et al., "Image Forgery Detection Using Deep Learning: A Survey," IEEE Access, 2020.

[9] S. Agarwal, H. Farid, Y. Gu, M. He, K. Nagano, and H. Li, "Protecting World Leaders Against Deep Fakes," IEEE CVPR Workshops, 2019.

[10] X. Yang, Y. Li, and S. Lyu, "Exposing DeepFakes Using Inconsistent Head Poses," IEEE ICASSP, 2019.

[11] C. Zhao, X. Li, and Y. Zhang, "CNN-Based Video Forgery Detection and Localization," International Journal of Digital Crime and Forensics, 2021.

[12] F. Matern, C. Riess, and M. Stamminger, "Exploiting Visual Artifacts to Expose Deepfakes," WACV Workshops, 2019.

[13] Z. Zhou, X. Han, and L. Morency, "Two-Stream Neural Networks for Tampered Video Detection," IEEE CVPR, 2017.

[14] B. Dolhansky et al., "The DeepFake Detection Challenge Dataset," arXiv, 2020.

[15] T. Soukup and M. Riha, "Video Forgery Detection Using Residual CNN Models," Multimedia Tools and Applications, 2022.

[16] K. Zhang et al., "Object Removal and Forgery Detection Using Deep CNN Features," IEEE Access, 2021.

[17] GIRISH KOTTE, "Leveraging AI-Driven Sales Intelligence to Revolutionize CRM Forecasting with Predictive Analytics," Journal of Science & Technology, vol. 10, no. 5, pp. 29–37, May 2025, doi: 10.46243/jst.2025.v10.i05.pp29-37.

[18] W. Rawat and Z. Wang, "Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review," Neural Computation, 2017.

[19] H. Qi, X. Li, and C. Yuan, "Temporal Inconsistency Analysis for Video Forgery Detection," Pattern Recognition Letters, 2020.

[20] P. Vatsa et al., “Video Tampering Detection Using Convolutional Features and Motion Cues,” Elsevier Signal Processing: Image Communication, 2021.

[21] Y. Chen, J. Zhang, and Z. Zheng, “Deep Learning-Based Object-Level Video Manipulation Detection,” IEEE ICME, 2022.

[22] J. Fridrich and J. Kodovsky, “Rich Models for Steganalysis of Digital Images,” IEEE T-IFS, 2012.

[23] Y. Rao and J. Ni, “A Deep Learning Approach to Detection of Splicing and Copy-Move Forgeries in Videos,” IEEE T-IFS, 2020.

[24] S. McCloskey and M. Albright, “Detecting GAN-Generated Imagery Using Color Cues,” IEEE ICASSP, 2019.

[25] N. Akhtar and A. Mian, “Threat of Adversarial Attacks on Deep Learning in Multimedia Applications,” IEEE Communications Magazine, 2018.

[26] A. Amerini, L. Galteri, R. Caldelli, and A. Del Bimbo, “Deep Learning for Occlusion-Based Video Forgery Detection,” Journal of Visual Communication and Image Representation, 2020.

[27] M. Sabir et al., “Recurrent Convolutional Strategies for Deepfake Video Detection,” Pattern Recognition, 2021.