International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

An Efficient Technique for Detection and Localization of the Forgery in Digital Videos

Manpreet Kaur Aulakh[1*], Dr. Navdeep Kanwal[2†] and Dr. Manish Bansal[3†]
[1*]Department of Computer Science and Engineering, Punjabi University, Patiala, 147002, Punjab, India
[2]Department of Computer Science and Engineering, Punjabi University, Patiala, 147002, Punjab, India
[3]Department of Computer Science, Baba Farid Group of Institutions, Bathinda, 151001, Punjab, India

Abstract - Video forgery detection is challenging because small frame-level changes like insertion, duplication, and deletion, disturb the temporal motion continuity of video sequences. This paper proposes a reliable technique based on Flow-Strain Energy (FSE) for identifying and locating forged frames. The method simultaneously records local deformation and motion intensity in video sequences by combining strain tensor analysis and optical flow. By calculating dense optical flow between successive frames, optical flow energy and primary strain are estimated. A single FSE description that characterizes motion-deformation features is created by combining these elements. Statistical methods such as mean, standard deviation, and Z-score normalization are used to assess temporal variance in FSE. By clustering consecutive outlier frames that surpass an adaptive thresh- old, forged regions can be accurately localized. The suggested technique produces a localization accuracy of 94.21%, an F1-score of 97.03%, and an overall detection accuracy of 95.92%. The effectiveness of the proposed method in detecting frame-level forgeries and its resilience to changes in illumination and compression issues are demonstrated through experiments on real and altered videos. Furthermore, the framework is suitable for practical video forensic applications due to its excellent processing performance.
Keywords: Video Forgery Detection, Inter-frame Forgery, Forgery Localization, Flow-Strain Energy, Optical Flow, Outlier Detection

1 INTRODUCTION

Video content is now readily available and modifiable due to the quick development [1] of digital video recording and editing capabilities [2]. As a result, verifying the integrity and authenticity of video data [3] has become crucial in fields including forensic investigation, surveillance [4], and legal evidence analysis [5]. Among other types of alterations, frame-level forgeries—such as frame insertion, deletion, and duplication—are particularly challenging to detect because they disrupt temporal consistency without affecting visual appearance [6].

Most traditional methods for identifying video forgeries focus on motion information, pixel-level differences, or compression artifacts [7].However, because they cannot accurately capture the motion behavior in the video [8], they frequently fail in complex motion sequences, different lighting conditions, and severely compressed recordings. In particular, conventional optical flow-based methods [9] primarily record motion data but neglect to take manipulation-induced structure deformation patterns [10]. To overcome these constraints, this study proposes an inter-frame video fraud detection and localization system based on Flow-Strain Energy (FSE). The basic idea is to use strain tensor analysis and optical flow [11] to simultaneously analyze local deformation and motion intensity. The starin tensor models local spatial deformation inside the motion field, whereas optical flow records pixel-by-pixel motion over successive frames. Their integration results in a single FSE description that provides a more comprehensive perspective of motion dynamics.

In order to identify abnormal changes caused by fabricated frames, statistical techniques are employed to examine temporal fluctuations in the FSE signal. Significant changes from normal motion behavior are found using a Z-score-based outlier detection technique. The location of the counterfeit is then precisely determined by grouping these identified anomalies.

The proposed method is appropriate for real-time forensic applications since it is unsupervised, computationally efficient, and does not require training data. The efficacy and resilience of the method in detecting various kinds of frame-level forgeries under diverse circumstances are demonstrated by experimental findings on a variety of video datasets.

The present paper makes main contributions, such as:

- A Flow-Strain Energy (FSE) descriptor is created by integrating optical flow energy and strain tensor deformation to capture motion and structural anomalies in video sequences.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

- A lightweight, training-free statistical detection approach based on Z-score normalization is proposed to detect temporal anomalies caused by frame-level forgeries without the requirement for supervised learning.
- To precisely identify and localize tampered regions, a successful forgery localization approach based on consecutive outlier grouping is employed.
- The suggested approach demonstrates robustness and good performance when evaluated on two benchmark datasets and contrasted with state-of-the-art techniques.

This paper is organized as follows for the remainder: Section 2 reviews recent approaches to video forgeries detection. Section 3 provides a detailed presentation of the proposed methodology, covering every algorithmic step. Section 4 presents the results of the experiment and performance analysis. In Section 5, the state-of-the-art techniques are compared. Section 6 concludes the work and discusses possible future directions for research on video forgery detection.

2 RELATED WORK

This section provides a detailed explanation of constraints of current video forgery detection methods, which were briefly covered in the introduction. A number of techniques for identifying frame-level forgeries, including insertion, deletion, and duplication, have been presented out during the past decade. The main features of representative approaches in terms of forgery type handled, localization capability, classification strategy, underlying methodology, outlier detection technique, datasets used, evaluation metrics, and related research gaps are summarized in Table 1. Even with significant improvements, many existing techniques still have issues like large processing costs and poor performance on real-world videos. The majorities of techniques rely on supervised learning and require sizable labeled datasets. Comparative analyses also demonstrate how difficult it is for many techniques to deliver accurate frame-level localization in many video scenarios. Additionally, they find it difficult to achieve a balance between efficiency, generalization, and accuracy.

3 PROPOSED METHODOLOGY

The proposed method detects frame-level video forgeries by analyzing temporal changes in motion and deformation patterns between consecutive frames. In authentic videos, motion changes smoothly over time, while forged videos introduce sudden changes due to frame duplication, deletion, or insertion. A Flow-Strain Energy (FSE) descriptor is used to model this behavior by combining optical flow and deformation information. Optical flow captures pixel-wise motion, while the strain tensor represents local structural changes in the motion field. Statistical methods are then applied to the temporal variation of the FSE signal to detect abnormal frames.

The proposed method's overall pipeline is shown in Fig. 1. It includes preprocessing, optical flow computation, Flow-Strain Energy estimation, temporal variation analysis, Statistical Outlier Detection and Forgery Detection with Localization.

3.1 Preprocessing

Let I(x,y,t) be the representation of the input video, where t=1,2,…,T.

