Indian Engineering Journal
ISSN: 2984-9047
Vol 4 Issue 2 (2026)

Semantic-Aware Hierarchical Forgery Network for Image Forgery Detection using Structural and Semantic Inconsistency Learning

P. Muniasamy[1], Dr. B. Srinivasan[2]

[1]Assistant Professor, Department of Information Technology, Sri Ramakrishna Mission Vidyalaya College of Arts and Science (Autonomous and Affiliated to Bharathiar University), Coimbatore, Tamil Nadu, India.

ORCID Id: https://orcid.org/0009-0006-5067-4144, E-mail: profmuniasamy@rmv.ac.in

[2]Associate Professor, Department of Computer Science, Gobi Arts & Science College (Autonomous and Affiliated to Bharathiar University), Gobichettipalayam, Erode, Tamil Nadu, India.

ORCID Id: https://orcid.org/ 0000-0002-5977-083X, E-mail: srinivasanb@gascgobi.ac.in

Abstract:

Image forgery detection is a vital research area because the manipulation of digital images using advanced editing software poses a serious threat to the integrity and credibility of information in various fields, including forensic science, news media, and security. This research proposes an efficient and accurate method for image forgery detection that identifies both structural and semantic inconsistencies introduced during forgery processes, including splicing, copy-move, and image inpainting. The proposed approach presents a Semantic-Aware Hierarchical Forgery Network (SAHF-Net) that combines multi-scale feature learning with dual-branch attention mechanisms to simultaneously analyze local structural inconsistencies and global semantic inconsistencies through multiple convolutional and Transformer layers, enabling effective modeling of inconsistencies across both spatial and semantic contexts. Experimental results on benchmark image forgery datasets demonstrate the effectiveness of the proposed method, achieving 96.3% accuracy in forgery detection. The findings indicate that the model outperforms existing single-stage detection approaches, especially under challenging forgery conditions. In conclusion, the SAHF-Net provides an effective and efficient solution for image forgery detection by jointly capturing structural and semantic inconsistencies, making it highly suitable for forensic applications.

Keywords: Deep Learning, Feature Fusion, Hierarchical Learning Network, Image Forgery Detection, Semantic Inconsistency, Structural Inconsistency

I. Introduction

The rapid growth of social media sharing, news reporting, forensic investigations, and security systems, authenticity verification of digital images has become more critical than ever. With the availability of powerful image editing software and Artificial Intelligence (AI)-based manipulation methods, image forgery has become easier to carry out and harder to detect, posing a great challenge in digital media forensics. Typical image manipulations such as copy-move, splicing and in painting result in both structural inconsistencies, such as edge discontinuities, texture inconsistencies and noise inconsistencies, and semantic inconsistencies, such as implausible object relationships and scene inconsistencies. Current forgery detection approaches mainly use hand-crafted features and shallow Machine Learning (ML) models, which not be sufficiently robust for complex forgery cases. Recent Deep Learning (DL)-based approaches have enhanced detection accuracy by automatically learning discriminative features; however, many of these approaches still only consider either structural inconsistencies or semantic understanding, leading to poor robustness and localization in practical cases. Thus, it is crucial to build a unified model that can jointly model both structural and semantic inconsistencies.

Several recent studies have attempted to improve image forgery detection using advanced DL architectures. A hybrid Vision Transformer with SVM was introduced for robust forgery detection and adversarial resistance [1], while a CNN–Transformer synergy framework improved localization using both spatial and contextual learning [2]. Sequence-to-sequence Transformer models were applied for copy-move forgery detection [3], and language-guided hierarchical frameworks enhanced semantic-level fine-grained localization [4]. Hierarchical UNet models improved progressive forgery localization [5], whereas multi-scale edge-guided self-supervised learning focused on structural inconsistency extraction [6]. Uncertainty-guided refinement methods were developed for image splicing traces [7], and ELA-CNN integration improved authenticity verification using compression artifacts [8]. Siamese learning networks were used for splicing manipulation identification [9], while DL combined with error level and noise residual analysis improved tampering detection [10]. Hybrid optimization-based compressed image forgery detection [11], feature pyramid integration for inpainting detection [12], enhanced Mask R-CNN for blind forgery detection [13], end-to-end attention localization networks [14], and CAMU-Net with coordinate attention for copy-move forgery detection [15] have also shown strong performance. However, most of these approaches lack sufficient generalization, computational efficiency, and joint consideration of low-level structural and high-level semantic inconsistencies. To overcome these challenges, this research presents a Semantic-Aware Hierarchical Forgery Network (SAHF-Net) that uses CNN-based local feature extraction, Transformer-based global semantic encoding, and dual-branch attention to jointly detect multi-scale forgery clues and enhance detection performance and localization stability.

