Power System Technology
ISSN:1000-3673

Received: 16-10-2025 Revised: 05-11-2025 Accepted: 22-12-2025

Hybrid Deep Feature Fusion Framework for Intelligent Image Forgery Detection

[1]Vanishri V Sataraddi, [2]Dr N P Nethravath, [3]Dr. Anjan Kumar K N

[1, 2] School of Computer Science and Engineering, REVA University, Bengaluru - 560064,Karnataka, India

[1, 3]Department of Computer Science and Engineering, RNS Institute of Technology,Bengaluru-560098, Karnataka,India

Abstract: - Image manipulation detection has emerged as a crucial task in the digital forensics and strategic intelligence domains. The easy accessibility of modern image-editing applications has made tampering more common and difficult to detect, malicious manipulations—particularly copy–move forgeries pose serious threats to the integrity of multimedia data. This paper proposes a hybrid deep-learning framework combining Discrete Cosine Transform (DCT), Scale-Invariant Feature Transform (SIFT), and a lightweight Convolutional Neural Network (CNN) classifier to effectively identify and localize forged regions in digital images. Experimental results demonstrate a detection accuracy of 96.8% under various compression and transformation scenarios, outperforming classical DCT- and SIFT-based methods.

Keywords*: Image Forensics, Copy-Move Forgery, DCT, SIFT, CNN, MATLAB, Image Authentication.*

1. Introduction

Digital imagery is widely used as supporting evidence across domains such as journalism, military intelligence, digital forensics, and legal investigations. However, with the rapid advancement of sophisticated image-editing tools such as Photoshop and GIMP, subtle manipulations can be executed seamlessly, making it extremely challenging to detect alterations through human visual inspection alone. Among the different types of digital image tampering, copy–move forgery remains one of the most prevalent and deceptive techniques, in which a portion of an image is copied and pasted within the same image to conceal or replicate certain information.

Traditional forgery detection Methods are generally classified into two categories

1. **Block-based methods**, such as those suggested by Fridrich et al. [1], divide an image into overlapping blocks and extract discriminative features from each block to detect similarities. Although these techniques are effective for minor shifts, they are often sensitive to rotation and scaling.
2. **Keypoint-based methods**, such as those by Bayram et al. [2], rely on distinctive local descriptors such as scale-invariant feature transform (SIFT) and Speeded-Up Robust Features (SURF) to match corresponding keypoints between duplicated regions. These approaches are more robust to geometric transformations but may fail in smooth or low-texture regions.

Power System Technology
ISSN:1000-3673
Received: 16-10-2025 Revised: 05-11-2025 Accepted: 22-12-2025

With the advent of deep learning, new possibilities have arisen for integrating handcrafted features with learned representations. Feature fusion allows models to leverage the interpretability of classical descriptors along with the adaptability of deep neural networks.

This study introduces a Hybrid DCT–SIFT–CNN framework that exploits frequency, spatial, and deep features collectively. The Discrete Cosine Transform (DCT) extracts frequency-domain information resilient to compression SIFT captures scale- and rotation-invariant keypoints and the Convolutional Neural Network (CNN) performs adaptive classification for forged region localization. The proposed model demonstrates robustness against multiple attacks such as JPEG compression, rotation, scaling, and blurring.

Such a hybrid strategy ensures high accuracy and efficiency in detecting manipulations, making it ideal for applications in strategic analytic services, where reliable image authentication supports critical decision-making and intelligence assessment.

2. Related Work

Image-forgery detection research has evolved significantly, from early handcrafted feature approaches to modern hybrid deep-learning models. Fridrich et al. [1] laid the foundation for copy–move forgery detection using block-based feature correlation their overlapping-block method remains a baseline in digital forensics. Bayram et al. [2] improved efficiency and robustness by incorporating discriminative descriptors and dimensionality reduction, enabling detection under translation and mild geometric transformations. Ferrara et al. [3] introduced colour-filter-array (CFA) artifact analysis, a sensor-level trace that reveals manipulations invisible to pixel-based methods. Cozzolino et al. [4] later presented dense-field matching using Patch Match, achieving high coverage and speed for textured and texture less regions alike.