Table 1 An Overview of the existing Video Forgery Detection Techniques

S. No	Paper	Year	Type of Forgery	Methodology	Outlier Detector	Dataset	Measurement Parameters	Research Gaps
1	[12]	2013	Frame Duplication	Histogram Difference	Pre-defined Threshold	Own Dataset	Precision= 0.849, Recall= 1, Accuracy=1	Consider only one type of forgery
2	[13]	2014	Frame Insertion and	No-Negative tensor	By Observing	Own Dataset	Deletion: precision= 88.64,% Recall=	Evaluated only on static background videos; complex

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

			Deletion	factorization	Graph		86.67%, Insertion: Precision= 100%, Recall=99%	motion scenarios remain unaddressed.
3	[14]	2014	Copy-move in temporal domain	Structural Similarity	_	Own Dataset	Precision= 0.997, Recall= 1	Duplicated frames are mi-detected in long still scenes; High Computational Cost
4	[15]	2015	Frame Insertion and Deletion	Inconsistency of Correlations between LBP frames	Tchebyshev inequality twice	Shanghai Jiao Tog University Database	Precision=88.16 %, and Recall=85.80%	No Forgery Type Classification. Need large dataset.
5	[16]	2016	Frame Duploication	Normalized Cross-Correlation of Moment Features	Mean Squared error	SULFA Dataset	Detection Accuracy= 82% and Localization Accuracy= 85%	Precision rate is low.
6	[17]	2016	Frame Duplication, Deletion and Insertion	Histogram Intersection method	Box Plot	Own dataset	Recall= 090.4% and Precision= 95.2%	Less effective under heavy compression and camera shake.
7	[18]	2017	Frame Dropping	Convolutional 3D Neural Network (C3D)	Threshold	World Dataset	Accuracy= 99.92%	Not address shot boundaries and duplication in looping cases.
8	[19]	2017	Frame Duplication	Hash Value	-	Downloaded from Internet	Accuracy= 100%	Only examine the video files with MJPEG codec.
9	[20]	2017	Frame Deletion	Multi-scale Mutual Information	Modified generalized ESD test	Own dataset	Detection Accuracy= 97.5% and Localization Accuracy=82.1 %	Low Localization Accuracy and not consider frame deletion and duplication type forgery.
10	[21]	2018	Frame and Region duplication	Correlation coefficient and CoV	_	From SULFA dataset and internet	Precision= 1.0, Recall= 0.990, Accuracy= 0.995	Not locate the position of forgery. Not consider other types of forgery.
11	[22]	2018	Frame Duplication,	Histogram Differenc	Selected Thresh	Own Dataset	Precision= 98.07%, Accuracy=	Fails to handle shots that are obtained erroneously,

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

			Insertion and Deletion	e, SURF with FLANN matching	old		99.01%, Recall= 100%	including scene changes.
12	[23]	2018	Frame Insertion, Deletion, Duplication and Shuffling	VFI and footprints	ESD algorithm	Own Dataset	Accuracy= 93.6% to 95.4% and F1 Score= 96.1% to 96.8%	Not effective for inter-frame forgery with synthetic zoomed frames.
13	[24]	2019	Frame Deletion	Coding pattern analysis and classification	-	Downloaded Online	Precision= 0.865, Recall= 0.917, Accuracy= 0.883 and F1 score= 0.888	Not locate the position of forgery. Not consider other types of forgery.
14	[25]	2020	Frame Duplication	Motion vector and SIFT features	Random Sample Consensus Algo	Own Dataset	Precision rate= 99.9%, Recall rate= 99.7%, and Accuracy= 99.8%	Consider only one type of forgery. Not explored performance parameters.
15	[26]	2021	Frame Duplication	I3D and Siamese RNN	-	VIRAT and MFC datasets	Accuracy= 86.6% to 93%	Transfer learning requires improvement.
16	[27]	2021	Frame Insertion, Deletion, Duplication, Copy-move and Splicing	NBAP and PCT with GoogleNet model	-	REWIND, VTL and SULFA datasets	Detection Accuracy= 97.07%, Precision= 96.82%, and Recall= 94.40%	Not suitable for real-time video screening.
17	[28]	2022	Frame Insertion and Deletion	3DCNN with difference layer	MS-SSIM	UFC-101 and VIFFDS datasets	Precision= 0.96, Recall= 0.95, Accuracy= 0.98	Frame Duplication cannot be detected. Need benchmark dataset.
18	[29]	2023	Frame Duplication	LBP. Haralick and custom Haralick features	-	Own dataset	True Positive Rate= 99.79%, Detection Accuracy= 99.32%	Benchmark dataset required. Border frame detection issue.
19	[30]	2023	Frame Insertion and Deletion	Compression domain features	Tukey Box plots	Own dataset	Precision= 0.9589, Recall= 0.9655, F1 = 0.9622	Fixed GOP size limitation. No forgery classification.
20	[31]	2024	Frame Duplication, Insertion	HoG, Uniform and rotation	-	Customized dataset from	Overall Accuracy 99%	Unable to locate frame duplication and frame shuffling attack.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

			and Deletion	invariant LBP		SULFA Dataset		
21	[32]	2024	Frame Deletion	Noise Transfer matrix analysis	-	Own customized dataset	TPR of 0.4 with an FPR of 0	Noise features are highly content dependent. Specially detect integral GOP deletions in static scenes.

