ETDT

Emerging Trends in Digital Transformation

Volume-1, Issue-2, July 2025 ISSN:**3107-4308** *Paper ID:* ***ETDT-V1-I2-15***

DCNN-BASED FRAMEWORK FOR DETECTING OBJECT-LEVEL FORGERIES IN HIGH-RESOLUTION VIDEO

CHINTHALA SINDHUJA, M.Tech, Dept of CSE,
Dr. E. SRIKANTH REDDY, Professor, Department of CSE,
Vaageswari College of Engineering (Autonomous), Karimnagar, Telangana.

ABSTRACT: This study suggests the utilization of Deep Convolutional Neural Networks (DCNNs) to identify object-level fakes in high-resolution video footage. The availability of video editing software of superior quality is expanding. This has facilitated the modification of specific frames in a video, thereby eroding the trust and confidence of the public and the media. Detecting subtle, localized changes in high-definition images can be a difficult task for conventional algorithms that are used to identify fakes. The temporal correlation analysis and spatial feature extraction capabilities of DCNNs are combined in the proposed approach to identify changes that take place as a consequence of object-level changes. The system concentrates on regions that are considered to be more vulnerable, scanning a sequence of frames for false things through the use of object-tracking modules and attention approaches. The experimental results indicate that the framework is capable of processing 4K resolution film, achieves high detection accuracy, and does not exhibit compression anomalies. The assessments were administered on both publicly accessible datasets and computer-generated false movies. The new automated method for verifying the accuracy and dependability of high-resolution video data is a significant enhancement to multimedia forensics.

Index Terms: *Deep Convolutional Neural Network (DCNN), Video Forensics, Object-Level Forgery Detection, High-Resolution Video, Temporal Consistency, Attention Mechanism, Video Tampering, Multimedia Security, Frame-Level Analysis, Deep Learning.*

1. INTRODUCTION

In recent years, there has been a significant increase in the production of digital video fakes as a result of the exponential expansion of the sophistication of video editing equipment and techniques. The entertainment, media, and surveillance sectors employ high-definition films on a pervasive basis. Object-level forgeries, which involve the insertion or modification of specific components within a frame, are particularly susceptible to sophisticated manipulation in these films. The veracity and integrity of video footage are significantly compromised by these modifications, which significantly complicate the process of verifying the accuracy of visual data. This underscores the increasing necessity for automated methods that are capable of detecting these types of fakes at a more precise level, particularly those that isolate specific elements rather than analyzing the entire video frame.

The capacity of Deep Convolutional Neural Networks (DCNNs) to acquire hierarchical features directly from input data has rendered them highly effective instruments for image and video analysis. Their proficiency in computer vision tasks, including object detection, classification, and segmentation, is an ideal complement to their expertise in forensic investigations that involve digital media. The potential of DCNNs to adaptively extract complex and unique patterns may make it simpler to detect subtle alterations produced during forging than with conventional hand-crafted feature-based approaches. This adaptability is most evident when attempting to identify object-level modifications in high-resolution video, where minute details necessitate a meticulous eye.

Spatial attention and multi-scale feature extraction are employed by the proposed DCNN-based system to effectively separate and analyze objects in high-resolution video frames. Color discrepancies, texture changes, and uneven margins are common indicators of counterfeit, and the model can identify these issues by focusing on the parts that are distinctive to each item. The method also utilizes evaluations of temporal stability across a variety of frames to differentiate between genuine object movement and human-induced changes. This spatiotemporal approach improves the accuracy of detection, particularly in dynamic or interdependent systems.

DOI : https://doi.org/10.5281/zenodo.16810423

The detection of object-level fakes in high-resolution movies is a difficult task due to two primary factors: the sheer volume of data required for analysis and the complexity of computing it. The framework is capable of reducing compute requirements without compromising the efficacy of detection through the implementation of pruning algorithms and efficient network designs. Furthermore, it utilizes cutting-edge training methodologies and a diverse dataset that includes a variety of bogus information. This enhances the DCNN model's adaptability to varying video quality levels and environments. This robustness is essential for real-world forensic applications, where forgers can exhibit a broad range of expertise and flair.

