Bulletin of Electrical Engineering and Informatics
Vol. 15, No. 4, August 2026, pp. 3080~3094
ISSN: 2302-9285, DOI: 10.11591/eei.v15i4.10796

Image copy-move forgery detection: a survey of methods, datasets, and emerging trends

Li-xian Jiao[1,2], Kok-Why Ng[2], Hau-Lee Tong[2]
[1]Department of Computer Science, School of Computer and Big Data, Jining Normal University, Ulanqab, China
[2]Department of Computer Science, Faculty of Computing and Informatics, Multimedia University, Cyberjaya, Malaysia

Article Info

Article history:

Received Jun 2, 2025
Revised May 6, 2026
Accepted Jun 19, 2026

Keywords:

Benchmark datasets
Block-based
Copy-move forgery detection
Deep learning
Digital image forensics
Keypoint-based

ABSTRACT

Digital image forgery has become a critical concern in the era of advanced multimedia technologies, where the authenticity of visual content directly affects trust in digital communication, journalism, and law enforcement. Among various forgery techniques, copy-move forgery (CMF) is among the most common and deceptive, as it involves duplicating a region of an image to conceal or misrepresent information. To address this challenge, numerous copy-move forgery detection (CMFD) approaches have been proposed, ranging from block-based and keypoint-based methods to hybrid models and deep learning (DL) techniques. This paper provides a comprehensive review of these approaches, analyzing their strengths and limitations, and evaluating their performance across multiple benchmark datasets. The evaluation considers factors such as image resolution, manipulation types, and robustness against post-processing attacks. By systematically comparing the algorithms and datasets, the study highlights persistent challenges and outlines future research directions. The findings aim to guide researchers in selecting appropriate techniques and inspire the development of more robust CMFD solutions.

Corresponding Author:

Kok-Why Ng
Department of Computer Science, Faculty of Computing and Informatics, Multimedia University
Persiaran Multimedia, 63100 Cyberjaya, Selangor, Malaysia
Email: kwng@mmu.edu.my

1. INTRODUCTION

In recent years, network infrastructure has improved, broadening the reach and influence of digital images. Free and paid image processing software, such as Adobe Photoshop and Paint Shop Pro, make digital image processing easier to obtain and popularize. Users can easily operate on and edit images with minimal loss of quality and few visual artefacts [1]. But this convenience carries potential risks, as images may be maliciously tampered with, distorting their original meaning.

There was a famous case in the 2006 China News Photography Competition: the gold award-winning work "The First Thunder of Chinese Rural Urbanization Reform" was found to be forged after winning the award, which was spliced from photos taken at different scenes at the same time. In Figure 1, there are two cases: Figure 1(a) is a modified composite image and Figure 1(b) is an unmodified real scene.

Such incidents underscore the growing issue of digital image forgery, which presents substantial challenges to fields such as news reporting, academic research, legal proceedings, and healthcare. Forged images may damage credibility and endanger social stability [2], [3]. So, it is necessary to ensure the authenticity of digital images as soon as possible. The emergence of digital image forensics technology is an important development.

(a) (b)

Figure 1. The falsified digital image; (a) forged image and (b) original image

There are two methods in existing digital image forensics technology: active (non-blind) and passive (blind) [4], [5]. Figure 2 presents these categories and subtypes. Active forensics includes digital signature and watermarking technologies [6]. Watermarking is more flexible than digital signatures. Active forensic technology is effective in some cases, but it has certain limitations, such as the need for the original image for embedding and for decryption/extraction tools. In some cases, passive forensics may be more appropriate because there is no need to obtain information in advance [4]. Passive methods detect image manipulation by analyzing inconsistencies in content and structure. Examples of passive methods include image splicing, copy-move forgery (CMF), image retouching, and object removal. In blind digital image forensics, CMF is widely used due to its own characteristics [7]–[9]. From 2015 to 2025, the Scopus database listed 992 research papers on CMFD. Figure 3 presents the distribution of annual releases, revealing the development of research trends.

Section 2 presents various evaluation metrics used for forgery detection. Section 3 provides detailed information about the benchmark datasets employed in CMFD. Section 4 surveys state-of-the-art CMFD methods, providing a comprehensive review. Section 5 discusses research issues and future directions in CMFD. Finally, the conclusions are presented in section 6.

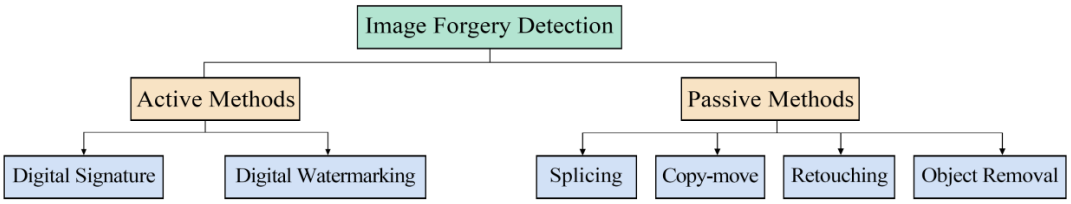


Figure 2. Classification of image forgery detection

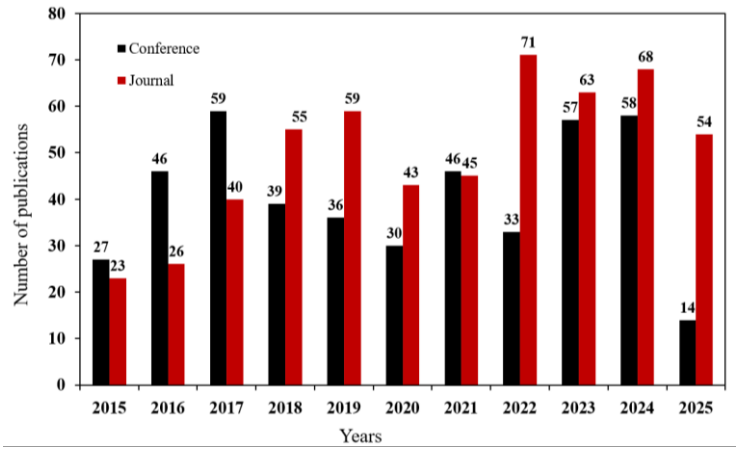


Figure 3. CMFD publications indexed by Scopus

2. EVALUATION METRICS

The evaluation of CMFD techniques typically occurs on two levels: image and pixel [10]–[12]. At the image level, the focus is on determining whether an image is real or tampered with. In contrast, pixel-level assessment classifies each pixel as either real or tampered, making it the most accurate and reliable method. This approach provides granular information, enabling precise identification of the location and shape of the forged region, rather than merely classifying the entire image.