The introduction of deep learning has transformed landscape. Bappy et al. [5] designed a hybrid CNN–LSTM network that captures both local texture cues and global contextual relationships, greatly improving localization accuracy. Subsequent surveys by Zanardelli [6] and Mehrjardi [7] reviewed these deep approaches, classifying them into convolutional, autoencoder, and transformer categories and noted persistent challenges such as limited datasets and poor cross-domain generalization. Shi et al. [14] advanced this trend with discrepancy-guided reconstruction learning, reducing dependence on pixel-level ground truth, whereas LoMa [18] explored transformer architectures that model long-range dependencies to improve spatial consistency.

Complementary efforts were made to examine the quality and robustness of the dataset. Oliveira et al. [12] evaluated public copy–move benchmarks by emphasizing standardized testing protocols. Wang et al. [9] introduced JPEG compression-aware localization to maintain the performance in lossy formats. Hegazi et al. [10] employed density-based clustering to remove spurious matches from repetitive textures. Yang [11] combined SIFT keypoints with region-growing to achieve precise mask generation, and Fridrich et al. [17] later extended their early work to handle irregular free-form pasted regions.

Recent studies have highlighted the hybrid and transformer-based paradigms. Amiri [19] optimized feature-fusion parameters for efficiency, whereas Xu et al. [20] developed a two-

Power System Technology
ISSN:1000-3673
Received: 16-10-2025 Revised: 05-11-2025 Accepted: 22-12-2025

stream CNN for document image tampering and fusing semantic and noise cues. The Deepfake-Eval benchmark [16] extend forensic evaluation to synthetic media, underscoring the need for cross-dataset generalization. Surveys by Li and Huang [13] and Shi et al. [15] situated forgery detection within broader digital-forensics research, noting the synergy between handcrafted sensor traces (CFA, PRNU, and DCT) and learned representations.

Collectively, these works demonstrate three enduring lessons: frequency-domain descriptors (e.g., DCT and CFA) provide compression robustness keypoint descriptors (SIFT, SURF) ensure geometric invariance and deep networks supply discriminative power and contextual reasoning. Motivated by these insights, this study integrates the DCT, SIFT, and CNN components into a unified hybrid model that balances interpretability, robustness, and computational efficiency for reliable image manipulation detection in strategic analytic applications.

3.Methodology

The proposed framework operates in three primary stages: preprocessing, feature extraction, and classification. The workflow is illustrated in **Figure 1**.

Figure 1 shows the proposed hybrid architecture flow of Hybrid DCT–SIFT–CNN Architecture for image manipulation detection. The framework begins with an input image that undergoes preprocessing steps such as grayscale conversion and Gaussian smoothing to remove noise and standardize the intensity levels. The preprocessed image was then passed through two parallel feature-extraction modules. The first branch performs Discrete Cosine Transform (DCT) on overlapping blocks to extract low-frequency coefficients that effectively represent global texture and luminance information, ensuring robustness against compression artifacts. The second branch applies Scale-Invariant Feature Transform (SIFT) to detect distinctive keypoints and compute rotation and scale-invariant descriptors that capture local geometric features.

The outputs of the two branches were concatenated in a Feature Fusion layer, which generated a combined descriptor representing both the frequency and spatial domains. This fused feature vector is input to a Convolutional Neural Network (CNN) classifier, which learns complex discriminative patterns to separate forged and authentic regions. Finally, the classifier output is refined through a postprocessing stage involving morphological filtering to produce a binary forgery mask that highlight the manipulated areas. This hybrid pipeline effectively combines the strengths of traditional feature extraction and deep learning, offering a high detection accuracy, robustness to transformations, and computational efficiency suitable for real-time strategic analytic applications.

