IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 14, No. 6, December 2025, pp. 5279~5289
ISSN: 2252-8938, DOI: 10.11591/ijai.v14.i6.pp5279-5289 ❒ 5279

Improved copy-move forgery detection through multilevel clustering

Doaa Gamal Abdelazem[1], Hala H. Zayed[1,2], Ahmed Taha[1]
[1]Department of Computer Science, Faculty of Computers and Artificial Intelligence, Benha University, Benha, Egypt
[2]Faculty of Engineering, Egypt University of Informatics, Cairo, Egypt

Article Info

Article history:

Received Oct 15, 2024
Revised Sep 29, 2025
Accepted Oct 16, 2025

Keywords:

AKAZE
Copy-move forgery
Double matching
Image forensics
Machine learning
Multilevel DBSCAN
Scale-invariant feature transform

ABSTRACT

Copy move forgery detection (CMFD) based on keypoints remains a widely used technique; however, it often struggles to effectively identify small and smoothly tampered regions within images. This paper introduces a CMFD method that enhances detection accuracy by integrating a double-matching process with advanced region localization techniques. Delaunay triangles formed by accelerated KAZE (AKAZE) and scale-invariant feature transform (SIFT) features are matched in the double-matching process to identify suspicious regions. To ensure sufficient keypoint pairs, the set of matching triangles is iteratively expanded to include neighboring triangles, covering the entire tampered area. Subsequently, a second matching with a looser threshold is performed on the vertices. In the region localization process, the multilevel density-based spatial clustering of applications with noise (DBSCAN) effectively handles scenarios involving multiple copied regions with varying sizes. Using the standard MICC-F600 and COVERAGE datasets, experiments demonstrate that the proposed CMFD method is robust and achieves better performance than state-of-the-art baselines.

Corresponding Author:

Doaa Gamal Abdelazem
Department of Computer Science, Faculty of Computers and Artificial Intelligence, Benha University
Benha, Egypt
Email: d.doaa27150@fci.bu.edu.eg

1. INTRODUCTION

Digital images now permeate everyday life and critical domains, including security and identity workflows, surveillance systems, intelligence operations, forensic investigations, political campaigns, and medical imaging [1]–[4]. Sometimes, images are falsified by copying a region from one photo and placing it into another—this is known as image manipulation or image forgery. It is difficult to identify these fake photos without specialized tools [5]–[7]. Researchers have developed various methods to detect such fake images.

Image-forgery detection techniques are commonly categorized into block-based and keypoint-based approaches. In block-based methods, the image is partitioned into rectangular tiles (overlapping or non-overlapping), and features are extracted from each tile. In contrast, keypoint-based methods detect salient points across the image and describe their local neighborhoods for matching.

Fridrich *et al.* [8] were the first to present a block-based method for detecting duplicated regions. The discrete cosine transform (DCT) algorithm was used to extract features from each block, and then these features were compared to identify matching vectors. The algorithm resisted specific attacks such as brightness and Gaussian noise. It failed against geometric attacks, such as scaling. He *et al.* [9] introduced another copy move forgery detection (CMFD) algorithm using an overlapping blocks method and improved

DCT, which is sensitive to smoothing, Gaussian noise, and quality factors. Popescu and Farid [10] reduce the dimensionality of block descriptors by applying principal component analysis (PCA) to every image block. The descriptors are then lexicographically ordered so that similar blocks are grouped, and a subsequent matching step exposes duplicated regions. The approach remains reliable after common post-processing, including lossy compression and additive noise, and remains robust. Luo *et al*. [11] break each block into four smaller sub-blocks and form the feature vector from the mean values of the R, G, and B channels. Their method shows robustness to Gaussian blur, additive noise, and JPEG compression. Hayat and Qazi [12] introduced a block-based CMFD where the tampered image is first transformed by discrete wavelet transform (DWT) and then partitioned into overlapping blocks of equal size, and each block is processed with the 2-D DCT (to reduce coefficients and retain the most informative ones). The blocks' DCT coefficient vectors are lexicographically sorted and matched via correlation to localize tampered regions. The method surpassed prior state-of-the-art accuracy but breaks down with occlusion and repeated patterns. Alkawaz *et al*. [13] similarly apply the 2-D DCT to overlapping blocks and reorder them lexicographically, but match using Euclidean distance between DCT coefficients (enabling faster matching); however, an inappropriate block size reduces CMFD accuracy.

Silva *et al*. [14] introduced a keypoint-based CMFD that employs speed up robust feature (SURF) descriptors with a clustering stage: after preprocessing, SURF keypoints/descriptors are computed, matches are formed via Euclidean distance, and spurious correspondences are pruned by clustering. The method is fast and robust to scale and rotation, but its performance degrades under JPEG compression. Pandey *et al*. [15] proposed a hybrid CMFD that fuses scale-invariant feature transform (SIFT), SURF, and HOG features, yielding higher accuracy at the cost of increased computation. Yu *et al*. [16] use Hu histogram (HH) together with the multi-support region order-based gradient histogram (MROGH) descriptor, outperforming SIFT- and SURF-based baselines when the tampered region undergoes rotation or JPEG compression; however, performance drops with scale changes and additive-noise post-processing.