The evaluation metrics for CMFD technology are based on measurements conducted at the image or pixel level [2], [13]. These metrics include:

True positive (TP): number of images/pixels correctly recognized as tampered.
False positive (FP): number of real images/pixels incorrectly recognized as tampered.
True negative (TN): number of real images/pixels correctly identified as real.
False negative (FN): number of tampered images/pixels incorrectly identified as real.

Based on the above measurement, various evaluation metrics can be calculated at either the image or pixel level. For instance, standard evaluation metrics are often pixel-level. Note that for image-level evaluation, these components refer to entire images instead of individual pixels.

Accuracy: this metric represents the proportion of correctly classified pixels to the total number of pixels. The accuracy ranges from 0 to 1, with higher values indicating greater accuracy in classifying the entire image. The calculation formula is as (1):

$$ACC = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

Precision: this metric is the ratio of correctly classified forged pixels (TP) to the total number of predicted forged pixels. Precision values range from 0 to 1, with higher values indicating greater accuracy in labeling forged pixels. The calculation formula is as (2):

$$P = \frac{TP}{TP+FP} \quad (2)$$

Recall/true positive rate (TPR): this metric is the ratio of correctly classified forged pixels (TP) to the total number of forged pixels. The value ranges from 0 to 1, with higher values indicating greater accuracy in identifying forged pixels. The calculation formula is as (3):

$$R/TPR = \frac{TP}{TP+FN} \quad (3)$$

F1-score: this metric represents the harmonic mean of precision and recall, providing a comprehensive evaluation of performance. The value ranges from 0 to 1, with higher values indicating a better balance between precision and recall. The calculation formula is as (4):

$$F1 = 2 * \frac{P*R}{P+R} = 2 * \frac{TP}{2*TP+FN+FP} \quad (4)$$

An excellent CMFD detection method should achieve high values for ACC, P, R, and F1 [14], [15]. This ensures accurate detection of forgery areas, comprehensive coverage of all forged regions, and minimized false positives. Such performance characteristics enhance the reliability and accuracy of forgery detection.

3. DATASETS

The CMFD field offers a variety of datasets with diverse characteristics, each unique in terms of capacity, ground truth representation, and complexity [16]. As shown in Figure 4, the datasets differ in the distribution of authentic (A) and tampered (T) images. MICC-F220 maintains a relatively balanced distribution, while datasets such as CASIA V2.0 and CoMoFoD contain more tampered images, reflecting varying degrees of complexity.

The datasets in Table 1 are commonly used to evaluate CMFD algorithms, but they vary significantly in scale, diversity, and attack authenticity. The MICC series [17] (F220, F2000, F600) is simple to operate and suitable for baseline evaluation, but has limited post-processing; FAU [18] and CoMoFoD [19] contain multiple geometric transformations and post-processing, which are more challenging. CASIA v2.0 [20] covers copy-move and splicing forgeries, but has low resolution and insufficient annotations. Small datasets (such as CVIP [21] and GRIP [22]) have a single test condition and are too small to easily lead to overfitting. COVERAGE [23] is known for its complex scenes, whereas IMD [24] provides high-resolution images that are more realistic. Figure 5 presents examples of forged images from several public datasets.

Figures 5(a)–(d) correspond to FAU [18], CoMoFoD [19], CVIP [21], and GRIP [22], respectively. Overall, CoMoFoD and FAU provide more realistic challenges, but existing datasets still have limitations, such as single-texture samples, unbalanced scenes, and limited scale, which affect algorithmic generalization. In the future, larger, more diverse, and balanced datasets should be constructed to narrow the gap between experimental environments and real-world forgery scenarios.