The main contribution and motivation are,

- To build a reliable SAHF-Net that detects both structural and semantic inconsistencies for forgery detection.
- To combine CNN-based local feature learning and Transformer-based global contextualization to model both local artifacts and global patterns of the forgery.
- To propose a two-branch attention model to improve spatial feature refinement and semantic relation modeling for accurately localizing forgery regions.
- To conduct efficient multi-scale feature fusion to enhance the ability to generalize to different types of forgery, such as copy-move, splicing and inpainting.
- To achieve improved forgery detection performance with 96.3% accuracy, which is more robust and efficient than current state-of-the-art techniques.

Organization: Section II deals with related work. Section III details the proposed SAHF-Net method, mathematical formulation and algorithmic steps, Section IV the experimental results and performance evaluation, and Section V the conclusion of the research and directions for future research.

II. Background Study

M. Abdelmaksoud et al. (2025) [1] proposed a hybrid Vision Transformer + SVM for adversarially robust image forgery detection to extract deep features and perform classification using a Support Vector Machine. This method increases the detection accuracy and robustness of the system, but it does not provide insightful localization and cannot deal with complicated semantic inconsistencies, although it attains better detection measures.

S. Sharma et al. (2025) [2] introduced a hybrid CNN-Transformer model to enhance the forgery localization, used the spatial and global features. While the model improves the localization accuracy, it is more complex and has limited generalization across different forgery types, although it achieves better detection performance.

G. Hao et al. (2025) [3] developed a copy-move forgery detection and localization method using sequence-to-sequence Transformer model and put stress on the sequential learning of features. It is effective for copy-move forgery detection, but it is not effective in other forgeries and it does not represent larger semantic anomalies.

X. Guo et al. (2024) [4] proposed a hierarchical language based model with vision and semantic processing to detect and localize forgery in fine-grained forgery detection and localization. It is useful in detecting higher semantic anomalies but it is computationally expensive and not efficient for real time operations even if it is very accurate.

Q. Li et al. (2024) [5] developed a progressive hierarchical UNet for multi-level feature extraction to detect and locate forgery. This method improves the structural detection capability, but it does not capture semantic information, and has poor performance for complex forgery, even though it performs well on segmentation.

L. Zhang et al. (2024) [6] proposed a multi-scale edge-guided method with self-supervised and adversarial learning for forgery detection. It effectively captures structural clues at the edge level but it lacks semantic clues, and low performance with complex forgeries.

H. Li et al. (2024) [7] proposed uncertainty-based refinement for image splicing detection, especially in scientific papers. It improves detection accuracy but it is not generalized for various types of forgery in real-world scenarios.

A. Mohamed and R. Hassan (2024) [8] proposed an integrated error level analysis and DL (ELA-CNN) method. It enhances detection of compression artifacts, but it's sensitive to quality variations and difficult to handle semantic changes.

A. E. Ramirez-Rodriguez et al. (2024) [9] proposed AISMSNet, which uses Siamese networks to learn image splicing using similarity. This improves detection of duplicates, but only works for image splicing forgery and fails for other manipulations.

Y. Jiang et al. (2024) [10] developed a combination of DL, error level analysis and noise residual for detection. This improves accuracy but uses limited feature representations and reduced accuracy for compressed or heavily manipulated images.