Comparative studies have shown that keypoint-based techniques outperform block-based methods in handling geometric transformations. However, detecting small, smooth, duplicated regions with varying shapes and sizes remains challenging due to the insufficient number of keypoints. This research proposes an approach that combines accelerated KAZE (AKAZE) and SIFT feature extraction to address this limitation. A subsequent double-matching strategy is employed to identify multiple duplicated regions, followed by multilevel density-based spatial clustering of applications with noise (DBSCAN) clustering to accurately localize objects with diverse shapes and scales. The proposed method effectively detects tampered regions across various objects, including small and smooth areas with diverse shapes and sizes, and can handle post-processing (e.g., scale changes, rotation, blurring, and additive noise). Experimental evaluations demonstrate robustness and high accuracy in detecting/localizing forged regions.

2. THE PROPOSED DETECTION METHOD

2.1. Framework

The framework has two main stages: keypoint matching and localization. Its overall architecture is shown in Figure 1. The first stage targets small, smooth modifications of varying shapes and sizes by detecting and describing local visual features. To achieve this, keypoint extraction is performed using SIFT and AKAZE.

SIFT is robust under geometric changes, including noise, illumination variations, scaling, and rotation. However, they can be sensitive to noise, especially in low-light conditions. For this reason, another keypoint, AKAZE, is combined with SIFT to obtain a more robust descriptor. AKAZE computes M-LDB descriptors in a nonlinear diffusion scale space that encodes gradient information. The resulting features are invariant to moderate scale changes, rotation, and affine transformations, and are more distinctive across scales due to the nonlinear scale-space construction. While robust in various scenarios, AKAZE can be sensitive to heavy blurring and noise, which may affect accurate keypoint detection and matching. SIFT is still used to improve matching success rates across various conditions. This fusion between AKAZE and SIFT leverages the distinct characteristics of each. After obtaining keypoints, they are matched separately using double matching to obtain matched keypoint pairs. Many matched keypoint pairs may cluster in a specific area due to multiple descriptors, thereby improving matching performance. To reduce computational complexity in the next stage, extracted features are combined, clustered, and filtered by describing each cluster's features in a new model and discarding redundant cluster features.

Localization approaches are applied to the clustered features to locate copy-move forgeries. The multilevel DBSCAN algorithm is employed to refine tampering localization precision, eliminating false matches and identifying replicated regions across multiple scales and orientations. In the final stage, the tampered region is precisely pinpointed through a meticulous comparison of the original image and its post-affine transformation.

Figure 1. The proposed work

2.2. SIFT and AKAZE feature extraction

Using AKAZE and SIFT, the input image yields the required feature descriptors. However, keypoint-based CMFD approaches often fail in regions with duplicated features, such as small or smooth areas. With default settings, conventional keypoint detectors often fail to detect sufficient keypoints in these regions, leading to missed detections. To overcome this limitation, and inspired by the approach in [17], are deliberately set to lower values to detect more relevant interest points. Detector response thresholds, denoted as T_{A-KAZE} and T_{SIFT}, deliberately set to lower values in order to detect a higher number of relevant interest points.

Figure 2 provides a comparative overview of keypoint detection performance across multiple algorithms applied to a tampered horse image. Figure 2(a) presents the tampered image used as the evaluation input. The subsequent subfigures display the keypoints detected by different feature detectors: local intensity order pattern (LIOP) in Figure 2(b), SURF in Figure 2(c), KAZE in Figure 2(d), SIFT in Figure 2(e), and AKAZE in Figure 2(f). This comparison highlights the variation in detection patterns and distribution of keypoints across the methods.

These findings suggest that conventional extractors, LIOP [18], SURF [19], KAZE [20], SIFT [21], and AKAZE [22], perform well for general keypoint detection. However, KAZE and AKAZE demonstrate superior capability in accurately capturing the whole structure of objects, with AKAZE offering the advantage of faster computation. Conversely, SIFT tends to respond to broader image regions, which can help localize manipulated areas. To leverage the strengths of both approaches, we propose a hybrid strategy that combines AKAZE and SIFT using customized low-threshold settings. This ensures a dense, meaningful keypoint distribution across tampered zones, thereby enhancing the robustness of CMFD pipelines.

(a) (b) (c) (d) (e) (f)

Figure 2. Feature extraction with various keypoint detector: (a) tampered horses, (b) LIOP, (c) SURF, (d) KAZE, (e) SIFT, and (f) AKAZE

2.3. Feature matching

After extracting feature descriptors with AKAZE and SIFT, a double-matching procedure [23] is applied to the descriptors to identify intra-image correspondences. This yields a set of keypoints $\{x_1, x_2, \dots, x_n\}$ and their corresponding descriptors $\{F_1, F_2, \dots, F_n\}$. Double matching assesses pairwise descriptor similarity to retain matching features. The similarity vector integrates triangle-based and keypoint-based matching. In the first stage, a Delaunay triangulation is constructed over the SIFT and AKAZE keypoints, and triangles are matched using g2NN, thereby filtering out isolated pairs of keypoint pairs and delineating candidate suspicious regions. In the second stage, each matched-triangle group is expanded by iteratively appending adjacent triangles. The vertices of the expanded groups define a new keypoint subset, which is re-matched with g2NN under a relaxed threshold. This yields a dense set of correspondences near and along region boundaries. The final correspondence set is obtained by fusing the AKAZE and SIFT-based matching results.

Figure 3 shows the progression of the image processing steps. Figure 3(a) shows the original input image, while Figure 3(b) shows the Delaunay triangulation applied to the extracted keypoints. Figure 3(c) presents the matched triangle pairs obtained via g2NN matching. Figure 3(d) demonstrates the expansion of the triangle set through the inclusion of adjacent triangles. Figure 3(e) illustrates the refinement process, where triangle vertices are transformed into enhanced keypoints. Finally, Figure 3(f) displays the output of the detection stage, which results from the integration of SIFT and AKAZE keypoint matches.