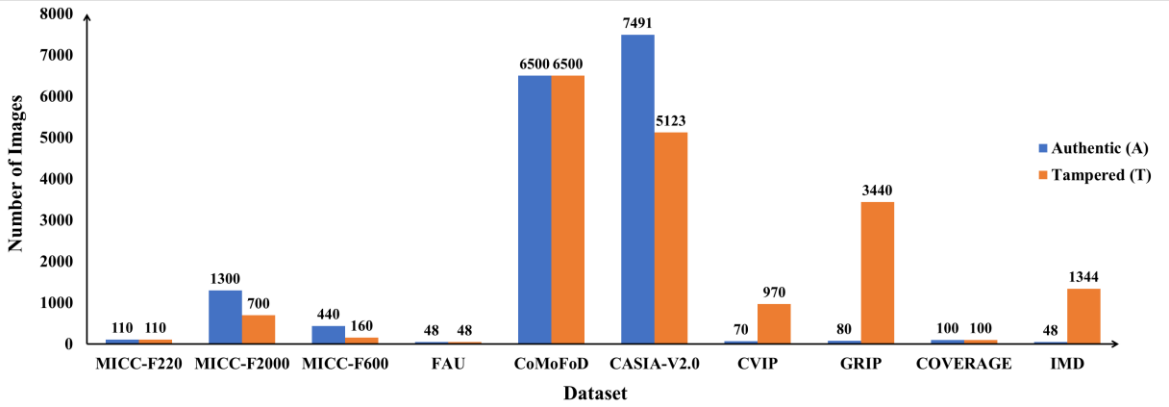
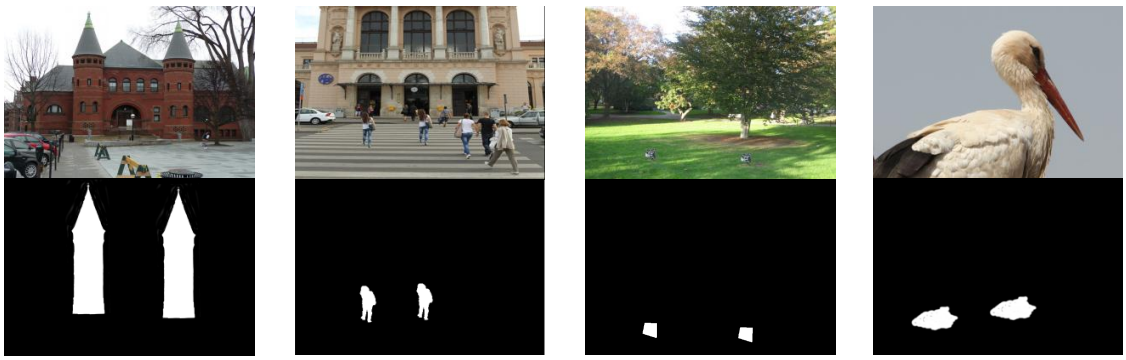


Figure 4. A bar chart shows the number of authentic and tampered items for each dataset

Table 1. Summary of the CMFD datasets

Dataset	No. of images	Size	Format	Geometric transforms	Post processing operations
MICC-F220 [16]	A:110 T:110	722*480 to 800*600	JPEG	S and R	No
MICC-F2000 [16]	A:1300 T:700	2048*1536	JPEG	S and R	No
MICC-F600 [16]	A:440 T:160	800*532 to 3888*2592	JPEG and PNG	S and R	No
FAU [17]	A:48 T:48	3000*2300	JPEG and PNG	S, R, T, D, and C	NA, JC, BC, CA, and CR
CoMoFoD [18]	A:6500 T:6500	512*512 to 3000*2000	JPEG and PNG	S, R, and D	NA, JC, BC, CA, and CR
CASIA V2.0 [19]	A:7491 T:5123	240*160 to 900*600	JPEG, BMP, and TIFF	S, R, and D	B
CVIP [20]	A:70 T:970	1000*700 or 700*1000	BMP	S and R	No
GRIP [21]	A:80 T:3440	768*1024	PNG	S and R	NA and JC
COVERAGE [22]	A:100 T:100	400*486	TIF	S, R, T, and C	BC
IMD [23]	A:48 T:1344	3000*2400	JPEG	S and R	NA and JC

Note: A=authentic, T=tampered, S=scaling, R=rotation, T0=transformation, D=distortion, C=combination, B=blurring, NA=noise adding, JC=JPEG compression, BC=brightness change, CA=contrast adjustment, and CR=color reduction



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

Figure 5. Sample forged images and binary masks; (a) FAU, (b) CoMoFoD, (c) CVIP, and (d) GRIP

4. RELATED WORK

CMFD methods fall into four categories: block-based, keypoint-based, combination-based, and deep learning (DL) methods. In this section, we provide a comprehensive review of each category. Figure 6 illustrates the classification of CMFD techniques.

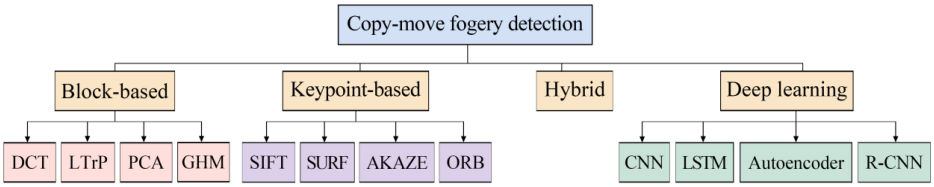


Figure 6. The categorization of CMFD techniques

4.1. Block-based forgery detection methods

In block-based image CMFD methods, the input image is divided into overlapping or non-overlapping blocks, typically circular or rectangular. Feature vectors are then extracted using techniques such as the discrete cosine transform (DCT), the polar complex exponential transform (PCET), principal component analysis (PCA), the histogram of oriented gradients (HOG), or Hu moments, either individually or in combination. The extracted features are matched based on correlation, Euclidean distance, or dictionary ranking. To refine detection and eliminate false matches, forgery localization is performed using random sample consensus (RANSAC) and J-Linkage, which estimate geometric transformations between matched pairs. Table 2 summarizes representative block-based CMFD methods, highlighting their details, benchmark datasets, reported metrics, and key strengths and limitations.

4.1.1. Frequency transform

The frequency transform is a widely used block-based feature-extraction method, valued for its robustness to rotation, translation, and noise. In papers [25], [26], they first convert the image from RGB to grayscale during pre-processing. The grayscale image is then divided into overlapping circular blocks, enabling information sharing between adjacent regions. In the feature extraction stage, PCET features are computed for each circular block. For block matching, the approximate nearest neighbor (ANN) search is employed to identify potentially similar blocks. In the post-processing stage, a block distance threshold helps eliminate false matches by discarding overly similar but untampered blocks, ensuring accurate detection of tampered regions.

Alkawaz *et al*. [27] convert color images to grayscale during pre-processing, followed by division into overlapping blocks. In the feature extraction stage, the DCT is applied to each block, a common technique for transforming spatial-domain images into the frequency domain. The resulting frequency coefficients serve as feature descriptors. During block matching, feature vectors are sorted using dictionary ordering, and duplicate blocks are identified based on Euclidean distance.