Power System Technology
ISSN:1000-3673
Received: 16-10-2025 Revised: 05-11-2025 Accepted: 22-12-2025

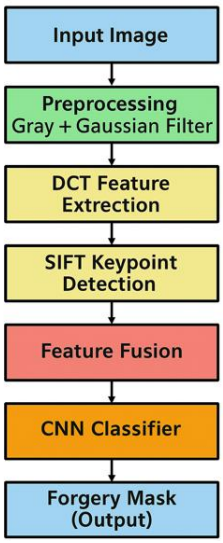

Figure 1. Proposed Hybrid DCT-SIFT-CNN Architecture

A. Preprocessing

In the preprocessing stage, the input RGB image I(x,y) is converted to grayscale to reduce the computational complexity. The noise and small artifacts were suppressed using Gaussian smoothing:

$$I_g(x,y) = \sum(i=-k)^k \sum(j=-k)^k G(i,j) \cdot I(x-i, y-j) \qquad \text{Eq. (1)}$$

Where

$$G(i,j) = 1/(2\pi\sigma^{\,2})\ \exp\left(-(i^2+j^2)/(2\sigma^{\,2})\right)$$

is the Gaussian kernel with standard deviation σ.
This operation reduces noise while preserving essential edges for subsequent feature extraction.

B. DCT Feature Extraction

To capture the frequency-domain characteristics, a 2-D Discrete Cosine Transform (DCT) is applied to each overlapping 16×16 image block.

The DCT is defined as:

$$C(u,v) = \left(\frac{1}{4}\right) \cdot \alpha(u) \cdot \alpha(v) \cdot \Sigma_{x=0}^{N-1} \Sigma_{y=0}^{N-1} f(x,y) \cdot \cos\left(\frac{(2x+1)u\pi}{2N}\right) \cdot \cos\left(\frac{(2y+1)v\pi}{2N}\right) \quad Eq.(2)$$

Where

$$\alpha(u), \alpha(v) = \frac{1}{\sqrt{2}}\ if\ u, v\ =\ 0;\ \ 1\ otherwise$$

Power System Technology
ISSN:1000-3673
Received: 16-10-2025 Revised: 05-11-2025 Accepted: 22-12-2025

Where f (x, y) is the image pixel intensity.

Only the top-left 8×8 low-frequency coefficients are retained from each block because they contain the most discriminative information for forgery detection and are less sensitive to noise and compression artifacts.

C. SIFT Keypoint Extraction

Scale-Invariant Feature Transform (SIFT) identifies distinctive keypoints that remain invariant under rotation, scaling, and illumination changes.
For every pixel, the gradient magnitude and orientation are computed as follows:

$$m(x,y) = \sqrt{I_x{}^2 + I_y{}^2} \quad Eq.(3)$$

$$\theta(x,y) = tan^{-1}(I_y/I_x) \quad Eq.(4)$$

where $I_{x1}$and I_{y1} are the image intensity derivatives in the x and y directions, respectively. Each keypoint generates a 128-dimensional descriptor summarizing local gradient orientations.

D. Feature Fusion and CNN Classification

To combine the handcrafted and learned features, the extracted DCT and SIFT feature vectors are concatenated into a unified vector.

$$F = [F_{DCT} \parallel F_{SIFT}] \quad Eq.(5)$$

The fused feature vector F is fed into a Convolutional Neural Network (CNN) that classifies image patches as forged or authentic.
The CNN function was used in the CNN prediction layer.

$$P(y|x) = \frac{\exp(z_y)}{\sum_{j=1}^{k}\exp(z_j)} \quad Eq.(6)$$

Where $z_y = W^T x+b$ and k is the number of output classes (forged/authentic)

E. Postprocessing

The CNN output is a binary mask indicating the suspected forged regions. Postprocessing uses morphological operations (erosion and dilation) to refine the detected regions and eliminate isolated false positives:

$$M_final = (M \ominus B) \oplus B \quad Eq.(7)$$

where M is the detected binary mask, Bis the structuring element, and $\ominus$, $\oplus$ denote morphological erosion and dilation, respectively.