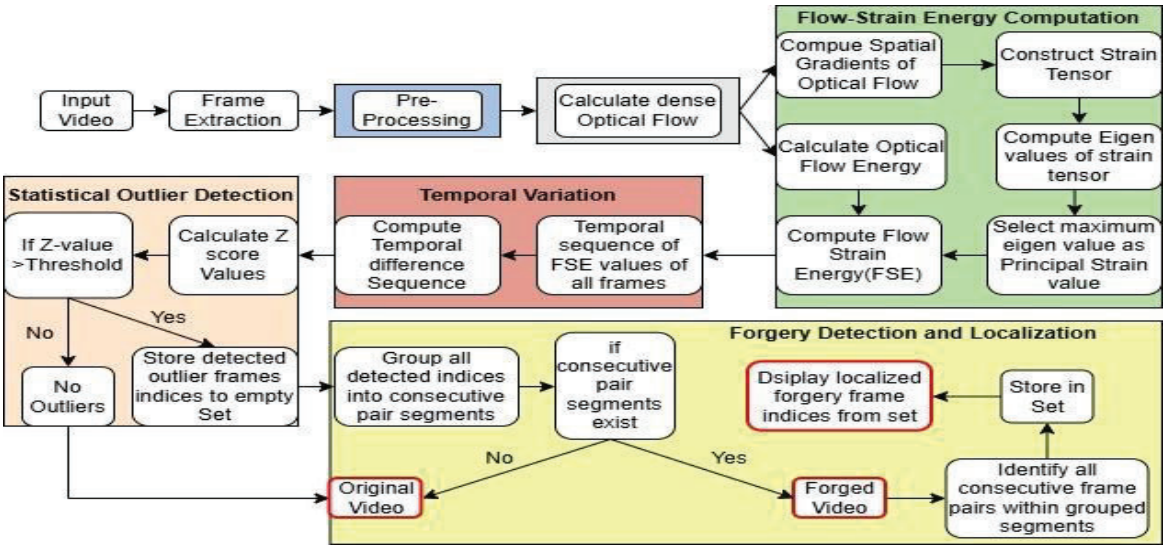

Fig. 1 Proposed Flow-Strain Energy (FSE) Framework for frame-level Video Forgery Identification and Localization

3.1.1 Gray-scale Conversion

The RGB frames are converted into gray-scale to preserve luminance information using Eq. (1):

$$I_g(x,y,t) = 0.299R(x,y,t) + 0.587G(x,y,t) + 0.114B(x,y,t) \quad (1)$$

3.1.2 Gaussian Smoothing

Gaussian filtering is applied to reduce noise and improve the stability of gradient computation. The Gaussian kernel is expressed in Eq. (2):

$$G(x,y) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \quad (2)$$

Eq. (3) is used to generate the smoothed frame:

$$I_s(x,y,t) = I_g(x,y,t) * G(x,y) \quad (3)$$

3.1.3 Histogram Equalization

The cumulative distribution function, which is defined in Eq. (4), is used to enhance contrast:

$$I_e(x,y,t) = (L-1)\sum_{k=0}^{I_s(x,y,t)} p(k) \quad (4)$$

3.1.4 Normalization

Intensity values are normalized using Eq. (5) to maintain consistent scaling:

$$I_n(x,y,t) = \frac{I_e(x,y,t) - I_{min}}{I_{max} - I_{min}} \quad (5)$$

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

3.2 Optical Flow Computation

As stated in Eq. (6), optical flow is calculated using the brightness constancy assump- tion, which asserts that a moving pixel's intensity stays almost constant across successive frames.

$$I(x, y, t) = I(x + u, y + v, t + 1) \quad (6)$$

Using Taylor series expansion, the relationship between spatial and temporal inten- sity changes is derived to obtain the optical flow constraint equation, as given in Eq. (7).

$$\frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v + \frac{\partial I}{\partial t} = 0 \quad (7)$$

Dense optical flow [11] is then computed between consecutive frames to estimate pixel-wise motion throughout the frame, as defined in Eq. (8).

$$F(x, y) = (u(x, y), v(x, y)) \quad (8)$$

3.3 Flow-Strain Energy Computation

The Flow-Strain Energy (FSE) is a combination descriptor that records both local deformation and motion intensity in a video clip. It is computed using the spatial variation of optical flow and strain-based deformation analysis. The optical flow field is first analyzed using spatial derivatives to understand how velocity changes between adjacent pixels.

The spatial variation of motion is represented by the velocity gradient tensor, which describes how the optical flow components change in both horizontal and vertical directions:

$$\nabla F = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix} \quad (9)$$

Eq. (9) computes local motion variation, where $\frac{\partial u}{\partial x}$ and $\frac{\partial v}{\partial y}$ represent normal motion changes, while $\frac{\partial u}{\partial y}$ and $\frac{\partial v}{\partial x}$ capture shears deformation in the motion field.

The strain tensor is obtained from the symmetric part of the velocity gradient tensor, which removes rotational motion and keeps only deformation information. It represents local stretching, compression, and distortion in the motion field:

$$S = \frac{1}{2}\begin{bmatrix} 2\frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \\ \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} & 2\frac{\partial v}{\partial y} \end{bmatrix} \quad (10)$$

The deformation behavior in the motion field is represented by Eq. 10, where the diagonal elements show expansion or contraction, and the off-diagonal elements represent shear deformation.

The principal strain values are obtained by computing the eigenvalues of the strain tensor, as given in Eq. (11). These values represent the maximum and minimum deformation directions.

$$\lambda_{1,}\ \lambda_2 = eig(S) \quad (11)$$

The largest eigenvalue is then selected as the dominant measure of deformation:

$$\lambda_{max} = \max(\lambda_{1,}\ \lambda_2) \quad (12)$$

Eq. (12) represents the strongest local structural deformation in the motion field and is highly sensitive to sudden changes caused by forgery. The optical flow energy, given in Eq. (13), measures the motion intensity within a block by calculating the average squared magnitude of motion vectors.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

$$E_{flow} = \frac{1}{N} \sum_{(x,y)\in B} (u(x,y)^2 + v(x,y)^2) \quad (13)$$

Optical flow energy measures the strength of motion within a block, where higher values indicate larger pixel displacement. Flow-Strain Energy (FSE) combines optical flow energy and deformation information obtained from the optical flow field. It uses motion strength (E_{flow}) and structural deformation ($\lambda_{\max}$) together in a single descriptor. FSE is calculated as a weighted sum of optical flow energy and principal strain, as given in Eq. (14).

$$FSE_B = \alpha E_{flow} + \beta \lambda_{max} \quad (14)$$

Where α and βcontrol their relative contribution.

The frame-level Flow-Strain Energy is calculated by averaging the block-level values over all blocks in a frame, as given in Eq. (15):

$$FSE_t = \frac{1}{M} \sum_{i=1}^{M} FSE_B^{(i)} \quad (15)$$

This reduces local noise, produces a stable temporal energy signal, and provides an overall view of motion and deformation for the entire frame.