Building on these foundations, Warif *et al*. [28] employ the Fourier-Mellin transform (FMT) to extract stable features against various attacks. However, this method did not account for compressed images, leading to either the loss of some detected forged pixels or increased noise in the detected forgeries. To address this, Abir *et al*. [29] proposed automatic frequency-based enhancement filters. They utilized the Otsu algorithm for image pre-processing and applied four enhancement filters to improve the image for display or further analysis. During feature extraction, they combined DCT and FMT to address compression issues. In the block matching stage, PatchMatch was used to identify matches for each patch. Finally, in the post-processing stage, either the RANSAC or the same sampling transform selection (SATS) algorithms were employed to determine the affine transformation for CMFD.

4.1.2. Texture and intensity

Different natural scenes, such as grass, clouds, and ground, exhibit distinct intensity and texture characteristics. When an attacker copies and moves a region within an image, it can lead to noticeable differences in brightness or alterations in texture features. Analyzing these discrepancies can be useful for detecting image forgery.

Uliyan *et al.* [30] propose a method that combines statistical analysis and color texture to identify and segment regions of interest in images. The degree of blurring in these regions is assessed using blurring metrics. Subsequently, the image is divided into fuzzy and non-fuzzy parts. Fourier coefficients and gradient features are employed to detect blurred regions, and the similarity between these regions is estimated using the Euclidean distance.

Table 2. Summary of representative block-based CMFD methods

Category	Ref.	Details	Dataset(s)	Metrics (%)	Strengths	Limitations
Frequency transform	Emam *et al*. [26]	Circular blocks, PCET, ANN, and morphological operations.	MICC-F600	R=79, P=89, F1=83.7	Robust to rotation, scaling, and other geometric transforms.	Computationally intensive; performance decreases under heavy noise.
	Alkawaz *et al*. [27]	overlapping blocks, DCT, dictionary ordering, and Euclidean distance.	CoMoFoD	R=96.6, P=64.5	Simple and effective against compression artifacts.	Limited robustness to geometric transformations (e.g., scaling, and rotation).
	Warif *et al*. [28]	Auto-enhanced and enhanced filter, DCT, FWT, and PatchMatch.	GRIP	F1=93%	Enhances image quality while suppressing noise in detection.	Inability to accurately determine temporal performance.
Texture and intensity	Uliyan *et al*. [30]	SRM segmentation, blur estimator, Fourier coefficients, and gradient-based features, Euclidean distance.	FAU and MICC-F220	TPR=97, FPR=2.6, TPR=96.5, FPR=2.86	Robust to varying levels of blur noise in images.	If the image is highly blurred, the detection rate drops dramatically.
	Kumar *et al*. [31]	Haar transform, PCA, GLCM, and Euclidean distance.	kidney disease dataset	ACC=92	The method has a high peak signal-to-noise ratio (PSNR) and a low mean square error (MSE).	Struggles with low-texture/ homogeneous backgrounds
	Ganguly *et al*. [32]	Overlapping blocks, LTrP, Lexical sorting, Euclidean distance, and morphological operation.	GRIP and CoMoFoD	F1=98.34, F1=90.93	Effective against a range of post-processing manipulations.	For rotation and scaling attacks, the method yields no desirable results.
Dimensionality reduction	Dixit and Bag [33]	Overlapping blocks, HOG and SVD, Lexographical sorting, Euclidean distance, and morphological operation.	CoMoFoD	F1=81.83	Capable of detecting multiple copy-move forgeries; effective feature compression.	Computationally more expensive; limited scalability for high-resolution images.
	Priyanka *et al*. [34]	8*8 overlapping blocks, DCT, SVD, SVM, K-means clustering.	FAU	ACC=93	Improved robustness against geometric transformations; reduces false positives.	Performance degrades on compressed/noisy images.
	Kumar *et al*. [35]	SWT decomposition, overlapping blocks, DCT, SVD, Lexographical sorting, and morphological operation.	CoMoFoD	ACC=98.90, FPR=0.66, R=81.75, P=91, F1=84.42	Lower computational cost; suitable for large datasets.	Less robust under strong post-processing (e.g., blur and scaling).
Moment invariant	Meena and Tyagi [36]	Overlapping square blocks, GHM, Euclidean distance, RANSAC, and morphological operation.	GRIP & CoMoFod	R=98.11, P=97.99, F1=98.05	The method exhibits greater robustness to rotation, scaling, noise, and JPEG compression.	High computational complexity; less scalable for large datasets.
	Mahmood *et al*. [37]	Square blocks, Tchebichef moments, SVD, Euclidean distance, morphologic operations.	CoMoFod and COVERAGE	R=94.82, P=96.92, F1=95.86	Effective across diverse post-processing operations; simple design.	Sensitive to rotation and scaling; reduced accuracy in homogeneous regions.
	Kashyap *et al*. [38]	Elemental detection, wavelet decomposition, overlapping block, Zernike moments, PCA, parameter estimation.	MICC-F220, MICC-F600 and MICC-F2000	R=91.89, P=94.44, F1=93.15	High robustness to rotation; improved accuracy and speed.	Narrow focus on rotation; limited robustness to compression/blur.

Kumar *et al*. [31] extract image features using the Haar transform and subsequently reduce them via PCA. Building upon this, they analyze the texture features of the input image using a gray-level co-occurrence matrix (GLCM). The GLCM is employed to capture the spatial relationships among pixel intensities, thereby aiding the extraction of texture information. Finally, the features are matched using the Euclidean distance metric, with mismatched features being identified as forged regions.