(a) (b) (c) (d) (e) (f)

Figure 3. The progression of the image processing steps: (a) the original image, (b) the outcome of delaunay triangulation, (c) the matched triangles, (d) the set of expanding triangles, (e) the transformation of the triangles into keypoints, and (f) the final result

2.4. Forged region localization
2.4.1. Keypoint pairs clustering

In the current context, although pairs of keypoints have been identified, the lack of a specific direction poses a challenge. This implies that some pairs might lack consistent directional information. Consequently, when computing the transfer matrix, pairs with reversed directions might not meet the criteria and could be classified as outliers. To address this issue and retain pairs with opposite directions, a strategy reverses these pairs during processing. Let a pair be $\{(x_1, y_1), (x_2, y_2)\}$, where $(x_1, y_1), (x_2, y_2)$ are the point coordinates. To obtain the pair's angle, first compute the difference vector (d_x, d_y) using:

$$(d_x, d_y) = (x_1 - x_2, y_1 - y_2) \quad (1)$$

Define the pair's angle as the angle formed by this vector and the X-axis:

$$a_p = \langle (d_x, d_y), (C, 0) \rangle \quad (2)$$

With a positive constant, and $\langle\rangle$ indicating the angle formed by two vectors.

To enforce a consistent direction, an angle constraint is applied: if a pair's angle is negative, the pair is reversed. For example, if the angle of $\{(x_1, y_1), (x_2, y_2)\}$ is negative, rewrite it as $\{(x_2, y_2), (x_1, y_1)\}$. This unifies the direction of neighboring pairs and ensures all angles are ≥ 0. For images with multiple copy–move areas, clustering is used to assign key point pairs to labels, thereby identifying their region membership and the corresponding inter-region transformations. Before clustering, each pair is represented by a feature vector comprising its endpoints, difference vector, and angle:

$$f_p = (x_1, y_1, x_2, y_2, d_x, d_y, a_p) \quad (3)$$

Where a_p is the pair angle defined above.

DBSCAN clustering is used to group similar pairs and features. DBSCAN was chosen because it automatically determines the number of groups and works well with densely packed data points. The algorithm identifies core points (pairs with many neighbors) and expands clusters around them. Points without enough neighbors are considered outliers. Pairs within the same cluster belong to copied-and-pasted areas. Each cluster is then analyzed to locate the copied region. The clustering technique requires refinement to effectively detect objects with diverse shapes and sizes. This study proposes a multilevel clustering approach to improve the identification of small, smooth regions.

2.4.2. Multilevel DBSCAN

The multilevel DBSCAN algorithm procedure builds on DBSCAN [24]. Initially, the algorithm extracts and consolidates features from individual branches by vertically concatenating the features.

This merging of features aims to streamline computational complexity for subsequent stages. To further reduce computational resources, the algorithm filters clustered features by building a new model to describe them and discarding redundant clusters. Central to the algorithm's functionality is the implementation of multilevel DBSCAN for clustering within the CMFD framework. The primary objective is to detect and outline objects of varying sizes within an image.

Structured into three hierarchical levels, each characterized by specific epsilon (eps) and minimum points (minPts) values, the algorithm iteratively applies DBSCAN to generate clusters representing potentially forged or copied regions. A noteworthy aspect of this algorithm is its hierarchical clustering methodology, depicted in Figure 4. This hierarchical framework facilitates the merging of clusters from lower to higher levels based on proximity. This approach accommodates objects of varying sizes, addressing the diverse range typically encountered in images. The adaptability of the epsilon and minPts parameters across hierarchical levels enhances the algorithm's versatility, enabling it to localize both small and large copied regions within the same image.

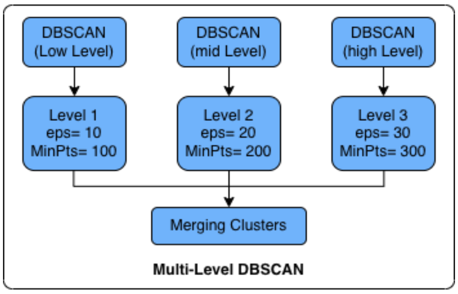


Figure 4. Multilevel DBSCAN

Multilevel DBSCAN offers several advantages for localization in CMFD. Its density-based methodology ensures robustness to noise, enabling the identification of genuine clusters across varying levels of noise. Hierarchical clustering enhances precision by reducing false matches and refining clusters at higher levels based on proximity to lower-level clusters. This refinement yields a more accurate representation of the image's underlying structures, which is crucial for identifying copied regions. Furthermore, the algorithm's adaptability to different object sizes and its consideration of multiple levels contribute to achieving scale and orientation invariance. This adaptability is crucial in CMFD, where copied regions may undergo transformations such as scaling or rotation, requiring a comprehensive approach for accurate detection in digital images. Figure 5 compares the performance of the proposed multilevel DBSCAN. Figures 5(a) and 5(b) depict the original input alongside the corresponding ground truth. Figure 5(c) reports a single-level DBSCAN result alongside the true forged regions, revealing limitations in capturing forgeries across different scales. In contrast, Figures 5(d) present the multilevel DBSCAN output, which more closely aligns with the ground truth by accurately localizing both small and large duplicated areas. This visual comparison highlights the method’s effectiveness and robustness in complex forgery scenarios.