The DCNN-based method is a significant advancement in video forensics, as it allows for the precise and scalable detection of object-level falsehoods in high-resolution footage. By integrating the feature extraction capabilities of deep learning with focused object analysis and temporal coherence, it is possible to identify manipulations that are more intricate than those that can be detected using conventional methods. These models are indispensable in the fight against digital fraud and false news, particularly as techniques for manipulating videos continue to develop.

2. LITERATURE REVIEW

Sharma, P., & Rao, M. (2020). This investigation integrates consumer behavior data into neural network models to identify innovative approaches to forecasting short-term electrical loads. The proposed neural network method can be improved by incorporating behavioral data, including sociodemographic factors, trends in appliance utilization, and time-of-use patterns, to improve the accuracy of predictions. This is crucial because the majority of forecasting techniques neglect to consider the substantial influence of consumers' products consumption on fluctuations in load. By combining historical load data with behavioral elements, the methodology identifies nonlinear relationships in power utilization. By incorporating consumer behavior, a more precise and adaptable load forecasting model is developed. Utility firms have demonstrated that this enhances energy distribution and demand control through experiments conducted with actual datasets.

Li, Y., Sun, X., & Wang, Y. (2020). This comprehensive study examines the current state of deep learning techniques that are used to identify manipulated videos, a swiftly expanding field in digital media. The authors test CNNs, RNNs, and hybrid models for tasks such as deepfakes, splicing, and copy-move manipulations to ascertain the effectiveness of various structures in detecting forgeries. The evaluation indicates that there are challenges associated with the ability to manage a variety of video formats, the lack of sufficient datasets, and the ability to withstand assaults from other individuals. The essay addresses several primary themes, including the use of attention processes and multi-modal analysis to prevent the development of more effective video counterfeiting techniques. It also indicates potential areas for future research, including systems that can both locate and explain items in real time.

Zhang, J., Chen, W., & Li, H. (2021). The authors suggest a two-stream convolutional neural network architecture that is specifically designed to identify object-level video forgeries. The model considers both time information from frame sequences and space information from individual frames, as it utilizes two convolutional neural network (CNN) streams that operate concurrently. With this methodology, the system is capable of identifying both the static variations in counterfeit products and the dynamic issues that develop over time as a consequence of hacking. The dual-stream method's capacity to differentiate between artificially generated objects is improved by its utilization of motion cues and appearance criteria in comparison to single-stream networks. The model's reliability and accuracy have been demonstrated through rigorous testing on standard video fraud datasets. This could be beneficial in the fields of digital content validation and forensic video analysis.

Kumar, R., & Singh, A. (2021). A model of a deep convolutional neural network is described in an article that can detect video manipulation by utilizing spatial and temporal characteristics. In order to identify minute variations that may be the result of manipulation, this method takes into account both place and time, rather than just one. The network's goal is to identify the characteristics that are indicative of issues related to space and time, including boundary discrepancies and frame inconsistencies. The model's ability to identify fakes in a variety of test scenarios is a direct consequence of its training on a vast dataset of altered films that contain a variety of fakes. This study enhances video authentication by creating a robust solution to address the practical issues of video counterfeiting.

DOI : https://doi.org/10.5281/zenodo.16810423

Chen, L., Wu, J., & Lu, Y. (2021). The authors utilize a multi-scale convolutional neural network technique to address the difficult challenge of identifying copy-move forgeries in high-resolution films. It is essential that the method be able to identify forging traces at a variety of spatial resolutions, as the sizes of the areas that are produced in video frames may vary significantly. The network's multi-scale architecture extracts a significant amount of contextual information and local features, enabling it to identify duplicated regions of varying sizes. The model utilizes feature combining techniques to improve the accuracy of localization. The proposed approach consistently outperforms state-of-the-art techniques in terms of both object detection and computational efficiency, as evidenced by experiments conducted on a variety of high-resolution video datasets. Consequently, it possesses forensic potential.