Table I. Background Study Image Forgery Detection

Author (Year)	**Concept**	**Methods Used**	**Research Gap**	**Limitations**	**Results**
A. Bhowal et al. (2024) [11]	Forgery detection in compressed images	Hybrid DL with optimization techniques	Lack of semantic feature integration	Sensitive to compression variations	Improved detection accuracy for compressed images
S. Aly et al. (2025) [12]	Inpainting forgery detection	Feature Pyramid Network (FPN) with deep learning	Limited to inpainting type forgery	Poor generalization to other forgery types	High localization accuracy in inpainting regions
V. S. Tallapragada et al. (2024) [13]	Blind forgery detection	Enhanced Mask-RCNN architecture	Lack of semantic-level analysis	High computational cost	Effective region-based forgery detection
Y. He et al. (2025) [14]	Forgery localization	End-to-end attention network	Insufficient multi-scale feature fusion	Limited robustness to complex manipulations	Improved localization precision
K. Zhao et al. (2024) [15]	Copy-move forgery detection	CAMU-Net with coordinate attention and multi-scale fusion	Focus only on copy-move forgery	Not suitable for general forgery detection	High accuracy in copy-move detection

Table I offers a background study comparison of the currently available methods for detecting image forgery in terms of concepts, techniques and performances. It also outlines the shortcomings and future work, suggesting a need for an integrated approach to learning structural and semantic inconsistencies in a joint fashion, for successful and general forgery detection.

Indian Engineering Journal
ISSN: 2984-9047
Vol 4 Issue 2 (2026)

III. Proposed Methodology

This section describes the proposed SAHF-Net approach for image forgery detection, including the dataset used for training and testing, and the proposed solution. This includes the diagram of the SAHF-Net model, explanation of the structural and semantic learning process, dual attention mechanism, mathematical formula and the algorithm in pseudocode form for the proposed forgery detection model.

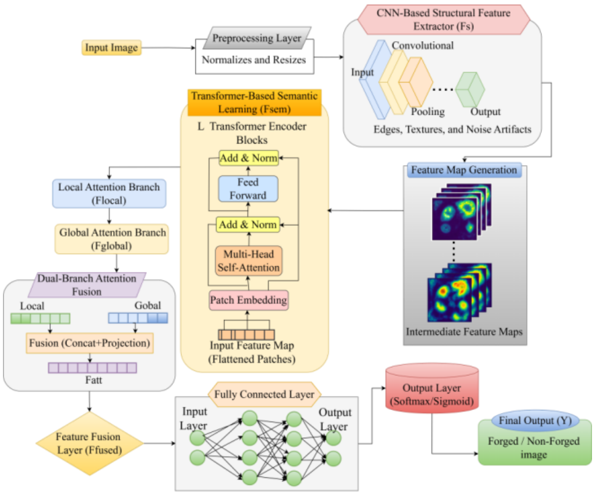

Fig. 1. Semantic-Aware Hierarchical Forgery Network Model for Image Forgery

Fig 1 represents the proposed SAHF-Net architecture, which combines CNN-based structural feature extraction, Transformer-based semantic learning, and dual-attention mechanisms to capture both local and global anomalies in images. The combined features are fed to fully connected and output layers for the final decision to detect image forgery.

A. Dataset Description

The Image Forgery Detection Dataset https://www.kaggle.com/datasets/labid93/image-forgery-detection\ includes a set of original and tampered images for training and testing models that detect image forgery. It covers different types of manipulations such as copy-move and splicing, enabling models to detect both structural and visual inconsistencies in manipulated images. It is widely used for training and evaluating deep learning-based forgery detection models, with a wide variety of images to enhance the accuracy and resilience of the image forgery detection system. It is commonly used in studies to evaluate algorithms for detecting image forgery and improving digital image forensics technologies.

B. Semantic-Aware Hierarchical Forgery Network

The proposed SAHF-Net is designed for accurate image forgery detection that combines structural and semantic inconsistency analysis in a single DL-based model. The model uses CNN-based layers to extract structural inconsistencies (edges, textures, noise patterns) and Transformer-based attention to analyse high-level semantic inconsistencies in the global image context. The method uses two-branch attention, with one branch dedicated to feature refinement and the other to capturing contextual information, and a feature fusion approach to merge multi-scale features seamlessly. SAHF-Net operates hierarchically, from pixel-level to semantic-level feature processing, to precisely locate the tampered areas. The proposed approach combines multi-scale feature extraction, attention, CNN–Transformer hybrid models and feature fusion. The model is designed to be efficient for real-time applications through lightweight backbones and parallel processing, thereby achieving both speed and accuracy. The algorithm is reproducible since it uses standard datasets, set training parameters and publicly accessible models, enabling consistent performance under different settings.

$I \in \mathbb{R}^{H \times W \times C}$ (1)

In equation 1 represents that I is an input image, H denotes a height of the image, W is a width of image, C Number of channels (RGB = 3). Specifies the input image tensor for hierarchical features.