To further enhance robustness, Ganguly *et al*. [32] first divide the input image into overlapping blocks. They then extract local tetra pattern (LTrP) features from each block to create a single feature vector. Since LTrP is texture-based, it remains unaffected by variations in brightness, contrast, and color. The feature vectors of all image blocks are subsequently sorted in dictionary order. Similar blocks are identified by matching features between neighboring blocks. Finally, an offset vector-assisted outlier removal method is employed to discard incorrectly matched blocks, which may result from homogeneous color regions (e.g., sea, field, and sky).

4.1.3. Dimensionality reduction

Feature representation methods based on dimensionality reduction are often integrated with other techniques to enhance performance. The dimensionality of features significantly affects the algorithm's operational efficiency. By simplifying the dimensionality, the time complexity of the feature matching phase is reduced. As a result, many methods are specifically designed to retain the most important information while minimizing redundancy, leading to more efficient image forgery detection [39].

Dixit and Bag [33] divide the suspicious image into overlapping blocks, from which features are extracted using HOG and singular value decomposition (SVD). They then perform dictionary ranking on the feature matrix and calculate the Euclidean distance to identify similar feature vectors. Finally, concatenated components are used for labeling to eliminate false matches and enhance the accuracy of detection. Although this method is effective for multi-region detection, its robustness is limited when images are compressed or noisy.

Priyanka *et al*. [34] begin by dividing the image into blocks and applying DCT to each block. They then reduce the dimensionality of the feature vectors using SVD. This combination of DCT and SVD enhances the method's robustness to compression, geometric transformations, and noise. Next, a support vector machine (SVM) classifier is used to determine whether the image is genuine or fake. If the image is classified as fake, K-means clustering is applied to the feature vectors. Finally, similar blocks are identified and labeled based on a distance threshold. However, this method still suffers from problems such as high feature dimension and high computational cost.

Building on this foundation, Kumar *et al*. [35] apply stationary wavelet transform (SWT) decomposition to the luminance channel (Y) after transforming the color image. Significant features are then selected using DCT for each overlapping block in the approximate frequency bands. The dimensionality of these features is subsequently reduced using SVD. Finally, the simplified features are analyzed and matched to detect corresponding regions within the image.

4.1.4. Moment invariant

Moment invariants offer significant advantages as a highly stable feature representation. They preserve image stability even under geometric transformations and common signal attacks. As a result, moment invariants are widely utilized in CMFD, providing a robust method for identifying manipulated regions in images.

Meena and Tyagi [36] segment the input image into overlapping blocks of fixed size and extract Gaussian-Hermite moments (GHM) for each block. The extracted features are then sorted in dictionary order. Subsequently, a threshold-based Euclidean distance is applied to match the sorted features. Finally, the RANSAC algorithm, along with morphological operations, is employed to filter out anomalies and improve detection accuracy. However, the computational efficiency of this method is reduced.

Mahmood *et al*. [37] segment suspicious images into overlapping blocks and calculate Tchebichef moments for each block to represent the features. Tchebichef moments are chosen for their efficiency in representing image features compared to other moments, such as Legendre and Zernike moments. They further divide each momentary block into non-overlapping blocks and process them using SVD for dimensionality reduction and feature analysis. But the method's sensitivity to rotation and scaling limits its robustness.

More recently, Kashyap *et al*. [38] began by applying wavelet decomposition to the input image and then tiled the rough regions with overlapping blocks. Zernike moments are applied to each block to describe its features. Dimensionality is subsequently reduced using PCA to extract the primary information. Finally, the similarity between blocks is analyzed to identify replicated regions within the image.

4.2. Keypoint-based forgery detection methods

Keypoint-based CMFD techniques bypass the chunking operation during preprocessing. The feature extraction phase in these methods involves two main steps: feature detection and feature description. In feature detection, key points or regions in the image that are unique and stable under geometric transformations are identified. Once detected, descriptors for these key points are generated. Feature description involves creating a representation for each key point that remains unique and invariant to geometric transformations. Ultimately, duplicate regions in the image are identified by classifying and matching these feature descriptors.

Scale-invariant feature transform (SIFT) and speeded-up robust features (SURF) are the most commonly used algorithms in keypoint-based CMFD techniques [40], [41]. In contrast, algorithms like the Harris corner detector and maximally stable extremal regions (MSERs) perform only feature detection. These feature detectors must be paired with other algorithms to characterize features. Table 3 summarizes representative keypoint-based CMFD methods, highlighting their details, benchmark datasets, reported metrics, and key strengths and limitations.

Table 3. Summary of representative keypoint-based CMFD methods

Category	Ref.	Details	Dataset(s)	Metrics (%)	Strengths	Limitations
SIFT-based algorithms	Chen *et al.* [42]	SIFT, ReversedG2NN, HAC, and J-Linkage.	GRIP and FAU	R=99.67, P=99.79, F1=99.72	High robustness for processed images, such as JPEG compression, scaling, and contrast adjustment.	Less stable and less effective for detecting large-scale forgery areas.
	Zedan *et al.* [43]	SIFT, Manhattan distance, intersection over union, and SVM.	MICC-F600, MICC-F220, and COVERAGE	F1=98.42; F1=95.49; F1=80.05	Reduces mismatches, improves precision, and performs well in complex forgery scenarios.	Increased computational overhead; limited generalization to highly diverse datasets.
	Rathi and Singh [44]	SIFT, K-means clustering, Euclidean distance for feature similarity, and a threshold-based decision rule.	CASIA V2.0	Accuracy= 97.85	Faster detection with reduced redundancy; robust to rotation and scaling.	Lower robustness in homogeneous or low-texture areas; requires careful parameter tuning.
SURF-based algorithms	Roy *et al.* [45]	SURF, RLBP, G2NN, Euclidean distance, hierarchical clustering, and RANSAC.	COVERAGE , CoMoFoD, and FAU	ACC=70.05 ; ACC=70.1; ACC=83.3	Robust in distinguishing forged regions from genuine duplicates; reduces false positives in ambiguous cases.	Limited scalability; not validated on large public CMFD datasets.
	Narasimh amurthy *et al*. [46]	CCL, M-SURF, A-KAZE, RANSAC, and SVD.	MICC-F220, MICC-F2000, and CoMoFoD	P=99.93, F1=.98.21, TPR=99.93, ACC=99.46 , FPR=4.67	Effective against rotation, scaling, and noise; improved keypoint repeatability.	Higher computational cost due to redundant features; weaker performance in low-texture regions.
	Patilano and Ballera [47]	BLOB, SURF, ORB, nearest neighbor, and RANSAC.	Real-world multi-object tampering	P=96.8, R =95.2, F1 =96.0	Strong adaptability to multi-region tampering; robust hybrid feature combination improves localization.	Computationally expensive; less robust under heavy JPEG compression.