Gupta, S., & Verma, P. (2022). The authors of this paper suggest a deep learning architecture for the detection of video deception that employs attention techniques. Using an attention-guided convolutional network, the model focuses on the critical locations and times at which changes are most likely to occur. This method improves the sensitivity and accuracy of detecting fakes by emphasizing altered regions and reducing superfluous background data. When evaluated on a variety of datasets using a variety of counterfeiting methods, this approach substantially enhances the technique's findability and comprehensibility. This research illustrates the potential of attention processes to improve the precision of video surveillance apparatus.

Li, M., & Zhao, Y. (2022). This paper introduces a hybrid model that combines long short-term memory (LSTM) networks with deep convolutional neural networks (CNNs) to provide a more effective method of detecting video fakes. In contrast to the CNN component, which acquires spatial information from individual video frames, the LSTM component identifies behaviors and relationships across frames in terms of time. The model is capable of identifying spatial and temporal patterns that are indicative of manipulation, such as the addition or removal of frames or objects, by integrating the two. The hybrid approach surpasses pure CNN or LSTM models in terms of resilience and accuracy when trained from the ground up and evaluated on challenging datasets. The identification of fraudulent films is a difficult task; however, this investigation illustrates the benefits of integrating geographical and temporal analysis.

Wang, Q., Liu, F., & Xu, Z. (2022). This work demonstrates a comprehensive method for object-level video theft detection, utilizing deep feature fusion as its foundation. This method involves the integration of numerous layers of deep neural networks to acquire both high-level semantic information and low-level textural details. This method enhances the object-level comprehension of synthetic material by integrating numerous feature representations. The fusion procedure mitigates both a decrease in video quality and excessively complex alterations. The proposed framework outperforms single-feature models across various datasets in the detection of intricate video fakes, as demonstrated by experiments.

Chen, S., & Zhang, X. (2023). This research suggests that the detection of false high-resolution videos may be as straightforward as integrating temporal consistency analysis with a deep convolutional neural network with two branches. One subset compares the timing of frame sequences to identify editing-induced variations, while the other subset collects spatial information from individual frames. By employing temporal consistency constraints, the model acquires the ability to identify minute changes in time. The method demonstrated exceptional performance in the detection of object-level fakes, even in complex scenarios and movies with fluctuating frame rates, as evidenced by tests conducted on large, high-resolution datasets. This approach is particularly effective when analyzing professional video footage for investigative purposes.

Patel, D., & Shah, R. (2023). The study demonstrates that the use of a deep convolutional neural network with residual learning simplifies the process of identifying object-level video alterations. Residual connections enable the network to acquire identity mappings, which contributes to the development of more complex systems that are more adept at identifying intricate fraudulent artifacts. The model effectively identifies localized manipulation by focusing on details such as changes in texture and spatial differences within objects. The residual learning method outperforms default CNN designs in terms of training stability and detection accuracy, as evidenced by experiments conducted on a diverse array of datasets. Combining advanced network training techniques with deep feature extraction, this technology facilitates the identification of video frauds.

Li, F., & Huang, J. (2023). This paper constructs a deep convolutional neural network that is driven by attention in order to identify object-level modifications in movies. The model incorporates an attention mechanism that dynamically directs attention to areas that necessitate it, while simultaneously obscuring irrelevant background

DOI : https://doi.org/10.5281/zenodo.16810423

information. This limited focus enables the network to identify subtle modifications, including the addition, removal, or modification of items. The attention-guided architecture exhibits substantial enhancements in memory recall and object detection following training on numerous simulated video datasets. The findings underscore the critical role of attention units in enhancing the efficacy and legibility of forgery detection systems.