3.4 Temporal Variation

Temporal inconsistency is calculated using the difference between successive frames. Eq. 16 represents the variation in Flow-Strain Energy between consecutive frames and captures sudden changes in motion and deformation patterns.

$$\Delta FSE_t = FSE_t - \Delta FSE_{t-1} \quad (16)$$

3.5 Statistical Outlier Detection

The temporal difference sequence is normalized using Z-score transformation to measure how much each frame differs from normal motion behavior.

First, the mean of the temporal Flow-Strain Energy difference sequence is calculated as given in Eq. (17). It represents the average variation in motion and deformation throughout the video sequence.

$$\mu_{\Delta FSE} = \frac{1}{T-1} \sum_{t=2}^{T} \Delta FSE_t \quad (17)$$

Next, the standard deviation is calculated as given in Eq. (18). It measures how much the temporal variations differ from the mean and reflects the stability of motion across frames.

$$\sigma_{\Delta FSE} = \sqrt{\frac{1}{T-1} \sum_{t=2}^{T} (\Delta FSE_t - \mu_{\Delta FSE})^2} \quad (18)$$

Finally, Eq. (19) standardizes the Flow-Strain Energy variation using the calculated mean and standard deviation. This helps identify abnormal motion changes compared to the overall distribution of frame-level motion dynamics.

$$Z_t = \frac{\Delta FSE_t - \mu_{\Delta FSE}}{\sigma_{\Delta FSE}} \quad (19)$$

Frames are identified as outliers when the Z-score exceeds a predefined threshold:

$$|Z_t| > T_{thr} \quad (20)$$

Eq. (20) identifies outlier frames with unusual motion and deformation changes that may indicate frame insertion, deletion, or duplication.

3.6 Forgery Detection with Localization

The outlier frames detected from the Z-score sequence are used for forgery detection and temporal localization. First, all frame indices satisfying $|Z_t| > T_{thr}$ are collected. These indices are then grouped into consecutive segments, where each group contains neighboring anomalous frames in time.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

If no consecutive frame pairs are found within the grouped segments, the detected anomalies are treated as isolated changes caused by natural motion, and the video is classified as an *Original Video*.

Otherwise, if consecutive frame pairs are found, it indicates consistent irregular changes over time. Each of these successive pairs is gathered as potential forged frames in a set F. In this instance, the video falls under the category of *Forged Videos*.

Finally, the frame indices of the tampered regions are contained in the set F, which provides precise temporal localization of the forgery.

3.7 Parameter selection and Threshold Setting

The weighting coefficients α and β, block size B, statistical threshold T_{thr} and Gaussian smoothing variance σ all affect how well the suggested framework performs. These settings are chosen empirically in order to achieve a compromise between robustness against false alarms and detection sensitivity.

While a larger value of the Gaussian kernel variance σ eliminates crucial motion information required for forgery detection, a smaller value does not effectively reduce noise. Consequently, σ is adjusted to 1.0 since it offers an adequate balance between reducing noise and maintaining significant motion details in the video.

The block size B is selected to 32×32 because it provides a suitable trade-off between computation and detail. Larger blocks make it more difficult to identify forged areas, whereas smaller blocks may identify small changes but also add noise.

The weighting factors are adjusted at $\alpha = 0.6$ and $\beta = 0.4$ for computing Flow-Strain Energy, giving motion energy a slightly higher importance than deformation. This configuration is selected because motion magnitude offers stable global dynamics and strain adds fine structural sensitivity.

The sensitivity of forgery detection is controlled by the statistical threshold T_{thr} which is employed in Eq. (20). It is a crucial component of the suggested approach. Comprehensive tests were performed on 100 video sequences in order to determine its optimal value. These comprise authentic videos as well as fabricated ones, such as frame duplication, deletion, and insertion. The tests are conducted in a variety of settings, including variations in lighting, motion intensity, and compression levels.

As experiments are conducted, it is observed that:

- At lower threshold values ($T_{\text{thr}} < 4$), sensitivity is increased, but natural motion variations result in more false positives.
- Higher threshold values ($T_{\text{thr}} > 6$) reduce false alarms, but they cannot detect subtle forgeries such as single-frame duplication.

The best threshold, $T_{\text{thr}} = 5$, which provides a constant trade-off between robustness and detection accuracy for all assessed movies, is chosen as a result of this research. The selected parameters exhibit consistent performance and high generalization ability in a range of video scenarios, indicating that they are suitable for real-world forensic applications.

3.8 Algorithm

The various steps of the proposed technique are summarized by Algorithm 1.

Algorithm 1 Flow-Strain Energy (FSE) based Video Forgery Identification and Localization:

Input: Video Sequence V

Output: Forgery decision and forged frame indices $\mathcal{F}$

1. Extract all frames from the input video sequence and denote them as $\{I_t\}_{t=1}^{T}$
2. Perform preprocessing on each frame including gray-scale conversion, Gaussian smoothing, histogram equalization, and normalization to obtain enhanced frames$\{I_n^t\}$.
3. Initialize an empty set $\mathcal{Z}$ to store detected outlier frame indices.
4. **for** each consecutive frame pair (I_n^{t-1}, I_n^t) where t = 2 to T **do**

   Compute dense optical flow field $F_t(x, y)$ consisting of horizontal and vertical motion components.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

Compute spatial gradient of the optical flow field to obtain motion variation information.

Construct the strain tensor from the symmetric part of the flow gradient field.

Compute eigenvalues of the strain tensor and select the maximum eigenvalues as the principal deformation measure.

Compute optical flow based energy on squared magnitude of motion vectors within the frame.

Compute Flow-Strain energy (FSE) as a weighted combination of optical flow energy and principal strain value.

end for

5. Construct temporal sequence of FSE values for all frames.
6. Compute temporal difference sequence ΔFSE_t between consecutive frames.
7. Compute standard deviation $\sigma_{\Delta FSE}$ and mean $\mu_{\Delta FSE}$ of the temporal difference sequences.
8. Normalize the temporal variation using Z-score transformation to obtain Z_t values.
9. **for** each frame index t **do**

if $|Z_t| > T_{\text{thr}}$ **then**

Store frame index t in the set $\mathcal{Z}$.

end if

end for

10. Group all detected indices in $\mathcal{Z}$ into consecutive segments $G_1, G_2, \ldots, G_k$.

if no consecutive frame pairs exist in any group **then**

Declare the video as **Original Video.**

else

Identify all consecutive frame pairs within grouped segments.