2.4.3. Affine transformation estimation

For each class $k, k = 1, \ldots, c_{max}$, (where c_{max} is the total number of classes), the affine transfer matrix H_k is estimated. For the i^{th} pair in class k, the mapping is:

$$\begin{pmatrix} x_{i2} \\ y_{i2} \\ 1 \end{pmatrix} = \begin{pmatrix} a & b & t_x \\ c & d & t_y \\ 0 & o & 1 \end{pmatrix} \begin{pmatrix} x_{i1} \\ y_{i1} \\ 1 \end{pmatrix} = H_K \begin{pmatrix} x_{i1} \\ y_{i1} \\ 1 \end{pmatrix} \tag{4}$$

Where a, b, c, d encode the linear (scale/rotation) part and t_x, t_y are the translations. Given a set of all keypoint pairs $\{p_1, \ldots, p_i, \ldots, p_{max_k}\}$ in the class k, were max_k is the number of pairs and $pi = \{(x_{i1}, y_{i1}), (x_{i2}, y_{i2})\}$, then search for the optimal matrix H_k that minimizes the following total error function:

$$\arg min_{H_K} \sum_{i=1}^{max_K} \|(x_{i2}, y_{i2}, 1)^T - H_K (x_{i1}, y_{i1}, 1)^T\|^2 \tag{5}$$

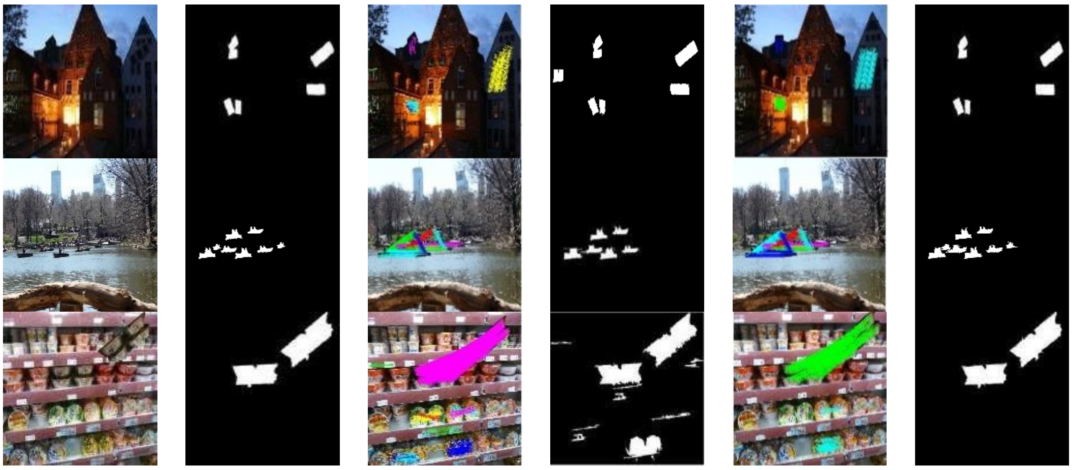
2.4.4. Duplicated-region localization

Given the estimated affine transform T, pixels are compared with their transformed counterparts to reveal duplicated areas:

i) Step 1: forward/backward mapping. Apply T and T^{-1} at every pixel; store the resulting intensities in the forward and backward transformation maps (FTM/BTM).
ii) Step 2: local correlation. Compute a local correlation coefficient between I and each map (FTM/BTM). Let $I(x)$ and $T(x),$) denote the intensities at the location X; the coefficient is:

$$c(x) = \frac{\sum_{\mu \epsilon \Omega(x)}\left(I(\mu)-\bar{I}\right)\left(T(\mu)-\bar{T}\right)}{\sqrt{\sum_{\mu \epsilon \Omega(x)}(I(\mu)-I)^2\left(T(\mu)-\bar{T}\right)^2}} \tag{6}$$

$\Omega_{(x)}$ is a 3×3-pixel neighborhood centered x; the local mean of I and T over this window is $\bar{I}_x$and $\bar{T}_x$. The correlation coefficient is bounded in [0, 1]; higher values indicate stronger similarity. Evaluating it at every pixel produces the forward and backward correlation maps.
iii) Step 3: post-processing. Apply a 3×3 Gaussian filter to denoise the maps, binarize each map using thresholding, and union the masks into a single binary image. Lastly, discard tiny connected components (area below 0.1%) and perform morphological filling to close holes.



(a) (b) (c) (d)

Figure 5. Comparison between single-level and multilevel DBSCAN on sample forged images: (a) original image, (b) ground truth, (c) single-level DBSCAN, and (d) multilevel DBSCAN

3. RESULTS AND DISCUSSION

Evaluation uses precision, recall, and F1-score—computed at the image and pixel levels—as defined:

$$precision = \frac{T_P}{T_P+F_P} \tag{7}$$

$$recall = \frac{T_P}{T_P+F_N} \tag{8}$$

$$F1 = 2 \cdot \frac{precision \cdot recall}{precision+recall} \tag{9}$$

We use TP, FP, and FN to denote, respectively, true positives, false positives, and false negatives. Image level: these quantities count images that are correctly detected as tampered (TP), incorrectly flagged as tampered though original (FP), and tampered but not detected (FN). At the pixel level, they count correctly labeled tampered pixels (TP), incorrectly labeled tampered pixels (FP), and tampered pixels that were missed within the manipulated regions (FN).