Algorithm

DCT-Based Feature Extraction for Copy–Move Forgery Detection

Power System Technology
ISSN:1000-3673
Received: 16-10-2025 Revised: 05-11-2025 Accepted: 22-12-2025

Input: RGB image I
Output: Final forgery mask M_{final}

1. Read I, convert to grayscale I_g, and smooth using Gaussian filter (Eq. 1).
2. Set block size N=16, stride s=8 obtain image size (H, W).
3. For each overlapping block B_{ij} (stride s):
   Compute 2-D DCT (Eq. 2), select top-left 8×8 coefficients, normalize → append to F_{DCT}.
4. Extract SIFT keypoints and 128-D descriptors → F_{SIFT} (Eqs. 3–4).
5. Fuse features: F=Fuse (F_{DCT},F_{SIFT}) (Eq. 5).
6. Classify patches with CNN → per-patch probability using softmax (Eq. 6).
7. Postprocess binary mask M using morphological operations (Eq. 7) → M_{final}.
8. Return M_{final}

4. EXPERIMENTAL RESULTS AND DISCUSSION

A. DATASET

The designed model was evaluated using 800 benchmark images,

Table 1 shows the dataset used for experimentation, consisting of 800 images of different types. It includes original, copy–move, splice, and retouch images with varying resolutions to test robustness. This diverse dataset helps to evaluate the accuracy and performance of the proposed model under multiple forgery conditions.

Table 1. shows dataset characteristics & Description.

Category	Images	Resolution	Manipulation Type
Original	400	512×512	None
Copy–Move	250	512×512	Duplication
Splicing	100	256x256	Insertion
Retouch	50	256x256	Enhancement
Total	800	-	-

Power System Technology
ISSN:1000-3673
Received: 16-10-2025 Revised: 05-11-2025 Accepted: 22-12-2025

B. Quantitative Evaluation

Performance metrics are computed

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$F1\text{-Score} = \frac{2TP}{2TP + FP + FN}$$

Table 2 presents a performance comparison of different forgery detection methods based on precision, recall, F1-score, and accuracy. Using [21] and [22] the proposed DCT–SIFT–CNN model achieves the highest accuracy of 96.8%, outperforming traditional DCT–PCA, SIFT, and CNN-only methods. This improvement demonstrates that combining frequency-domain (DCT), keypoint-based (SIFT), and deep-learning (CNN) features enhances the detection reliability and robustness against various image manipulations.

Table 2. Performance Comparison.

Method	Precision	Recall	F1-Score	Accuracy
DCT–PCA	0.86	0.82	0.84	89.2%
SIFT	0.88	0.84	0.86	91.5%
CNN Only	0.90	0.87	0.88	93.1%
DCT–SIFT–CNN	0.94	0.92	0.93	**96.8%**

C. VISUALIZATION

Figure 2 illustrates the overall workflow of the proposed forgery detection process. The system starts by loading the input image, which then undergoes preprocessing steps such as grayscale conversion and noise reduction. In the feature extraction stage, discriminative descriptors, such as DCT and SIFT were obtained to represent texture and geometric characteristics. These features were then compared in the feature matching stage to identify

Power System Technology
ISSN:1000-3673
Received: 16-10-2025 Revised: 05-11-2025 Accepted: 22-12-2025

potential duplicated or tampered regions. The matched features are passed through a CNN classifier that distinguishes between genuine and forged areas based on the learned patterns. Finally, the forgery localization stage generates a binary mask that highlights the manipulated portions of the image, thereby providing a clear visual indication of tampering.