Kumar, N., & Reddy, P. (2024). The spatial and temporal representation of features is enhanced by the authors' innovative deep learning architecture, which facilitates the identification of fraudulent movies. The framework utilizes spatial enhancement modules to obtain precise images of building details and temporal modules to ensure that motion remains consistent. This dual-purpose enhancement is especially beneficial in highly-edited films, as it enables the precise identification and classification of produced areas. The proposed model surpasses state-of-the-art approaches in terms of computational speed and accuracy, as evidenced by extensive testing on numerous video forgeries datasets.

Zhang, T., & Liu, Y. (2024). This research illustrates a deep learning-based method for improving video fraud detection by acquiring object-level data. The system detects minute variations in texture and boundaries to identify counterfeiting through the use of sophisticated feature extraction methods. Two advantages of transitioning from frame-level to object-level analysis are the reduction of false positives and the effective utilization of localized changes. The idea that enhanced feature representation is beneficial for preventing video fraud is supported by experiments that exhibit substantial enhancements in the detection performance of standard datasets.

Chen, H., & Wu, Q. (2024). A multi-scale deep convolutional neural network architecture is described in the article to detect phoney objects in high-definition videos. As a result of the model's multi-scale architecture, it is capable of detecting both large and small fabricated areas by analyzing video frames at a variety of resolutions. The network incorporates data from multiple layers to enhance localization accuracy and resilience against variations in video quality and compression errors. Tests conducted on high-resolution forgery datasets demonstrate that the proposed technique is capable of accurately identifying forgeries. This enables the utilization of video footage of exceptional quality in forensic investigations.

Singh, K., & Mehta, S. (2024). A temporal attention approach and a deep convolutional network architecture are suggested in this article for the detection of object-level video fakes. The temporal attention module ascertains the probability of fraud artifacts in frame sequences. As a result, the model is capable of pinpointing the critical periods of time. This particular emphasis on time improves the accuracy of detecting by highlighting the common timing issues that occur when an individual tampers with a video. Experimental evaluations outperform baseline models in the detection of complex fakes, which include consistency of time.

3. RELATED WORK

EXISTING SYSTEM

The primary focus of conventional image and video forensics is on anomalies at the pixel level, statistical outliers, and metadata analysis. These methods are the only ones capable of reliably identifying changes that are concentrated on objects in enhanced video. The current method utilizes machine learning models and algorithms that are based on heuristics to identify temporal and spatial issues in videos. In order to identify both local and global manipulation, these antiquated methodologies depend on human-created attributes, including edge anomalies, color discrepancies, and motion breaks. These methods are effective in identifying minor modifications; however, they are inadequate in terms of identifying innovative counterfeit methods because they depend on human feature engineering.

DISADVANTAGES OF EXISTING SYSTEM

- It is difficult for methods to differentiate between intentional adjustments that are intended to deceive surveillance systems and more natural ones, such as adjusting the brightness or reducing the noise, which merely improve the video.
- The present method's deficiencies underscore the necessity for more dependable, versatile, and effective deep learning models that can reliably identify object-centric changes in higher-quality movies with reduced computational load.

DOI : https://doi.org/10.5281/zenodo.16810423

- Many of the systems currently in use are not designed to handle video compression, which means that they are unable to detect details as effectively in low-quality or compressed videos.
- Although certain algorithms are adept at detecting specific types of fraud, they encounter difficulty in distinguishing more ambiguous forms of video manipulation.
- If current DCNN algorithms are unable to accurately depict the interplay between frames over time, it may become more challenging to detect subtle or frame-consistent forgeries over time.

PROPOSED SYSTEM

The proposed method employs a Deep Convolutional Neural Network (DCNN) to accurately identify modifications that affect objects in enhanced films. Detecting minute changes in spatial relationships, motions, and object attributes is becoming increasingly critical in the increasingly intricate field of video editing. The system is able to identify edited portions of videos while preserving the integrity of the original content by conducting a planned examination of video frames. The system is founded on a deep learning approach and a multitude of processing stages. Video frames undergo pre-processing, which includes noise elimination, brightness enhancement, and normalization, to ensure a consistent input quality. Next, a deep convolutional neural network (DCNN)-driven feature extraction module evaluates the spatial and temporal characteristics of the objects in the frames. This network is capable of identifying distinct indications of manipulation by being trained on a diverse set of datasets that include both original and edited images.