$F_0 = Conv(I; \theta_0)$ (2)

In equation 2 represents that F_0 is an initial feature map, Conv is a convolution function, θ_0 represents the learnable parameters of the convolution layer. The feature extractor takes low level features such as edges and textures.

$F_h = \bigcup_{i=1}^{n} F_i$ (3)

In equation 3 represents that F_h is a hierarchical feature set, F_i is a feature map with level i, n is a number of layers, $\cup$ is a union of features. Aggregates features from different layers into hierarchical features.

$F_{ms} = \sum_{i=1}^{n} w_i F_i$ (4)

In equation 4 represents that F_{ms} is a feature map at multiple scales, w_i is a weight for each feature level. Weights multi-scale features with normalized weights for representation.

C. Structural and Semantic Feature Extraction

The proposed SAHF-Net uses CNN to learn structural features such as edges, textures and noise inconsistencies through local spatial information learned via convolution and pooling. This effectively highlights inconsistencies at the pixel level, such as splicing and copy-move. These features are then passed to a Transformer module, which captures long-range contextual features, and introduces global semantic inconsistencies via self-attention. This enables the network to take into account low and high level information to improve detection and localization.

$F_s = ReLU(W_s * F_{ms} + b_s)$ (5)

In equation 5 represents that F_s is a structural features, W_s denotes a convolution weight, b_s indicates that bias, ReLU is an activation function, $*$ denotes a convolution. Capture inconsistencies such as edges, textures, and noise.

$F_p \in \mathbb{R}^{N \times d}$ (6)

In equation 6 represents that F_p is a patch feature sequence, N is a number of patches, d is a feature dimension. Convert feature maps to the sequence for the transformer.

$Q = F_p W_Q,\ K = F_p W_K,\ V = F_p W_V$ (7)

In equation 7 represents that Q is a query matrix, K is a key matrix, V is a value matrix and W_Q, W_K, W_V represent projection weights. Create query, key and value matrices for attention.

$F_{sem} = A + F_p$ (8)

In equation 8 represents that F_{sem} is a semantic features, A is an attention output. The attention output is combined with the input features to obtain semantic features.

D. Attention Mechanism and Classification

The proposed SAHF-Net combines two-branch attention to narrow the features by jointly extracting local spatial significance and global contextual correlations. The local attention branch deals with pixel-level areas to capture finer structural aspects whereas the global attention branch captures long-range connectivity to comprehend semantic inconsistencies in the image. The two complementary representations are then fused by feature fusion module, which fuses structural and semantic features into a unified representation. Lastly, the fused features are processed by a fully connected classification layer, which makes the final prediction, whether the image is forged or non-forged.

Indian Engineering Journal
ISSN: 2984-9047
Vol 4 Issue 2 (2026)

$A = softmax\left(\frac{QK^T}{\sqrt{d}}\right)V$ (9)

In equation 9, d represents the scaling factor. Learn long-range dependencies through attention between all regions to semantic comprehension.

$F_{local} = \sigma\left(Conv(F_s)\right)$ (10)

In equation 10 means that F_{local} represents local attention features, Conv is a convolution operation, and σ is an activation function. Strengthen significant local structural areas of the image.

$F_{fused} = W_f[F_{local} + F_{sem}] + b_f$ (11)

In equation 11 means that F_{fused} is a combined feature, W_f is fusion weights, + denotes concatenation, b_f is bias features. Integrate both structural and semantic properties to form a single representation to provide strong forgery detection.

$Y = softmax\left(W_c F_{fused} + b_c\right)$ (12)

In equation 12, Y is an output probability, W_c is a classification weights, b_c is a bias. Produce final prediction probabilities in order to classify the image as forged or non-forged.