Store these pairs as forged frame candidates in set $\mathcal{F}$.

Declare the video as **Forged Video**.

Report $\mathcal{F}$ as the localized forgery frame indices.

end if

3.9 Advantages of the Proposed FSE-Based Method

- A statistical analysis-based unsupervised and training-free framework.
- Computationally effective because of lightweight operations and block-based processing.
- Physically understandable by combining deformation and motion (Flow-Strain Energy).
- Extremely sensitive to small temporal irregularities in motion patterns.
- Resistant to variations in illumination since preprocessing techniques are employed.
- Efficient in identifying various forms of inter-frame forgeries, such as duplication, deletion, and insertion.
- Gives precise temporal localization of regions of forged frames.

4 EXPERIMENTAL RESULT

The experimental evaluation of the suggested Flow-Strain Energy framework for inter- frame video forgery detection is presented in this section. Benchmark inter-frame datasets with frame insertion, deletion, and duplication forgeries are used for the evaluation. To guarantee a fair and trustworthy analysis, the method's performance is evaluated using standardized measures.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

4.1 Dataset Description

A total of 174 video sequences- 53 real movies and 121 fake video samples- are used for the experimental investigation. The evaluation is carried out on two datasets, as summarized in Table 2, which include publicly available inter-frame forgery datasets, TDTVD [33] and VIFFD [34].

The videos are recorded in different indoor and outdoor environments under both static and dynamic conditions. They include variations in camera types, resolutions, lighting, and motion, covering scenes from completely still to highly dynamic. The videos are available in formats such as AVI, and MP4, with resolutions ranging from 320×240 to 720×404 pixels, durations between 5 and 18 seconds, and frame rates of 25, and 30 fps.

Table 2 Features of the Datasets Used for Assessment

Dataset	Source	Type of Forgery	Resolution	Length	Frame Rate	No. of Videos	Format
VIFFD [34]	-	Frame Duplication, Frame Insertion and Frame Deletion	720×404	5-10 s	25 fps	30 Original , 52 forged	AVI, mp4
TDTVD [33]	24 from YouTube, VTD Dataset and 16 from SULFA	Frame Duplication, Frame Insertion and Frame Deletion	320×240 640×360	6-18 s	30 fps	23 Original and 69 Forged	AVI

4.2 Standard Evaluation Metrics for the Proposed Framework

Standard metrics such as Detection Accuracy, Localization Accuracy, Precision, Recall, and F1-Score are used to assess the effectiveness of the suggested approach.

4.2.1 Detection Accuracy

Detection Accuracy, given in Eq. (21), measures how correctly the videos are classified as forged or original:

4.2.2 Localization Accuracy

$$\text{Detection Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \qquad (21)$$

Localization Accuracy evaluates how accurately the forged frame locations are identified, as given in Eq. (22):

$$Localization\ Accuracy = \frac{N_{correct\_frames}}{N_{actual\_forged\_frames}} \qquad (22)$$

4.2.3 Precision

Eq. (23) expresses precision as the percentage of successfully identified forged frames among all detected frames:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \qquad (23)$$

4.2.4 Recall

According to Eq. (24), recall quantifies the percentage of real forged frames that are accurately identified:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \qquad (24)$$

4.2.5 F1-Score

F1-Score, as described by Eq. (25) is the harmonic mean of Precision and Recall:

$$\text{F1} - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \qquad (25)$$

Where:

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

TP = True Positives (correctly detected forged frames/videos)

TN = True Negatives (correctly detected original videos)

FP = False Positives (original detected as forged)

FN = False Negatives (forged detected as original)

$N_{correct\_frames}$ represents the number of correctly localized forged frames, and

$N_{actual\_forged\_frames}$ represents the total number of actual forged frames.

4.3 Frame-Level Performance Evaluation

The proposed Flow-Strain Energy framework is evaluated for its ability to detect and localize frame-level forgeries across different datasets. To examine its generalization capability, experiments are conducted using selected subset of videos from the VIFFD [34] and TDTVD [33] datasets.

Confusion matrices (Fig. 2), which clearly depict true positives, false positives, true negatives, and false negatives, are used to analyze the results. The standard measures covered in the preceding section are used to evaluate the suggested method's performance.

Table 3 and Fig. 3 present the results, which show consistent performance under various motion conditions and video types. These results demonstrate how resilient and flexible the suggested strategy is in a variety of situations.

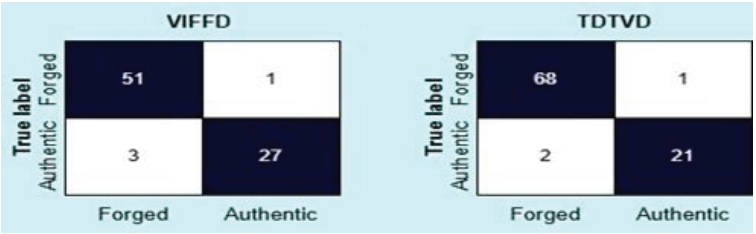

Fig.2 Confusion Matrix across Datasets

Table 3 Dataset-wise Performance Evaluation using Standard Metrics

Parameters	VIFFD [34]	TDTVD [33]
Detection Accuracy	95.12%	96.73%
Localization Accuracy	94.23%	94.20%
Precision	94.44%	97.14%
Recall	98.07%	98.55%
F1 Score	96.22%	97.84%

These findings demonstrate how well the suggested approach locates altered frames and detects forgeries, with very few false alarms and missed detections.