4.2.1. Scale-invariant feature transform-based algorithms

SIFT is an algorithm proposed by David Lowe in 2004 [48]. It mainly detects local interest points or feature vectors from an image. Chen *et al.* [42] initially selected SIFT to describe the key points. The key points are then clustered by scale and color and further divided into smaller clusters for separate matching. The J-Linkage algorithm is subsequently used to eliminate mismatches and estimate the affine transformation matrix. Finally, the tampered regions are identified by searching the neighborhoods of the matched pairs for similarities. Although this method improves computational efficiency, its effectiveness in low-texture areas is reduced.

Zedan *et al*. [43] initially validated matching sets obtained from high- and low-contrast keypoints. Next, they selected the three validated matching sets with the smallest distances. The image was then represented by averaging the features of these selected matching sets. This approach of using multiple

matching sets and averaging their features enhances resistance to outliers. Finally, the average distances and distance ratios were used to train an SVM, which determines if CMF has occurred.

Recently, Rathi and Singh [44] combined SIFT with K-means clustering to reduce redundancy in feature representation and speed up detection while maintaining robustness to geometric transformations. Overall, these studies have shown that SIFT-based CMFD methods have improved in efficiency and reliability. However, a common limitation is low accuracy in homogeneous or low-texture regions.

4.2.2. Speeded-up robust features-based algorithms

SURF, proposed by Bay *et al.* [49], is a commonly used feature extraction and matching algorithm. It has high computational speed and maintains good stability under moderate rotation and scaling. Compared with SIFT, it is more computationally efficient. Roy *et al.* [45] first used SURF to detect keypoints in the image. Subsequently, features were extracted using the rotated local binary pattern (RLBP) method, which employs rotational invariance to characterize texture in the image. Next, the extracted features were matched to identify similar keypoints. Finally, a hierarchical clustering method was applied to the matched keypoints to distinguish between real and fake images.

Building on this, Narasimhamurthy *et al.* [46] initially segmented the tested images into different objects using connected components labeling (CCL). Subsequently, hybrid feature descriptors were extracted using the modified SURF (M-SURF) and accelerated KAZE (A-KAZE) algorithms. The A-KAZE descriptor is particularly robust in detecting scale invariance in forged regions and locating poorly targeted keypoints. Finally, the matched features were compared, and false matches were removed using RANSAC. However, the method is computationally expensive and requires careful parameter tuning for large-scale datasets.

Recently, Patilano and Ballera [47] proposed a hybrid feature detection framework that combines an improved binary large object (BLOB) algorithm with SURF and ORB keypoint detectors. This integration enhances feature localization in complex images containing multiple overlapping regions, thereby improving detection performance in scenarios where a single detector may fail. Currently, SURF is often combined with complementary techniques to address its shortcomings in complex textures and to reduce computational overhead.

4.3. Combination-based forgery detection methods

In CMFD, researchers have combined keypoint-based and block-based methods [50]. This hybrid approach combines keypoint detection to capture unique feature points in an image with block-based methods that process localized regions. This combination can better accommodate diverse CMFs, thereby improving detection accuracy and reliability [51]. Additionally, it performs better in complex tampering scenarios. Table 4 summarizes representative combination-based CMFD methods, highlighting their details, benchmark datasets, reported metrics, and key strengths and limitations.

Narayanan and Gopakumar [52] first divide the input image into overlapping regular blocks. Keypoints are then extracted from each block using the SIFT algorithm. Euclidean distance is used to match the keypoints. If the number of keypoint matches exceeds an adaptive threshold, the two blocks are identified as potential candidates. If not, a block-matching algorithm is applied to detect similar blocks. Finally, morphological operations are employed to identify the forged regions. Similarly, Singh *et al.* [53] combined SIFT with the discrete wavelet transform (DWT) to enhance its resistance to geometric transformations and frequency domain distortions. While these methods are effective, they are generally computationally expensive.

Some scholars have combined statistical features with clustering techniques. For example, Harshith *et al.* [54] used the SIFT method to extract keypoints and applied K-means clustering to identify CMF in highly textured regions. To address the challenge of detecting SIFT keypoints in smooth regions, they processed the image using Hu's invariant-based method. This method is robust under scaling and rotation, but is prone to false positives in complex textured backgrounds, which limits its scalability.

The latest trend is transformation-based hybrid methods. Kumari and Garg [55] first segment the image into smooth and textured regions. SURF is then applied to the textured regions, while the fast FMT is used to extract features from the smooth areas. The latter is highly invariant to rotation and scaling, making it suitable for large-scale transformations. Recently, Wang *et al.* [56] first adaptively segmented the input image into non-overlapping blocks using simple linear iterative clustering (SLIC). The image blocks were then divided into textured and smooth regions using the K-multiple-means (KMM) clustering algorithm. Different thresholds were chosen to extract keypoints separately, ensuring an ideal number and distribution of keypoints. They introduced the fast quaternion generic polar complex exponential transform (FQGPCET) and the gray level co-occurrence matrix (GLCM) for feature extraction. This hybrid approach significantly improves accuracy under compression, noise, and geometric attacks, albeit at the expense of increased model complexity and computational overhead.