Algorithm: SAHF-Net for Image Forgery Detection
Input: I → Input image θc → CNN parameters (structural feature extractor) θt → Transformer parameters (semantic feature extractor) θl → Local attention parameters θg → Global attention parameters θf → Fusion parameters θcls → Classification layer parameters Begin Structural Feature Extraction Fs ← CNN(I, θc) Semantic Feature Learning Fsem ← Transformer(Fs, θt) Dual-Branch Attention Flocal ← Local_Attention(Fs, θl) Fglobal ← Global_Attention(Fsem, θg) Feature Fusion Ffused ← Concatenate(Flocal, Fglobal) Ffused ← FullyConnected(Ffused, θf) Final Classification Yprob ← Softmax(Ffused, θcls) Decision Making

If Yprob ≥ Threshold then
  Y ← "Forged"
Else
  Y ← "Non-Forged"
End If
Return Y
End
Output:
  Y → Predicted label (Forged / Non-Forged)

In the proposed SAHF-Net algorithm, an input image is first transformed into structural features with CNN layers and then semantic features are learned with the help of a Transformer module to gain a global context. It then uses dual-branch attention to refine local and global features, which are fused together to make a single representation. Lastly, the fused features undergo classification layer to determine whether the image is forged or not.

IV. Result and Discussion

This section presents the results and discussion of the proposed SAHF-Net for image forgery detection in terms of performance metrics. The model was implemented and evaluated using Python, where the results are presented in tabular form and compared with other forgery detection techniques using graphical analysis. The results show that the SAHF-Net method achieves better detection accuracy and overall performance by jointly analyzing structural and semantic inconsistencies.

Input Image	**Probability Map**	**Structural Inconsistency Map**	**Semantic Inconsistency Map**	**Final Forgery Mask**	**Output Image**

Fig. 2. Proposed Image Forgery Detection using SAHF-Net with Structural and Semantic Inconsistencies

Fig 2 illustrates the steps of the SAHF-Net method, from the input image, to probability map, structural and semantic inconsistencies, and finally the mask. The last image shows visual cues for the forgeries, which demonstrates the effectiveness of the structural and semantic information in the detection of image forgery.

Table II. Evaluation Metrics of Image Forgery Detection

Method	**Algorithm**	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**AUC**
Existing Methods	(ViT) [1]	91.2	90.5	89.8	90.1	92.3
	U-Net [5]	89.6	88.9	87.5	88.2	90.4
	Hybrid ELA-CNN [8]	92.8	92.1	91.6	91.8	93.5
	FPN [12]	93.5	92.9	92.2	92.5	94.1
	Mask R-CNN [13]	94.2	93.7	93.1	93.4	95.0
	CAMU-Net [15]	95.1	94.6	94.0	94.3	95.8
Proposed Methods	SAHF-Net	96.3	95.8	95.2	95.5	96.7

Table II shows the comparison of the proposed SAHF-Net model with other state-of-the-art image forgery detection methods in terms of accuracy, precision, recall, F1-score and area under the curve (AUC). The results indicate that the proposed SAHF-Net has better performance, with the ability to accurately detect and classify image forgery.

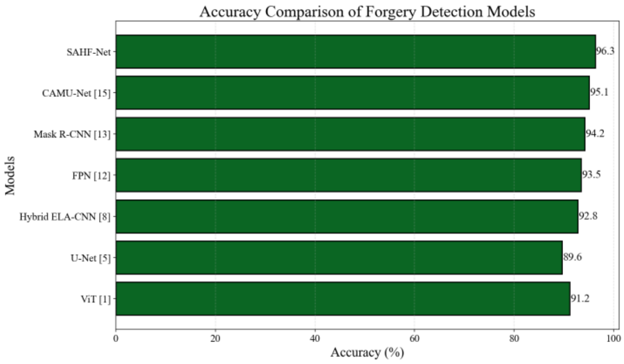

Fig. 3. Accuracy Comparison

Fig 3 compares the accuracy of the proposed SAHF-Net with the recent techniques. The SAHF-Net has the highest accuracy of 96.3% as compared to other baseline models such as ViT, U-Net and CAMU-Net. This improvement is due to its ability to learn both structural and semantic inconsistencies.

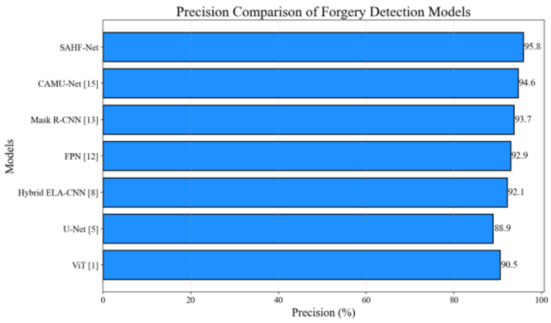

Fig. 4. Comparison of Precision of Forgery Detection Models

Fig 4 compares the precision of forgery detection models with the proposed SAHF-Net. The proposed SAHF-Net outperforms the other approaches with the highest precision of 95.8% by better prediction of positives. This performance improvement is achieved through the dual-branch attention mechanism, which enables fine feature discrimination.