To demonstrate the efficacy of the suggested approach, examples of three types of inter-frame forgeries are presented. For the deletion case, the video *huabin_walk2_1* from the VIFFD dataset [34] is considered, where frames 75 to 150 are removed. The original and forged frame sequences are shown in Fig. 4. This deletion introduces a temporal discontinuity, which appears as sharp peak in the graph, as shown in Fig. 5.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

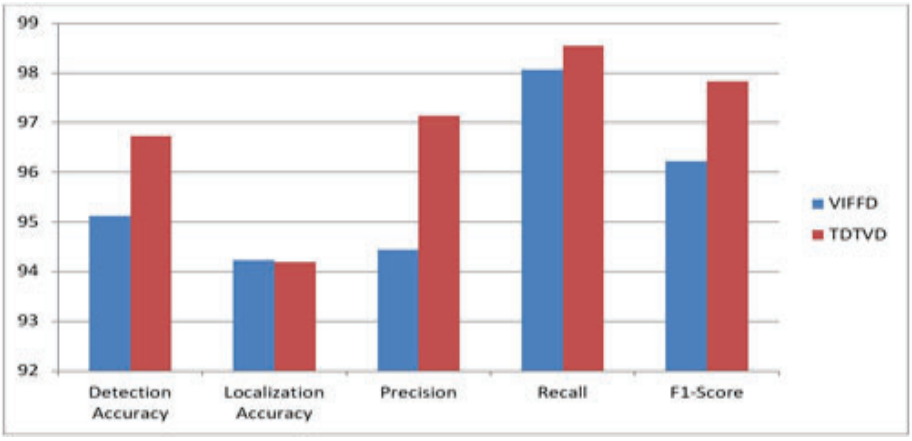

Fig.3 Graphical Comparison of Performance across two datasets

Fig.4 Original and Forged Frame Sequences for Frame Deletion case example

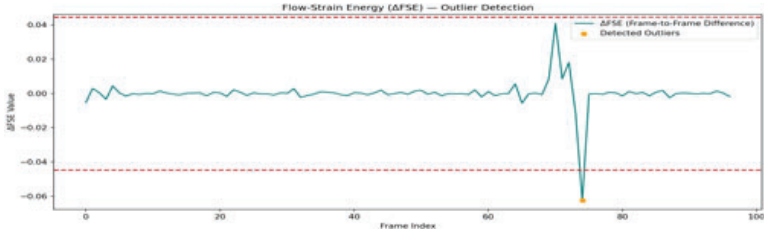

Fig.5 Flow-Strain Energy (FSE) Outlier Detection graph for Frame Deletion case example

For the duplication case, the video *Tampered_EOP_05_original_framedup* from the TDTVD dataset [33] is used, where frames 31 to 100 are duplicated and inserted between frames 259 and 333. The corresponding original and forged frame sequences are shown in Fig. 6. This duplication results in temporal inconsistency, producing two sharp peaks in the graph, as illustrated in Fig. 7.

Fig.6 Original and Forged Frame Sequences for Frame Duplication case example

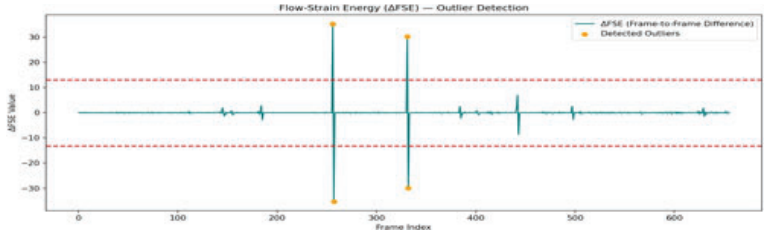

Fig.7 Flow-Strain Energy (FSE) Outlier Detection graph for Frame Duplication case example

Similarly, for the insertion case, the video *Tampered_EOP_can_220_man(1)_frameins* from the TDTVD dataset [33] is analyzed, where frames are inserted from 201 to 300. The original and forged frame sequences

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

are shown in Fig.8. This insertion creates temporal discontinuities at the boundaries, which are reflected as two sharp peaks in the graph, as shown in Fig.9.

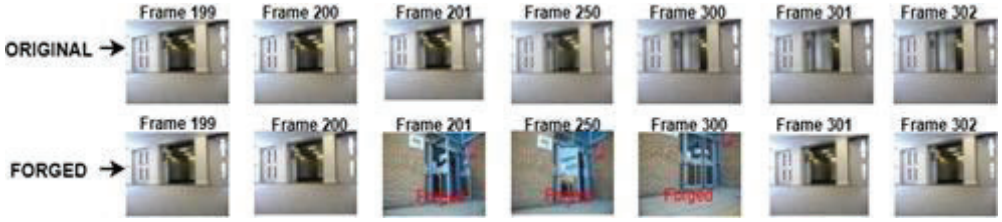

Fig.8 Original and Forged Frame Sequences for Frame Insertion case example

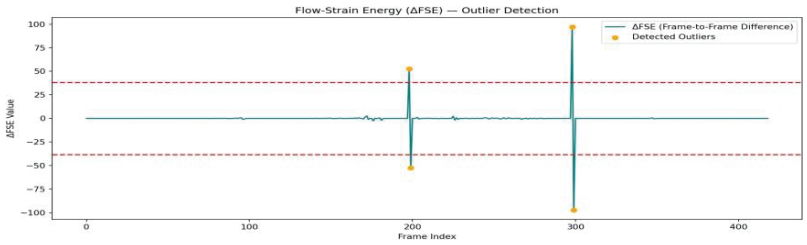

Fig.9 Flow-Strain Energy (FSE) Outlier Detection graph for Frame Insertion case example

5 COMPARISON WITH EXISTING TECHNIQUES

Table 4 compares the suggested technique with state-of-the-art techniques based on forgery types, localization capability, and performance measures such as detection accuracy, F1-score, recall, and precision. Most existing methods are limited to specific forgery types, such as both insertion and deletion [35, 36], or only insertion [37], or only deletion [38]. In contrast, the pro- posed method handles all major inter-frame forgeries, including duplication, insertion, and deletion.

The suggested approach performs consistently on datasets, with balance F1-Score, precision, and recall as well as detection accuracies of 95.12% on VIFFD and 96.73% on TDTVD. Additionally, it offers efficient frame-level localization with 94.23% and 94.20% accuracy on VIFFD and TDTVD, respectively.

Overall, the results indicate that the proposed Flow-Strain Energy-based frame- work offers a strong balance between detection performance, localization accuracy, and robustness across different video conditions.