In general, combination methods improve detection performance by integrating complementary features. However, they often come at the expense of computational efficiency and still require optimization to support large-scale or real-time applications.

Table 4. Summary of representative combination-based CMFD methods

Ref.	Details	Dataset(s)	Metrics (%)	Strengths	Limitations
Narayanan and Gopakumar [52]	Non-overlapping blocks, SIFT, Euclidean distance, and morphological operations.	MICC-F600	R=100, P=93.47, F1=97	Reduces false matches by recursive refinement; effective for small copied regions.	Computationally expensive for large images; sensitive to heavy post-processing.
Harshith *et al.* [54]	K-means clustering, SIFT, Overlapping blocks, Hu's invariant moments, Log-polar transform, and Euclidean distance.	MICC-F600	R=89.37, P=73.71, FPR=10.45, ACC=88.67	Invariant moments provide robustness to rotation; K-means improves the grouping of similar regions.	Limited scalability; performance drops on noisy and compressed images.
Singh *et al.* [53]	Haar wavelet, DWT, SIFT, and RANSAC.	Download from the web	R=100, P=97, ACC=98.12	Robust to scaling and JPEG compression; captures both frequency and spatial features.	High feature dimensionality increases runtime; it is less effective in uniform areas.
Kumari and Garg. [55]	SURF, FMT, PatchMatch, improved g2NN, DLF, RANSAC, and morphological operations.	Download from the web	R=97, P=97, F1=96 (all the best)	Effective against geometric transforms (rotation, scaling); better detection of duplicated patterns.	Sensitive to blur; requires careful parameter selection.
Wang *et al.* [56]	SLIC, KMM, SIFT, FQGPCET, GLCM, DBQ-LSH, DBSCAN, RANSAC, and SSIM.	FAU, GRIP, MICC-F600, and CVIP	TPR=100, FPR=0, F1=96.97; TPR=100, FPR=0, F1=97.89; TPR=98.75, FPR=1.25, F1=94.08; TPR=100, FPR=0, F1=97.06	High robustness to multiple attacks (noise, rotation, compression); accurate segmentation of tampered regions.	Computational complexity is high; parameter tuning is needed for large-scale images.

4.4. Deep learning forgery detection methods

DL has emerged as a core technology for solving complex problems in various fields [4], [57]. In digital image forensics, researchers have applied DL to image CMFD, driving significant technological advances. By leveraging the feature-learning capabilities of DL, models can identify unique patterns and inconsistencies in images, effectively locating tampered regions. This approach considers both local and global contexts, enhancing accuracy and robustness in detecting tampered areas. Table 5 summarizes representative DL CMFD methods, highlighting their details, benchmark datasets, reported metrics, and key strengths and limitations.

DL-based CMFD methods have evolved from sequence feature modeling to segmentation-driven detection and transformer and diffusion paradigms. Elaskily *et al*. [58] proposed a framework based on convolutional long short-term memory (ConvLSTM). They first standardized the input images to the same size and converted them to decimal numbers. They then applied a sequence of ConvLSTM, convolutional CNV, and pooling layers for feature extraction. The classification layer was then used to categorize the feature vectors into predicted categories. Finally, a dense layer was used to evaluate whether the output belonged to the original or falsified class.

In contrast, Nazir *et al*. [59] first extracted features using DenseNet-41 and passed them to the RPN layer, which generates regions of interest. The RoIAlign layer was subsequently introduced to align features associated with these regions, ensuring accurate localization in subsequent processing. Finally, the RoI features were passed to the fully connected network layer to detect and segment the forged pixels. While these methods have achieved significant progress, they rely on ConvLSTM or Mask R-CNN, which limits their generalization ability to unknown operation types.

Table 5. Summary of representative DL CMFD methods

Ref.	Details	Dataset(s)	Metrics (%)	Strengths	Limitations
Elaskily *et al.* [58]	ConvLSTM, CNV, pooling layer, classification layer, and dense layer.	MICC-F220, MICC-F600, MICC-F2000	ACC=100; ACC=98.89; ACC=98.14	Captures spatial–temporal dependencies; robust to noise and geometric transformations.	Limited global context modeling; performance drops on unseen datasets.
Nazir *et al.* [59]	DenseNet-41, RPN layer, RoIAlign layer, and RoI.	CoMoFoD, MICC-F2000, CASIA V2.0	ACC=98.12; ACC=99.02; ACC=83.41	Accurate localization with RoI features; precise segmentation masks.	High computational cost; sensitive to compression artifacts.
Liang *et al.* [60]	TransCMFD: CNN backbone + Transformer encoder–decoder; multi-head attention for global–local integration.	CASIA, CoMoFoD,	F1=93.8; F1=94.5	Strong robustness to rotation, scaling, and compression; high-quality masks.	Data-hungry; requires large-scale training.
Yu *et al.* [61]	CNN encoder + diffusion prior; denoising residuals for detection and localization.	CASIA, CoMoFoD,	F1=95.1; F1=95.5	Leverages generative priors; resilient to compression and re-sampling.	High computational complexity; limited interpretability.
Joshi *et al.* [62]	Feature matching + BBO search; rotation–move validation for tampering mask.	CASIA	P=93.2, R=92.7, F1=92.9	Effective for rotation-based manipulations; reduces false positives.	Sensitive to initialization; scalability issues.
Akram *et al.* [63]	CNN features and handcrafted descriptors: a fusion network for forgery segmentation.	CASIA v2.0, MICC-F220	F1=94.0; F1=94.7	Combines the robustness of DL with the interpretability of traditional features.	Fusion layer complexity; weaker performance on unseen post-processing.