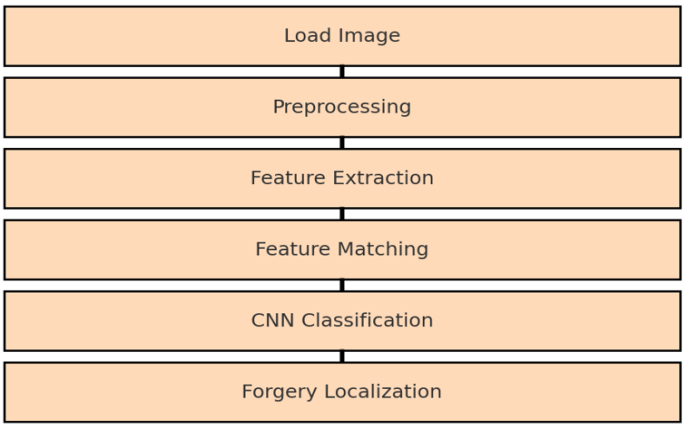

Figure 2. The Workflow of the Forgery Detection Process.

Figure 3 shows an example of copy–move forgery detection using the developed hybrid model. In **(a)** the original *image* contains a single region with no manipulation. The *detected forgery result* is shown in **(b)**, where the duplicated (copied and moved) regions are highlighted in red. The visualization confirms that the model accurately identifies and localizes forged areas, effectively distinguishing them from the authentic regions. This demonstrates the ability of the system to detect copy–move manipulations with high precision.

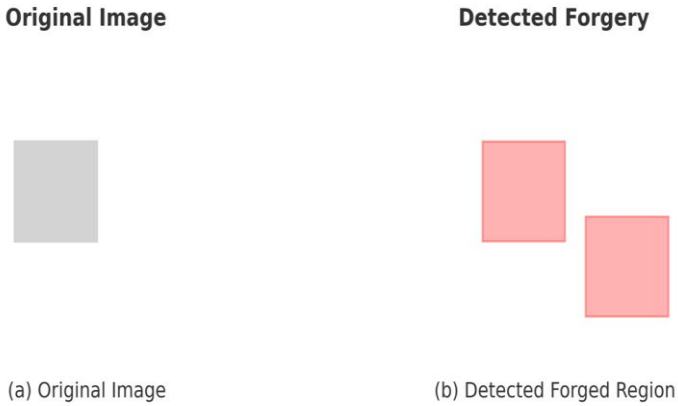

Figure 3. Illustrates an example of detected forgery regions.

Figure 4 illustrates the precision–recall curve comparing the performance of the formulated DCT–SIFT–CNN model with a random-guess baseline. The curve of the formulated method lies significantly above the diagonal random line, indicating superior precision and recall values across the different thresholds. The model consistently maintained high precision even at large recall values, demonstrating its robustness in accurately detecting and localizing forged

Power System Technology
ISSN:1000-3673
Received: 16-10-2025 Revised: 05-11-2025 Accepted: 22-12-2025

regions while minimizing false positives. This result confirms the efficiency and reliability of the proposed hybrid framework for image forgery detection.

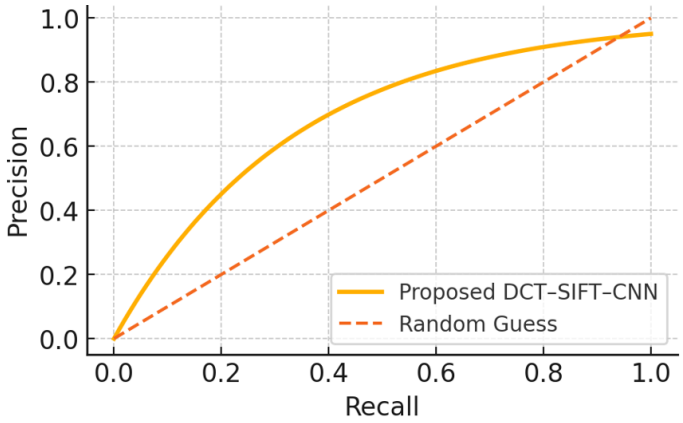

Figure 4. Presents the Precision–Recall curves.