6 CONCLUSION AND FUTURE WORK

This paper presents a Flow-Strain Energy (FSE)-based method for identifying and locating inter-frame video forgeries. The method combines motion information from optical flow with deformation information from strain analysis to capture changes over time. The FSE descriptor provides a consistent representation of motion in each frame. To detect abnormal frames, the changes in the FSE descriptor over time are analyzed using simple statistical methods. Effective localization of forged segments corresponding to frame duplication, insertion, and deletion is made possible by grouping consecutive outliers. Experimental evaluation on standard datasets shows that the proposed method delivers consistent detection and localization performance while being computationally inexpensive and independent of training data.

Future research will concentrate on enhancing the method's performance in challenging conditions such as heavy compression, camera motion, and low-quality videos. It will also explore better adaptive threshold techniques. Since there are limited benchmark datasets, there is a need to develop new and diverse high-quality datasets. In addition, the method will be extended to handle more complex forgeries, including deepfake-based manipulations.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

Table 4 Performance Comparison with State-of_the_Art Video Forgery Detection Techniques

Technique	Year	Type of Forgery Detected	Dataset Used	Localization	Localization Accuracy	Detection Accuracy	Precision	Recall	F1 Score
[35]	2022	Frame Insertion and Deletion	VIFFD	Yes	Not Given	98%	98%	98%	98%
[36]	2022	Frame Insertion and Deletion	VIFFD	Yes	Not Given	83%	83%	87%	84%
[37]	2022	Frame Insertion	VIFFD	Yes	Not Given	86.5%	78.94%	100%	87%
[38]	2023	Frame Deletion	TDTVD	No	Not Given	96.25%	No	95.12%	96.3%
Proposed	–	Frame Insertion, Deletion and Duplication	VIFFD	Yes	94.23%	95.12%	94.44%	98.07%	96.22%
Proposed	–	Frame Insertion, Deletion and Duplication	TDTVD	Yes	94.20%	96.73%	97.14%	98.55%	97.84%

Declarations
Funding: This research did not receive funding.

Data availability statement: This study utilizes publicly available benchmark datasets,. The Public datasets are accessible from their respective sources.

REFERENCES

[1] Alfarano, A., Maiano, L., Papa, L., Amerini, I.: Estimating optical flow: A comprehensive review of the state of the art. Computer Vision and Image Understanding **249** (2024) https://doi.org/10.1016/j.cviu.2024.104160
[2] El-Shafai, W., *et al.*: A comprehensive taxonomy on multimedia video forgery detection techniques: challenges and novel trends. Multimedia Tools and Appli- cations **83**, 4241–4307 (2023) https://doi.org/10.1007/s11042-023-15609-1
[3] Tyagi, S., Yadav, D.: A detailed analysis of image and video forgery detection techniques. The Visual Computer **39**, 813–833 (2023) https://doi.org/10.1007/ s00371-021-02347-4
[4] Mohiuddin, S., *et al.*: A comprehensive survey on state-of-the-art video forgery detection techniques. Multimedia Tools and Applications **82**(22), 33499–33539 (2023) https://doi.org/10.1007/s11042-023-14870-8
[5] Javed, A.R., Jalil, Z., Zehra, W., Gadekallu, T.R., Suh, D.Y., Piran, M.J.: A comprehensive survey on digital video forensics: Taxonomy, challenges, and future directions. Engineering Applications of Artificial Intelligence **106**, 104456 (2021) https://doi.org/10.1016/j.engappai.2021.104456
[6] Sharma, H., Kanwal, N., Batth, R.S.: An ontology of digital video forensics: Classification, research gaps & datasets. In: Proceedings of the International Con- ference on Computational Intelligence and Knowledge Economy (ICCIKE), pp. 485–491. IEEE, 2019. https://doi.org/10.1109/ICCIKE47802.2019.9004331
[7] Alsmirat, M.A., A1-Hussien, R.A., A1-Sarayrah, W.T., Jararweh, Y., Etier, M.: Digital video forensics: A comprehensive survey. International Journal of Advanced Inteeligence Paradigms 15(4) (2020) https://doi.org/10.1504/ijaip.2020.106040
[8] Shelke, N.A., Kasana, S.S.: A comprehensive survey on passive techniques for digital video forgery detection. Multimedia Tools and Applications **80**, 6247–6310 (2021) https://doi.org/10.1007/s11042-020-09974-4
[9] Alkawaz, M.H., Veeran, M.A.T., Hajamydeen, A.I., A1-Sanjary, O.I>: An overview of advanced optical flow techniques for copy move forgery detection. In: Proceddings of the IEEE 11th Symposium on Computer Applications & Industrial Electronics (ISCAIE), pp. 319-324. IEEE, 2021. https://doi.org/10.1109/ISCAIE51753.2021.9431772.
[10] Kaur, H., Jindal, N.: Image and video forensice: A critical survey. Wireless Personal Communications 112, 1281-1302 (2020) https://doi.org/10.1007/s11277-020-07102-x
[11] Farneb¨ack, G.: Two-frame motion estimation based on polynomial expansion. In: Image Analysis: 13th Scandinavian Conference, SCIA 2003, pp. 363–370. Springer, Halmstad, Sweden (2003)
[12] Lin, G.-S., Chang, J.-F.: detection of frame duplication forgery in videos based on spatial and temporal analysis. Interantional Journal of Pattern Recogntion and Artificial Intelligence 26 (2013) https://doi.org/10.1142/S0218001412500176
[13] Yin, L., Bai, Z., Yang, R.: Video forgery detection based on nonnegative tensor factorization. In: Proceedings of the 4th IEEE International Conference on Information Scinece and Technology(ICIST). Pp. 148-151 (2014). https://doi.org/10.1109/ICIST.2014.6920352
[14] Li, F., Huang, T.: Video copy-move forgery detection and localization based on structural similarity. In: Proceedings of the 3rd International Conference on Mul- timedia Technology (ICMT 2013). Lecture Notes in Electrical Engineering, vol.278. Springer, ??? (2014). https://doi.org/10.1007/978-3-642-41407-7 7
[15] Zhang, Z., Hou, J., Ma, Q., Li, Z.: Efficient video frame insertion and deletion detection based on inconsistency of correlations between local binary pattern coded frames. Security and Communication Networks **8**, 311–320 (2015) https://doi.org/10.1002/sec.981
[16] Mathai, M., Rajan, D., Emmanuel, S.: Video forgery detection and localization using normalized cross-correlation of moment features. In: Proceedings of the IEEE Southwest Symposium on Image Analysis ans Interpretation (SSIAI), pp. 149-152. IEEE, (2016). https://doi.org/10.1109/SSIAI.2016.7459197
[17] Xu, J., Liang, Y., Tian, X., Xie, A.: A novel video inter-frame forgery detection method based on histogram intersection. In: Proceedings of the IEEE/CIC International Conference on Communications in China (ICCC), pp. 1-6. IEEE, (2016). https://doi.org/10.1109/ICCChina.2016.7636851
[18] Long, C., Smith, E., Basharat, A., Hoogs, A.:A c3d-based convolutional neural network for frame dropping detection in a single video shot. In: Proceddings of the IEEE Conference on Computer Vision and Patteren Recognition Workshops (CVPRW), pp. 1898-1906. IEEE, (2017). https://doi.org/10.1109/CVPRW.2017.237
[19] Andy, S., Haikal, A.: Simple duplicate frame detection of mjpeg codec for video forensic. In: Proceedings of the 2nd International Conference on Information Tech- nology, Information Systems and Electrical Engineering (ICITISEE), pp. 321–324.IEEE, ??? (2017). https://doi.org/10.1109/ICITISEE.2017.8285520
[20] Zhao, Y., Pang, T., Xiaoyun, L., Li, Z.: Frame-deletion detection for static- background video based on multi-scale mutual information. In: Advances in Multimedia Information Processing – PCM, pp. 371–384 (2017). https://doi.org/ 10.1007/978-3-319-68542-7 31
[21] Singh, G., Singh, K.: Video frame and region duplication forgery detection based on correlation coefficient and coefficient of variation. Multimedia Tools and Appli- cations **78**, 11527–11562 (2019) https://doi.org/10.1007/s11042-018-6585-1