3.1. Detection results

The effectiveness and resilience of the proposed CMFD approach. were assessed on two standard datasets: MICC-F600 [25] and COVERAGE [26]. These datasets cover a wide spectrum of copy–move scenarios, ranging from differences in object size and texture smoothness to geometric variations (e.g., scaling and rotation) and various post-processing operations. Performance is systematically compared with state-of-the-art methods using both image-level and pixel-level evaluation metrics. allowing for a detailed comparison of detection accuracy and localization precision. Figure 6 presents difficult COVERAGE forgeries alongside their ground-truth masks and the proposed method's detections. Figure 6(a) illustrates a specific manipulation—e.g., rotation of small objects, rotation in smooth areas, scaling, and combined rotation-plus-scaling. Figure 6(a) displays the forged images, Figure 6(b) contains the corresponding ground-truth, while Figure 6(c) shows the detection results of the proposed method. These visual results indicate that the proposed approach reliably detects and localizes tampered regions across varied post-processing conditions and object complexities. Notably, the method successfully handles small-scale manipulations and smooth texture areas, which are often problematic for keypoint-based CMFD techniques. These qualitative results are consistent with the quantitative improvements reported in Tables 1 and 2, further validating the effectiveness of the hybrid SIFT-AKAZE feature extraction and multilevel DBSCAN-based localization. The proposed algorithm demonstrates remarkable accuracy across various challenges, including rotation of small objects, multiple-copy forgeries, rotation in smooth regions, scaling, and combined scaling and rotation. This underscores the algorithm's effectiveness in detecting and localizing forged regions, even in diverse and complex manipulations.


(a)


(b)


(c)

Figure 6. Difficult COVERAGE forgeries: (a) specific manipulation, (b) corresponding ground-truth, and (c) detection results of the proposed method

A summary of results on COVERAGE and MICC-F600 appears in Tables 1 and 2. As shown in Table 1, our method achieves the highest F1 scores: 0.7724 at the image level and 0.706 at the pixel level. Indicating superior performance in accurately detecting and localizing forged regions. Unlike Buster Net, which has high recall but low precision, our method strikes a balance between both metrics, resulting in more reliable results. Similarly, it outperforms the dual-order attentive generative adversarial network (DOA-GAN), Li, and second-keypoint matching and double adaptive filtering (SMDAF) methods in precision and F1-score, particularly at the pixel level, which reflects improved localization accuracy.

Table 1. Detection results at the image and pixel levels on the COVERAGE dataset.

Method		Image level			Pixel level		
	Detected images	F1-score	Recall	Precision	F1-score	Recall	Precision
Buster Net [27]	200	0.660	0.940	0.508	0.638	0.652	0.624
Li [28]	194	0.688	0.750	0.635	0.595	0.595	0.595
DOA-GAN [29]	200	0.625	0.810	0.509	0.559	0.530	0.591
SMDAF [30]	200	0.711	0.910	0.583	0.678	0.768	0.606
Proposed work	200	0.7724	0.784	0.7644	0.706	0.704	0.7088

Table 2. Results for pixel-level detection on MICC-F600.

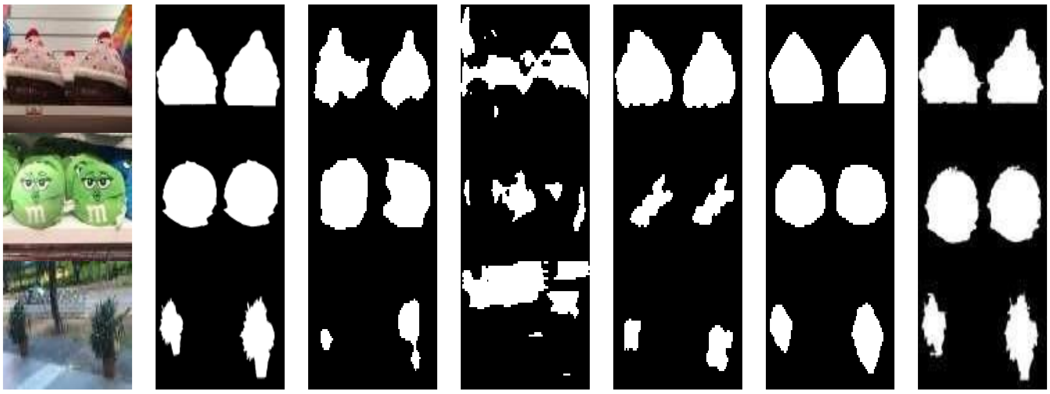
Method		Pixel level	
	F1-score	Recall	Precision
Li [28]	91.80	-	-
Double matching [23]	77.93	81.74	74.46
Jiang [31]	90.39	-	-
Proposed work	93.17	89.067	98.03

In Table 2, evaluated on the MICC-F600 dataset, the proposed method again achieves the highest pixel-level F1-score (93.17%) and precision (98.03%)—the best among all compared methods.

This demonstrates its robustness against scale and rotation attacks, where traditional methods, such as Li's and double matching, tend to struggle. The high precision further indicates a low false-positive rate, making the approach well-suited for digital forensics.

Figure 7 provides a visual comparison of the proposed method with several state-of-the-art techniques—BusterNet, DOA-GAN, Li's method, and SMDAF—using examples from the COVERAGE dataset. Figure 7(a) displays the forged images, and Figure 7(b) shows their corresponding ground-truth masks. The subsequent subfigures present the detection results generated by different methods: Figure 7(c) by BusterNet, Figure 7(d) by the DOA-GAN method, Figure 7(e) by Li's method, Figure 7(f) by SMDAF, and Figure 7(g) by our proposed method.