Fig. 5. Recall of Forgery Detection Models

Fig 5 shows recall of various deep learning-based forgery detection approaches. The recall performance of proposed SAHF-Net is the highest (95.2%) which shows the ability to detect all tampered regions. This is due to its multi-scale feature learning capability, which effectively detects fine-grained tampering regions.

Indian Engineering Journal
ISSN: 2984-9047
Vol 4 Issue 2 (2026)

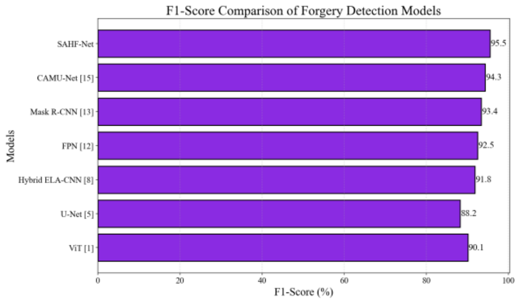

Fig. 6. F1-Score of Forgery Detection Models

Fig 6 compare the F1-score of all models and SAHF-Net has the highest F1 of 95.5%. It provides a good balance between precision and recall, and outperforms CNN and transformers. The hierarchical fusion approach results in stable classification.

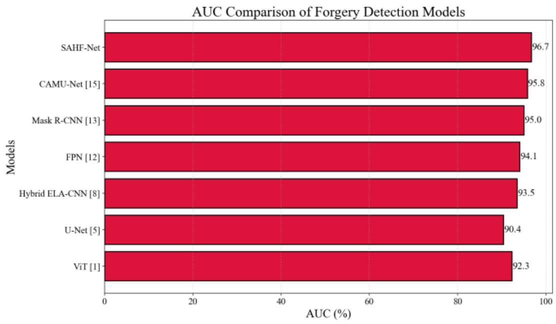

Fig. 7. AUC Comparison of Forgery Detection

Fig 7 compares the AUC results from different approaches, with SAHF-Net achieving the best result of 96.7%. The high AUC value demonstrates strong discriminative capability between forged and genuine images. This is due to the joint learning of structural and semantic inconsistency.

Table III. Ablation Study

Model Variant	**Accuracy (%)**
Baseline CNN	89.4
CNN + Transformer	92.1
CNN + Transformer + Attention	94.2
CNN + Transformer + Fusion	93.6
Proposed SAHF-Net (Full Model)	96.3

Table III presents the ablation study and demonstrates that each component of the proposed SAHF-Net positively contributes to the overall performance, from the baseline CNN (89.4%) to the CNN with Transformer-based semantic learning (92.1%). The use of dual-branch attention further enhances features and leads to 94.2% accuracy, while multi-scale fusion helps with forgery generalization. The full SAHF-Net gives the best accuracy of 96.3%, validating the need for both structural and semantic learning, attention and fusion.

V. Conclusion

This research proposed a SAHF-Net for reliable image forgery detection by jointly examining both structural and semantic anomalies in forged images. The proposed approach combines CNN-based feature learning to detect low-level inconsistencies such as edges, textures, and noise patterns with Transformer-based semantic extraction to capture high-level contextual inconsistencies. The dual-branch attention mechanism further enhances local feature refinement and global dependency learning, and the multi-scale fusion mechanism enables robustness to various types of image forgery, such as copy-move, splicing and inpainting. Evaluation on benchmark forgery datasets showed that the proposed SAHF-Net achieved high accuracy (96.3%) and performed better than several state-of-the-art methods in terms of precision, recall, F1-score and AUC. These results validate the proposed model as an effective and accurate solution for digital image forensics. Future research will focus on model compression for real-time inference and generalization across datasets. Moreover, explainable AI techniques can be incorporated to improve interpretability and increase confidence in forensic results.