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

[22] Zhao, D.-N., Wang, R.-K., Lu, Z.-M.: Inter-frame passive-blind forgery detection for video shot based on similarity analysis. Multimedia Tools and Applications **77**, 25389–25408 (2018) https://doi.org/10.1007/s11042-018-5791-1

[23] Sitara, S.K., Mehtre, B.M.: Detection of inter-frame forgeries in digital videos. Forensic Science International **289** (2018) https://doi.org/10.1016/j.forsciint.2018.04.056

[24] Hong, J.H., Yang, Y., Oh, B.T.: Detection of frame deletion in hevc-coded video in the compressed domain. Digital Investigation **30**, 23–31 (2019) https://doi. org/10.1016/j.diin.2019.06.002

[25] Kharat, J., Chougule, S.: A passive blind forgery detection technique to iden- tify frame duplication attack. Multimedia Tools and Applications **79**, 8107–8123 (2020) https://doi.org/10.1007/s11042-019-08272-y

[26] Munawar, M., Noreen, I.: Duplicate frame video forgery detection using siamese- based rnn. Intelligent Automation & Soft Computing **29**, 927–937 (2021) https://doi.org/10.32604/iasc.2021.018854

[27] Shelke, N.A., Kasana, S.S.: Multiple forgery detection and localization technique for digital video using pct and nbap. Multimedia Tools and Applications **81**, 22731–22759 (2022) https://doi.org/10.1007/s11042-021-10989-8

[28] Gowda, R., Pawar, D.: Deep learning-based forgery identification and localization in videos. Signal, Image and Video Processing **17**, 2185–2192 (2023) https://doi. org/10.1007/s11760-022-02433-7

[29] Mohiuddin, S., Malakar, S., Sarkar, R.: An ensemble approach to detect copy- move forgery in videos. Multimedia Tools and Applications **82**, 24269–24288 (2023) https://doi.org/10.1007/s11042-023-14554-3

[30] Singla, N., Nagpal, S., Singh, J.: A two-stage forgery detection and localiza- tion framework based on feature classification and similarity metric. Multimedia Systems **29**, 1173–1185 (2023) https://doi.org/10.1007/s00530-023-01050-9

[31] Shehnaz, Kaur, M.: Detection and localization of multiple inter-frame forg- eries in digital videos **83**, 71973–72005 (2024) https://doi.org/10.1007/ s11042-024-18263-3

[32] Bao, Q., Wang, Y., Hua, H., Dong, K., Lee, F.: An anti-forensics video forgery detection method based on noise transfer matrix analysis. Sensors **24**(16), 5341 (2024) https://doi.org/10.3390/s24165341

[33] Panchal, H.D., Shah, H.B.: Video tampering dataset development in tempo- ral domain for video forgery authentication. Multimedia Tools and Applications **79**(33–34), 24553–24577 (2020) https://doi.org/10.1007/s11042-020-09205-w

[34] Nguyen, X.H., Hu, Y.: VIFFD - A dataset for detecting video inter-frame forgeries. Mendeley Data (2020). https://doi.org/10.17632/r3ss3v53sj.6

[35] Gowda, R., Pawar, D.: Deep learning-based forgery identification and localization in videos. Signal, Image and Video Processing **17**(5), 2185–2192 (2023) https://doi.org/10.1007/s11760-022-02433-7

[36] Kumar, V., Gaur, M.: Multiple forgery detection in video using inter-frame corre- lation distance with dual-threshold. Multimedia Tools and Applications **81**(30), 43979–43998 (2022) https://doi.org/10.1007/s11042-022-13284-2

[37] Kumar, V., Gaur, M., Kansal, V.: Deep feature based forgery detection in video using parallel convolutional neural network: Vfid-net. Multimedia Tools and Applications **81**(29), 42223–42240 (2022) https://doi.org/10.1007/ s11042-021-11448-0

[38] Panchal, H.D., Shah, H.B.: Multiple forgery detection in digital video based on inconsistency in video quality assessment attributes. Multimedia Systems **29**(4), 2439–2454 (2023) https://doi.org/10.1007/s00530-023-01123-9