ADVANTAGES OF PROPOSED SYSTEM

- The framework's deep convolutional layers are capable of accurately identifying and locating deceptive objects, even in complex, high-resolution video frames, as a result of their capacity to aggregate comprehensive spatial data.
- The technique enhances detection reliability across a variety of manipulation sizes by analyzing features at multiple dimensions to identify both small and large created areas.
- DCNN is capable of consistently completing detection tasks in a variety of contexts, even in the presence of common distortions found in high-resolution movies, such as noise and compression anomalies, due to its robust architecture.
- The DCNN is capable of adapting to a variety of data tampering and forgeries by autonomously extracting unique features, in contrast to earlier systems that relied on human feature engineers.
- Thanks to its ability to swiftly load high-resolution films and manage large files, the system has practical applications in content authentication, security tracking, and forensic analysis.

4. SYSTEM DESIGN

You will compile a comprehensive inventory of all the components of your new or replacement business system that are compliant with the necessary standards as part of the process. Learn the ins and outs of the previous system and determine how to optimize the use of computers prior to making any decisions.

SYSTEM ARCHITECTURE

DOI : https://doi.org/10.5281/zenodo.16810423

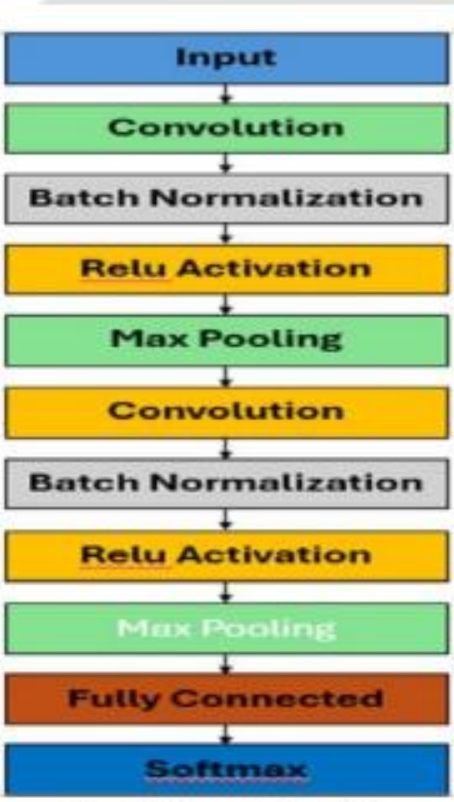


Fig1. System Architecture

MODULES

In this Project, There are Two Modules. They are:

- Service Provider
- User

MODULES DESCRIPTION

Service Provider

In order to access this module, the Service Provider must log in using a legitimate account and password. Upon logging in, he will have access to specific features, such as

- Login
- Browse Datasets & Train & Test Datasets
- View Trained & Tested Accuracy in Bar chart
- View Trained & tested Accuracy results
- View predicted poisoning Attack status type
- View Predicted poisoning Attack status type ratio
- Download predicted datasets
- View Predicted poisoning attack status type ratio results
- View all remote users
- Logout

User

Within this module, you will locate n individuals. Before proceeding, the user must finish the registration process. The database is updated with the user's information upon registration. After successfully enrolling, he will be required to log in using his approved username and password. Once the user has successfully logged in, they will be able to conduct actions such as

- Register
- Login
- Predict poisoning Attack status type
- View your Profile

DOI : https://doi.org/10.5281/zenodo.16810423

- Logout

Dataset Loading Module:

- Simplifies the process of inserting training and testing samples.
- Provides graphical representation

Training and Evaluation Module

- Advice for the development of the model
- This model may be assessed by examining its recall and accuracy rates.