Particularly in complex scenarios such as multiple forgeries and smooth regions. Unlike Buster Net and SMDAF, which tend to produce false positives or miss small manipulations, our approach accurately captures both small and large forged regions with better alignment to the ground truth. This visual evidence supports the quantitative results presented in Table 1 and further demonstrates the robustness of the multilevel DBSCAN strategy and the effectiveness of the double matching process.



(a) (b) (c) (d) (e) (f) (g)

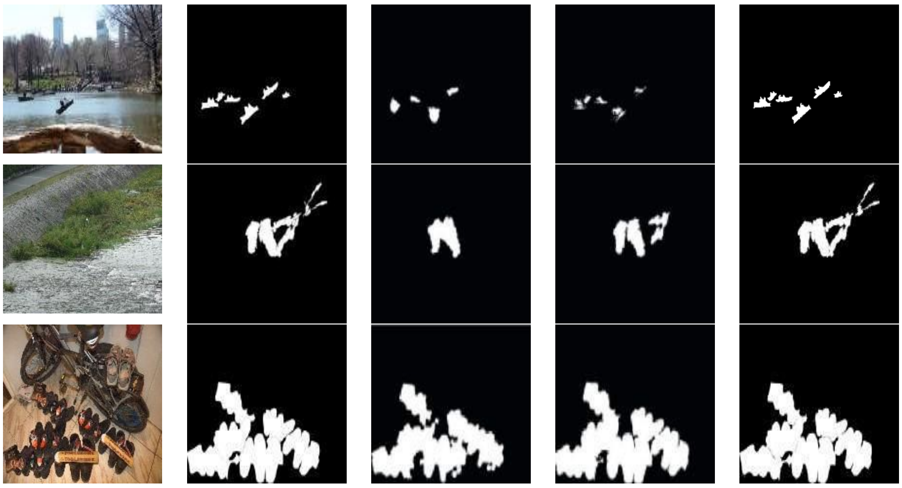
Figure 7. Examples extracted from the COVERAGE dataset: (a) forged images, (b) corresponding ground-truth masks, (c) Buster Net, (d) DOA-GAN method, (e) Li method, (f) SMDAF, and (g) our proposed method

The COVERAGE dataset was used to test the proposed method. Its results were compared with the approaches summarized in Table 1. The Buster Net method demonstrated the highest recall and accuracy rates at the image level. However, it struggled to accurately detect forged images, with an image-level accuracy of 0.508. This suggests that Buster Net struggles to distinguish between forged and raw images, leading to copy-move scenarios. The Li method exhibited limitations in image detection, identifying only 97 of 100 pairs, and may have introduced errors in analyzing specific pairs (Nos. 9, 13, and 36). When evaluating recall, precision, and the F1-score, the three undetected image pairs were included to ensure a fair comparison with other methods. Despite HFPM's good image-level accuracy, the proposed method outperformed Li in various aspects.

Upon incorporating DOA-GAN, the proposed method's image-level results were comparable to Buster Net but exhibited a decline at the pixel level. Meanwhile, the SMDAF method achieved high recall and accuracy rates at the image level. However, like Buster Net, it faced challenges in accurately detecting forged images, reflected in its image-level accuracy of 0.711. In conclusion, the proposed method showed superior performance in distinguishing challenging images at both the image and pixel levels.

The detection results for different images are depicted in Figure 8, highlighting the successful identification of a substantial portion of tampered regions, including smaller ones. Figure 8(a) shows the forged images, Figure 8(b) displays their corresponding masks, Figure 8(c) presents the detection results of Li's method, Figure 8(d) shows the double-matching results, and Figure 8(e) illustrates the outputs of the proposed method. the efficacy of detecting complete regions relies on obtaining an adequate number of keypoint pairs. To achieve this, a double-matching process is employed to ensure robust keypoint correspondence, while the multilevel DBSCAN technique is used to accurately identify objects of varying sizes. In the first image, even the smallest forged objects are detected, demonstrating the method's robustness. The second image shows enhanced detection of intricate details compared to other methods. Although Li and Zhou [28] exhibit commendable performance in detecting tampered regions on average. However, it falls short in three tampered regions where the results are notably undetected. Notably, Li's method and double matching struggle to identify the copied regions in the first and second images, characterized by a 30-degree rotation and 120-degree scaling.

Conversely, the proposed method excels in detecting small objects with clarity. As noted in [31], addressing missed detections and false alarms remains a critical challenge in copy-move forgery detection. However, it still has a low F1-score. In summary, the proposed method demonstrates notable robustness when compared to existing approaches, such as Li's method, the double matching method, and Jiang's, particularly in scenarios involving rotation post-processing and detection of small, smooth regions. These attributes significantly enhance the practical utility of the proposed method in CMFD. With consistent attainment of a high F1-score, the algorithm underscores its effectiveness in accurately detecting tampered images across diverse conditions.