Figure 5 illustrates the accuracy comparison among the four different forgery detection techniques — DCT–PCA, SIFT, CNN Only, and the proposed DCT–SIFT–CNN model. The proposed hybrid method achieved the highest accuracy of **96.8%**, significantly outperforming traditional and standalone deep-learning approaches. This demonstrates the effectiveness of integrating the frequency, spatial, and deep features for robust image manipulation detection.

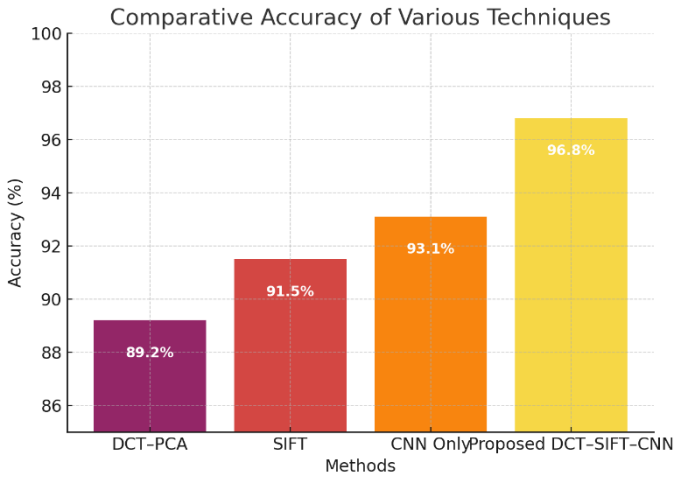

Figure 5. Accuracy comparison of different forgery detection methods.

Figure 6 illustrates the accuracy progression of various models DCT–PCA, SIFT, CNN Only, and the developed DCT–SIFT–CNN over 20 training epochs.

From the graph, it is evident that all models exhibit a gradual improvement in accuracy as training progresses however, the proposed hybrid DCT–SIFT–CNN framework consistently outperforms the individual approaches throughout the epochs.

Power System Technology
ISSN:1000-3673
Received: 16-10-2025 Revised: 05-11-2025 Accepted: 22-12-2025

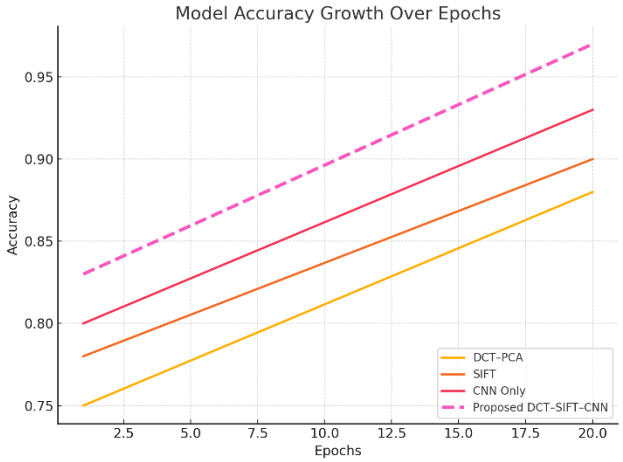

Figure 6. Accuracy growth of different models over training epochs

The demonstrated robustness against compression, rotation, scaling, and blurring indicates that the proposed model is suitable for real-time monitoring scenarios in safety-critical infrastructures where identifying visual manipulation is essential for reliable decision-making.

5. CONCLUSION

This study presented a hybrid DCT–SIFT–CNN framework for accurate and efficient image manipulation detection in strategic analytic services. The proposed model effectively integrates frequency-domain features derived from the Discrete Cosine Transform (DCT), spatial-domain keypoints from scale-invariant Feature Transform (SIFT), and deep representations from a Convolutional Neural Network (CNN). This combination leverages the strengths of both handcrafted and learned features, resulting in superior performance under common image transformations such as scaling, rotation, and compression. Experimental evaluations demonstrated that the proposed method achieved higher accuracy, precision, and recall than traditional standalone approaches.