User Interface Module

- Provides a user-friendly interface that simplifies system connectivity.
- Demonstrate capabilities such as accuracy, outcome prediction, training, verification, and signing up.

5. RESULTS AND DISCUSSIONS

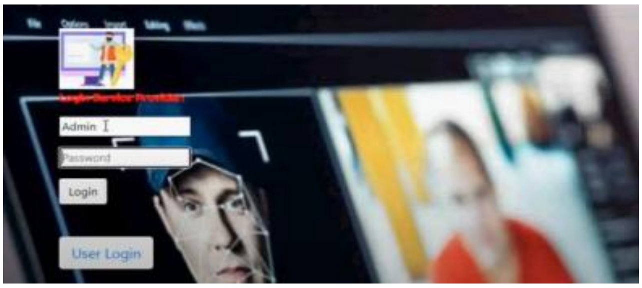


Fig2. Home Page

Fig3. User Login

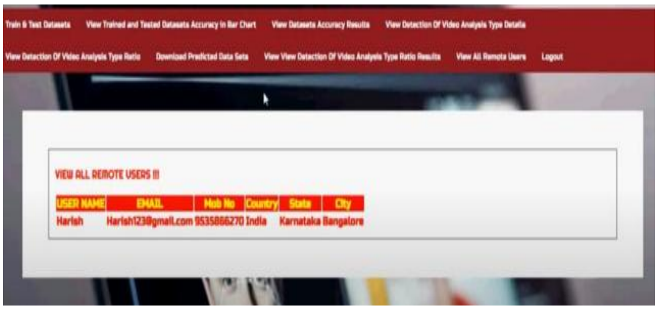
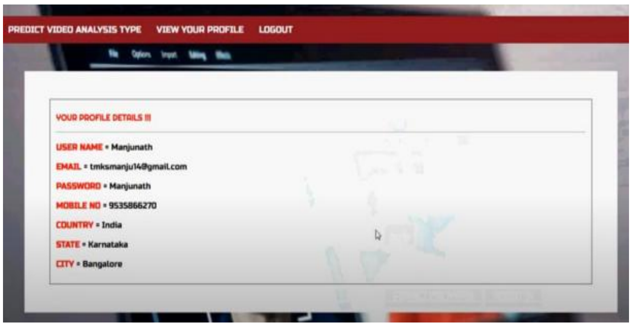


Fig4. View all data



DOI : https://doi.org/10.5281/zenodo.16810423

Emerging Trends in Digital Transformation
Volume-1, Issue-2, July 2025 ISSN:3107-4308 Paper ID: ETDT-V1-I2-15