(a) (b) (c) (d) (e)

Figure 8. Examples extracted from the MICC-F600 dataset: (a) forged images, (b) corresponding masks, (c) results from Li [28], (d) double matching, and (e) our proposed method

4. CONCLUSION

In copy–move forgery detection, tackling key challenges posed by geometric transformations and small, smooth, duplicated regions has been particularly arduous. This study presents an innovative approach that integrates AKAZE and SIFT features to harness richer image descriptors, enhancing feature extraction capabilities. By employing a double-matching strategy, the method identifies multiple duplicated regions with greater accuracy. To further refine the detection process, a multilevel DBSCAN algorithm is utilized, effectively mitigating false positives and accommodating forgeries of various sizes. Experimental results underscore the method's robustness against a spectrum of tampering conditions, including rotations, scaling, blurring, and brightness variations. Combining SIFT, AKAZE, and the advanced multilevel DBSCAN algorithm significantly enhances forgery detection performance. The method achieves a high F1-score, demonstrating commendable precision and recall, thereby significantly advancing image forgery detection.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Doaa Gamal Abdelazem		✓	✓		✓	✓	✓	✓	✓			✓		
Hala H. Zayed	✓	✓		✓	✓			✓		✓		✓	✓	
Ahmed Taha	✓	✓		✓			✓		✓	✓	✓	✓	✓	

C : **C**onceptualization	I : **I**nvestigation	Vi : **Vi**sualization
M : **M**ethodology	R : **R**esources	Su : **Su**pervision
So : **So**ftware	D : **D**ata Curation	P : **P**roject administration
Va : **Va**lidation	O : Writing - **O**riginal Draft	Fu : **Fu**nding acquisition
Fo : **Fo**rmal analysis	E : Writing - Review & **E**diting	

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] H. Farid, "Image forgery detection," *IEEE Signal Processing Magazine*, vol. 26, no. 2, pp. 16–25, Mar. 2009, doi: 10.1109/MSP.2008.931079.
[2] A.-T. P. -Ho and F. Retraint, "A comparative study of bayesian and dempster-shafer fusion on image forgery detection," *IEEE Access*, vol. 10, pp. 99268–99281, 2022, doi: 10.1109/ACCESS.2022.3206543.
[3] K. M. Hosny, A. M. Mortda, M. M. Fouda, and N. A. Lashin, "An efficient CNN model to detect copy-move image forgery," *IEEE Access*, vol. 10, pp. 48622–48632, 2022, doi: 10.1109/ACCESS.2022.3172273.
[4] S. I. Lee, J. Y. Park, and I. K. Eom, "CNN-based copy-move forgery detection using rotation-invariant wavelet feature," *IEEE Access*, vol. 10, pp. 106217–106229, 2022, doi: 10.1109/ACCESS.2022.3212069.
[5] A.-R. Gu, J.-H. Nam, and S.-C. Lee, "FBI-net: frequency-based image forgery localization via multitask learning with self-attention," *IEEE Access*, vol. 10, pp. 62751–62762, 2022, doi: 10.1109/ACCESS.2022.3182024.
[6] M. Kaya, K. J. Sani, and S. Karakus, "Copy-move forgery detection in digital forensic images using CNN," in *2022 7th International Conference on Computer Science and Engineering (UBMK)*, Diyarbakir, Turkey: IEEE, Sept. 2022, pp. 239–245, doi: 10.1109/UBMK55850.2022.9919560.
[7] Y. Zhang *et al.*, "CNN-transformer based generative adversarial network for copy-move source/ target distinguishment," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 33, no. 5, pp. 2019–2032, May 2023, doi: 10.1109/TCSVT.2022.3220630.
[8] J. Fridrich, D. Soukal, and J. Lukas, "Detection of copy-move forgery in digital images," in *Proceedings of digital forensic research workshop*, Cleveland, OH, 2003, pp. 652–63.
[9] Z. He, W. Lu, Wei Sun, and J. Huang, "Digital image splicing detection based on Markov features in DCT and DWT domain," *Pattern Recognition*, vol. 45, no. 12, pp. 4292–4299, Dec. 2012, doi: 10.1016/j.patcog.2012.05.014.
[10] A. C. Popescu and H. Farid, "Exposing digital forgeries by detecting duplicated image regions," *Technical Reports:TR2004-515,* 2004. [Online]. Available: https://digitalcommons.dartmouth.edu/cs_tr/254/
[11] W. Luo, J. Huang, and G. Qiu, "Robust detection of region-duplication forgery in digital image," in *18th International Conference on Pattern Recognition (ICPR'06)*, Hong Kong, China: IEEE, 2006, pp. 746–749, doi: 10.1109/ICPR.2006.1003.
[12] K. Hayat and T. Qazi, "Forgery detection in digital images via discrete wavelet and discrete cosine transforms," *Computers & Electrical Engineering*, vol. 62, pp. 448–458, Aug. 2017, doi: 10.1016/j.compeleceng.2017.03.013.
[13] M. H. Alkawaz, G. Sulong, T. Saba, and A. Rehman, "Detection of copy-move image forgery based on discrete cosine transform," *Neural Computing and Applications*, vol. 30, no. 1, pp. 183–192, Jul. 2018, doi: 10.1007/s00521-016-2663-3.
[14] E. Silva, T. Carvalho, A. Ferreira, and A. Rocha, "Going deeper into copy-move forgery detection: Exploring image telltales via multi-scale analysis and voting processes," *Journal of Visual Communication and Image Representation*, vol. 29, pp. 16–32, May 2015, doi: 10.1016/j.jvcir.2015.01.016.
[15] R. C. Pandey, S. K. Singh, K. K. Shukla, and R. Agrawal, "Fast and robust passive copy-move forgery detection using SURF and SIFT image features," in *2014 9th International Conference on Industrial and Information Systems (ICIIS)*, Gwalior, India: IEEE, Dec. 2014, pp. 1–6, doi: 10.1109/ICIINFS.2014.7036519.
[16] L. Yu, Q. Han, and X. Niu, "Feature point-based copy-move forgery detection: covering the non-textured areas," *Multimedia Tools and Applications*, vol. 75, no. 2, pp. 1159–1176, Jan. 2016, doi: 10.1007/s11042-014-2362-y.
[17] X.-Y. Wang, S. Li, Y.-N. Liu, Y. Niu, H.-Y. Yang, and Z. Zhou, "A new keypoint-based copy-move forgery detection for small smooth regions," *Multimedia Tools and Applications*, vol. 76, no. 22, pp. 23353–23382, Nov. 2017, doi: 10.1007/s11042-016-4140-5.
[18] Z. Wang, B. Fan, and F. Wu, "Local intensity order pattern for feature description," in *2011 International Conference on Computer Vision*, Barcelona, Spain: IEEE, Nov. 2011, pp. 603–610, doi: 10.1109/ICCV.2011.6126294.
[19] H. Bay, T. Tuytelaars, and L. V. Gool, "SURF: speeded up robust features," in *Computer Vision – ECCV 2006*, Berlin, Heidelberg: Springer, 2006, pp. 404–417, doi: 10.1007/11744023_32.
[20] P. Zhang and X. Yan, "Application of improved KAZE algorithm in image feature extraction and matching," *IEEE Access*, vol. 11, pp. 122625–122637, 2023, doi: 10.1109/ACCESS.2023.3328778.
[21] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004, doi: 10.1023/B:VISI.0000029664.99615.94.
[22] P. Alcantarilla, J. Nuevo, and A. Bartoli, "Fast explicit diffusion for accelerated features in nonlinear scale spaces," in *Procedings of the British Machine Vision Conference 2013*, Bristol: British Machine Vision Association, 2013, pp. 13.1-13.11, doi: 10.5244/C.27.13.
[23] Q. Lyu, J. Luo, K. Liu, X. Yin, J. Liu, and W. Lu, "Copy move forgery detection based on double matching," *Journal of Visual Communication and Image Representation*, vol. 76, Apr. 2021, doi: 10.1016/j.jvcir.2021.103057.