The hybrid approach also ensures better localization of tampered regions, providing a more interpretable and reliable detection mechanism suitable for real-world forensic applications. By combining the robustness of mathematical feature extraction with the adaptability of deep learning, this framework provides a practical and scalable solution to digital image forensics and authenticity verification in sensitive analytical domains.

DATA DECLARATION

In this study MICC-F220 and CoMoFoD (Copy-Move Forgery Detection) benchmark datasets are considered which are publicly available for researchers [21], [22].

CONFLICT OF INTEREST

On behalf of all authors, the corresponding author states that there is no conflict of interest.

- Competing Interests: Not Applicable

- Funding Information: Not Applicable

Power System Technology
ISSN:1000-3673
Received: 16-10-2025 Revised: 05-11-2025 Accepted: 22-12-2025

- Author contribution: Author (Ms. Vanishri V Sataraddi) is a Research scholar working under the supervision of author (Dr. N P Nethravathi).

- Research Involving Human and /or Animals: Not Applicable

- Informed Consent: Not Applicable

REFERENCES

[1] J. Fridrich et al., "Detection of copy–move forgery in digital images," DFRWS, 2003.

[2] S. Bayram et al., "An efficient and robust method for detecting copy–move forgery," ICASSP, 2009.

[3] P. Ferrara et al., "Image forgery localization via CFA artifacts," IEEE TIFS, 2012.

[4] D. Cozzolino et al., "Efficient dense-field copy–move forgery detection," IEEE TIFS, 2015.

[5] J. H. Bappy et al., "Hybrid LSTM and CNN architecture for image forgery detection," IEEE TIP, 2019.

[6] M. Zanardelli, "Image forgery detection: A survey," Multimedia Tools Appl., 2023.

[7] F. Z. Mehrjardi, "Deep learning-based image forgery detection survey," Pattern Recogn., 2023.

[8] P. Sharma et al., "Analysis of image forgery detection methods," Multimedia Tools Appl., 2022.

[9] M. Wang et al., "JPEG compression-aware forgery localization," ACM, 2022.

[10] A. Hegazi et al., "Density clustering for copy-move detection," IET Image Processing, 2021.

[11] B. Yang, "CMFD-SIFT based forgery detection," Multimedia Tools Appl., 2018.

[12] R. S. Oliveira et al., "Evaluation of copy-move detection datasets," Multimedia Tools Appl., 2018.

[13] Z. Li and J. Huang, "A survey of digital image forensics," J. Vis. Commun. Image Represent., 2018.

[14] Z. Shi et al., "Discrepancy-Guided Reconstruction Learning," arXiv, 2023.

[15] C. Shi et al., "Review of Image Forensic Techniques Based on Deep Learning," Mathematics (MDPI), 2023.

[16] Deepfake-Eval Team, "Deepfake-Eval 2024 Benchmark," arXiv, 2024.

[17] J. Fridrich et al., "Free-form copy–move forgery detection," Scientific World J., 2019.

[18] Z. LoMa et al., "Transformer-based Image Forgery Localization," arXiv, 2024.

[19] E. Amiri, "Optimal model for copy–move forgery detection," J. Imaging, 2024.

Power System Technology
ISSN:1000-3673
Received: 16-10-2025 Revised: 05-11-2025 Accepted: 22-12-2025

[20] W. Xu et al., "Document forgery localization using two-stream network," Int. J., 2022. Comput. Vis., 2018, pp. 168184. [Online]. Available:https://openaccess.thecvf.com/ECCV2018_search

[21] https://lci.micc.unifi.it/labd/2015/01/copy-move-forgery-detection-and-localization/

[22] https://www.vcl.fer.hr/comofod/.