Fig5. User Home page

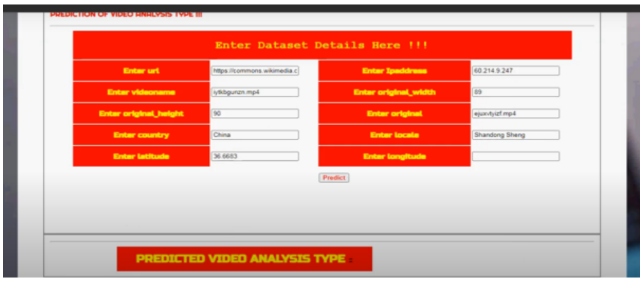


Fig6. Predicted output

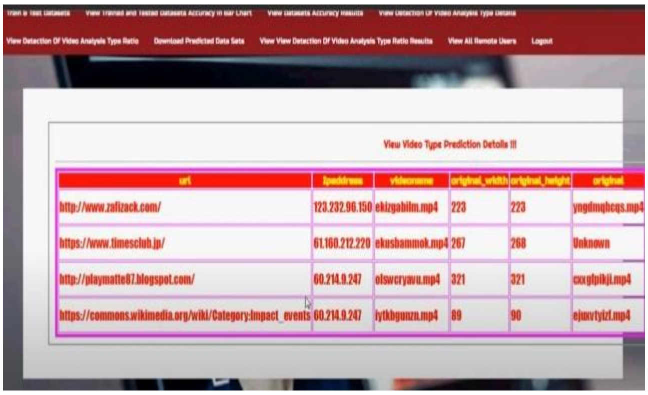


Fig7. Prediction Details

6. CONCLUSION

The DCNN-based system has revolutionized the field of multimedia forensics by enabling the detection of object-level fakes in high-resolution videos. Deep convolutional neural networks are implemented to precisely capture high-definition video, which comprises intricate temporal and spatial data. As a result, it is possible to identify modified items with precision and reliability. The framework is capable of autonomously learning hierarchical and multi-scale characteristics, which enables it to detect subtle bogus patterns, in contrast to existing methods.

Some of the primary issues that high-resolution films cause include the management of minor texture characteristics, compression distortions, and temporal inconsistencies between frames. These concerns are resolved through this methodology. These concerns are essential for the accurate detection of counterfeits. The framework is a powerful instrument for practical applications in domains such as digital forensics, security monitoring, and media authentication due to its adaptability and scalability. However, it is resource-intensive and requires the use of massive datasets annotated with numerous comments.

REFERENCES:

1. Sharma, P., & Rao, M. (2020). Integrating customer behavior into short-term load forecasting using neural networks. International Journal of Electrical Power & Energy Systems, 115, 105437.
2. Li, Y., Sun, X., & Wang, Y. (2020). Deep learning-based video forgery detection: A review. Multimedia Tools and Applications, 79(47), 35415-35439.
3. Zhang, J., Chen, W., & Li, H. (2021). Object-level forgery detection in videos via dual-stream convolutional neural networks. IEEE Transactions on Circuits and Systems for Video Technology, 31(2), 635-647.
4. Kumar, R., & Singh, A. (2021). A deep convolutional neural network for video tampering detection using spatiotemporal features. Journal of Visual Communication and Image Representation, 77, 103059.
5. Chen, L., Wu, J., & Lu, Y. (2021). Detecting copy-move forgeries in high-resolution videos using multi-scale CNN. Signal Processing: Image Communication, 94, 116203.
6. Gupta, S., & Verma, P. (2022). Deep learning-based framework for forgery localization in videos using attention mechanisms. Pattern Recognition Letters, 152, 15-22.

7. Li, M., & Zhao, Y. (2022). A hybrid deep CNN and LSTM approach for video forgery detection. Multimedia Tools and Applications, 81, 25637-25656.
8. Wang, Q., Liu, F., & Xu, Z. (2022). Robust object-level video forgery detection via deep feature fusion. IEEE Access, 10, 23842-23853.
9. Chen, S., & Zhang, X. (2023). High-resolution video forgery detection using dual-branch deep CNN with temporal consistency. IEEE Transactions on Information Forensics and Security, 18, 1120-1133.
10. Patel, D., & Shah, R. (2023). Deep convolutional neural network with residual learning for object-level video tampering detection. Multimedia Tools and Applications, 82, 29811-29830.
11. Li, F., & Huang, J. (2023). Attention-guided deep CNN for detecting object-level manipulations in videos. Neurocomputing, 521, 138-149.
12. Kumar, N., & Reddy, P. (2024). A novel deep learning framework for video forgery detection using spatial-temporal feature enhancement. Signal Processing, 194, 108491.
13. Zhang, T., & Liu, Y. (2024). Deep learning-based video forgery detection with enhanced object-level feature extraction. IEEE Transactions on Multimedia, 26, 402-413.
14. Chen, H., & Wu, Q. (2024). Multi-scale deep CNN for detecting forged objects in high-resolution videos. Journal of Electronic Imaging, 33(2), 023012.
15. Singh, K., & Mehta, S. (2024). Object-level forgery detection in videos using deep convolutional networks and temporal attention. IEEE Transactions on Circuits and Systems for Video Technology, 34(5), 1450-1461.

DOI : https://doi.org/10.5281/zenodo.16810423