[24] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *KDD'96: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, 1996, pp. 226–231.
[25] I. Amerini, L. Ballan, R. Caldelli, A. D. Bimbo, L. D. Tongo, and G. Serra, "Copy-move forgery detection and localization by means of robust clustering with J-Linkage," *Signal Processing: Image Communication*, vol. 28, no. 6, pp. 659–669, Jul. 2013, doi: 10.1016/j.image.2013.03.006.
[26] B. Wen, Y. Zhu, R. Subramanian, T.-T. Ng, X. Shen, and S. Winkler, "COVERAGE — A novel database for copy-move forgery detection," in *2016 IEEE International Conference on Image Processing (ICIP)*, Phoenix, AZ, USA: IEEE, Sep. 2016, pp. 161–165, doi: 10.1109/ICIP.2016.7532339.
[27] Y. Wu, W. Abd-Almageed, and P. Natarajan, "BusterNet: detecting copy-move image forgery with source/target localization," in *Computer Vision – ECCV 2018*, Cham, Switzerland: Springer, 2018, pp. 170–186, doi: 10.1007/978-3-030-01231-1_11.
[28] Y. Li and J. Zhou, "Fast and effective image copy-move forgery detection via hierarchical feature point matching," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 5, pp. 1307–1322, May 2019, doi: 10.1109/TIFS.2018.2876837.
[29] A. Islam, C. Long, A. Basharat, and A. Hoogs, "DOA-GAN: dual-order attentive generative adversarial network for image copy-move forgery detection and localization," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA: IEEE, Jun. 2020, pp. 4675–4684, doi: 10.1109/CVPR42600.2020.00473.
[30] G. Yue, Q. Duan, R. Liu, W. Peng, Y. Liao, and J. Liu, "SMDAF: A novel keypoint based method for copy-move forgery detection," *IET Image Processing*, vol. 16, no. 13, pp. 3589–3602, Nov. 2022, doi: 10.1049/ipr2.12578.
[31] L. Jiang, Z. Lu, Y. Gao, and Y. Wang, "Image copy-move forgery detection and localization scheme: how to avoid missed detection and false alarm,", *arXiv:2406.03271*, 2024.

BIOGRAPHIES OF AUTHORS

Doaa Gamal Abdelazem has been a Teaching Assistant in the Department of Computer Science at the Modern Academy for Computer Science and Management Technology, Egypt, since 2018. She holds a B.Sc. in Computer Science. She can be contacted at email: d.doaa27150@fci.bu.edu.eg or doaag6349@gmail.com.

Prof. Hala H. Zayed received the B.Sc. degree (Hons.) in Electrical Engineering and the M.Sc. and Ph.D. degrees in Electronics Engineering from Benha University, in 1985, 1989, and 1995, respectively. She is currently a Professor at the Faculty of Engineering, Egypt University of Informatics, Egypt. Her research interests include computer vision, biometrics, machine learning, image forensics, and image processing. She can be contacted at email: hala.zayed@eui.edu.eg or hala.zayed@gmail.com.

Dr. Ahmed Taha received his M.Sc. degree and his Ph.D. degree in Computer Science, at Ain Shams University, Egypt, in February 2009 and July 2015. He currently works as an Associate Professor at the Department of Computer Science, Benha University, Egypt. He is the founder and coordinator of "Information Security and Digital Forensics" program, Faculty of Computers and Artificial Intelligence, Benha University. His research interest's concern: computer vision and image processing (human behavior analysis, video surveillance systems), digital forensics (image forgery detection, document forgery detection), security (encryption, steganography, cloud computing), content-based retrieval (Arabic text retrieval, video scenes classification, video scenes retrieval, trademark image retrieval, closed-caption technology). He can be contacted at email: ahmed.taha@fci.bu.edu.eg.