Reference

[1] M. Abdelmaksoud, B. Youssef, K. Wassif, and R. A. El-Khoribi, “Hybrid framework for image forgery detection and robustness against adversarial attacks using vision transformer and SVM,” Scientific Reports, vol. 15, p. 40371, 2025. doi: https://doi.org/10.1038/s41598-025-25436-z

[2] S. Sharma, B. K. Singh, and H. Garg, “Robust image forgery localization using hybrid CNN-transformer synergy based framework,” Computers, Materials & Continua, vol. 82, no. 3, pp. 4691–4708, 2025. doi: https://doi.org/10.32604/cmc.2025.061252

[3] G. Hao, P. Liang, Z. Li, H. Zhao, and H. Zhang, “Image copy-move forgery detection and localization method based on sequence-to-sequence transformer structure,” Computers, Materials & Continua, vol. 82, no. 3, pp. 5221–5238, 2025. doi: https://doi.org/10.32604/cmc.2025.055739

[4] X. Guo, X. Liu, I. Masi, and X. Liu, “Language-guided hierarchical fine-grained image forgery detection and localization,” International Journal of Computer Vision, vol. 133, pp. 1015–1037, 2024. doi: https://doi.org/10.1007/s11263-024-02255-9

[5] Q. Li, C. Wang, X. Zhou, and Z. Qin, “Hierarchical progressive image forgery detection and localization method based on UNet,” Big Data and Cognitive Computing, vol. 8, no. 9, p. 119, 2024. doi: https://doi.org/10.3390/bdcc8090119

[6] L. Zhang, T. Zhou, and W. Chen, “Multi-scale edge-guided image forgery detection via improved self-supervision and self-adversarial training,” Electronics, vol. 14, no. 9, p. 1877, 2024. doi: https://doi.org/10.3390/electronics14091877

[7] H. Li, J. Liu, and K. Zhang, “Exposing image splicing traces in scientific publications via uncertainty-guided refinement,” Patterns, vol. 5, no. 9, p. 101038, 2024. doi: https://doi.org/10.1016/j.patter.2024.101038

[8] A. Mohamed and R. Hassan, “Detecting image manipulation with ELA-CNN integration: A powerful framework for authenticity verification,” PeerJ Computer Science, vol. 10, p. e2176, 2024. doi: https://doi.org/10.7717/peerj-cs.2176

[9] A. E. Ramirez-Rodriguez, R. E. Arevalo-Ancona, H. Perez-Meana, M. Cedillo-Hernandez, and M. Nakano-Miyatake, “AISMSNet: Advanced image splicing manipulation identification based on Siamese networks,” Applied Sciences, vol. 14, no. 13, p. 5545, 2024. doi: https://doi.org/10.3390/app14135545

[10] Y. Jiang, F. Wang, and S. Liu, “Detection of image tampering using deep learning, error levels and noise residuals,” Neural Processing Letters, vol. 56, p. 94, 2024. doi: https://doi.org/10.1007/s11063-024-11448-9

[11] A. Bhowal, S. Neogy, and R. Naskar, “Deep learning-based forgery detection and localization for compressed images using a hybrid optimization model,” Multimedia Systems, vol. 30, no. 3, p. 128, 2024. doi: https://doi.org/10.1007/s00530-024-01314-6

[12] S. Aly, O. Emara, H. Mahmoud, and S. Bekhet, “Detecting image forgery: A deep learning framework with feature pyramid integration for inpainting detection,” Neural Computing and Applications, vol. 37, no. 24, pp. 20365–20381, 2025. doi: https://doi.org/10.1007/s00521-025-11461-6

[13] V. S. Tallapragada, D. V. Reddy, and G. P. Kumar, “Blind forgery detection using enhanced mask-region convolutional neural network,” Multimedia Tools and Applications, vol. 83, pp. 1–24, 2024. doi: https://doi.org/10.1007/s11042-024-18590-9

[14] Y. He, L. Yang, and Q. Zhang, “Learning to localize image forgery using end-to-end attention network,” Neurocomputing, vol. 576, p. 127336, 2025. doi: https://doi.org/10.1016/j.neucom.2024.127336

[15] K. Zhao, X. Liu, J. Wang, and G. Chen, “CAMU-Net: Copy-move forgery detection utilizing coordinate attention and multi-scale feature fusion-based up-sampling,” Expert Systems with Applications, vol. 238, p. 121918, 2024. doi: https://doi.org/10.1016/j.eswa.2023.121918