Recently, Transformer-based models have been used to address these limitations. Liang *et al.* [60] proposed an adaptive framework, TransCMFD. They first extract local features using convolutional layers. They then use a Transformer encoder-decoder to capture global dependencies. A multi-head self-attention module fuses global semantics with local details to generate context-aware features. Finally, the decoder generates pixel-level masks of the replicated regions. Compared to CNN-based methods, TransCMFD shows significant improvements, especially under rotation, scaling, and compression artifacts [64], [65].

At the same time, the rise of generative models has also inspired the application of diffusion priors in forgery detection. Yu *et al.* [61] proposed DiffForensics, which incorporates diffusion priors into forgery detection. First, an encoder is used to extract image representations. These representations are then aligned with a diffusion model prior, enabling the system to reconstruct the expected "real" image distribution. The differences between the reconstructed and observed features are computed to identify inconsistencies attributable to tampering. Finally, the model achieves joint detection and localization by mapping these differences to pixel-level masks.

Meanwhile, several hybrid frameworks have emerged, combining DL with optimization or traditional matching strategies. Joshi *et al.* [62] proposed a biogeographically based optimization model specifically for CMF that improved region matching and reduced false positive rates. Furthermore, Akram *et al.* [63] proposed a hybrid CMFD framework that combines deep features with classical forensic methods. They first used convolutional neural networks to extract discriminative features. These features were subsequently fused with handcrafted descriptors (e.g., moment invariants and keypoint features) to enhance robustness. The feature fusion layer then unified the representation and detected tampered regions through classification and localization, ultimately generating an accurate mask. These studies demonstrate that hybrid approaches effectively balance the representational power of deep models with the interpretability and explicit correspondence of traditional methods.

In summary, DL-based CMFD methods have evolved from early ConvLSTM and Mask R-CNN approaches to more advanced Transformers, diffusion models, and hybrid frameworks. This evolution reflects a trend toward the fusion of global and local features and the integration of generative priors with explicit constraints. These methods not only outperform traditional approaches in accuracy and robustness but also offer more interpretable and scalable future solutions for digital image forensics.

This section surveys a range of state-of-the-art CMFD methods, classifying them into four broad categories: block-based, keypoint-based, combination-based, and DL methods. It provides a comprehensive review of each method type.

5. CHALLENGES AND FUTURE SCOPE

The application of digital images is relatively extensive, and image forgery detection is relatively important. CMF is one of the primary passive techniques for detecting forged images. To develop a more accurate, efficient, and non-overly complex image forgery detection method, numerous research studies have been conducted. However, there are still many places that can be improved.

The block-based method can analyze specific local parts of an image because it focuses on small block areas rather than the entire image. When there are nonlinear, complex geometric transformations in CMF, block division can be difficult to cover effectively, resulting in performance degradation. The method based on key points is robust to common geometric transformations, such as rotation, but it struggles to extract key points in regions with a single texture, such as the sky. The combination of block division and key points can improve the accuracy and robustness of image CMFD, and this approach performs well in processing complex forged images, but it also has drawbacks. The integration of the two detection methods will increase the time required for calculation and processing.

Most researchers are now focusing on CMF detection through DL. DL can automatically extract hierarchical features, learn rich semantic expressions, and reduce the matter of manual feature engineering. The framework has flexibility and extensibility. However, DL models need a large amount of labeled data for training, which is difficult to obtain in CMFD, especially the data of real-fake image pairs. Training DL models often requires substantial computing resources, such as high-performance graphics processing units (GPUs) or dedicated tensor processing units (TPUs), making them difficult to use in resource-constrained environments.

6. CONCLUSION

This article reviews the CMFD technology, which is divided into block-based, key-point-based, combined, and DL methods. The analysis shows that the block-based and key-point-based methods are effective in controllability, but have limitations for complex forgeries. DL technology is highly adaptable but requires a large amount of data.

Furthermore, the evaluation of commonly used benchmark datasets reveals limitations in data diversity and realism, which may restrict the generalizability of existing CMFD systems. Overall, this survey underscores the need for more robust, data-efficient, and practically deployable solutions to address real-world forgery scenarios. By synthesizing current methodologies and emerging trends, this work provides a structured foundation and forward-looking perspective for future research in digital image forensics.

FUNDING INFORMATION

This research received no specific research grant or contract from any funding agency in the public, commercial, or not-for-profit sectors.

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
Li-xian Jiao	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓		✓	
Kok-Why Ng		✓				✓	✓			✓		✓	✓	
Hau-Lee Tong		✓				✓	✓			✓		✓	✓	

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

CONFLICT OF INTEREST STATEMENT

The authors state no conflict of interest.

DATA AVAILABILITY

No new data were created or analyzed in this study. The datasets discussed in this survey, including CoMoFoD, CASIA, COVERAGE, and other publicly available benchmarks, can be accessed through their respective original publications and repositories as cited in this article.

REFERENCES

[1] T. Li *et al.*, "RealCam-I2V: Real-World Image-to-Video Generation with Interactive Complex Camera Control," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2025, pp. 28785-28796.

[2] F. Z. Mehrjardi, A. M. Latif, M. S. Zarchi, and R. Sheikhpour, "A survey on deep learning-based image forgery detection," *Pattern Recognition*, vol. 144, Dec. 2023, doi: 10.1016/j.patcog.2023.109778.

[3] S. Singh and R. Kumar, "Image forgery detection: comprehensive review of digital forensics approaches," *Journal of Computational Social Science*, vol. 7, no. 1, pp. 877–915, 2024, doi: 10.1007/s42001-024-00265-8.

[4] A. Pandey, P. Parwekar, and A. K. Chakraverti, "A Review On Deep Learning Techniques For Image Forgery Detection," in *3rd IEEE 2022 International Conference on Computing, Communication, and Intelligent Systems, ICCCIS 2022*, 2022, pp. 890–895, doi: 10.1109/ICCCIS56430.2022.10037746.

[5] N. Chithra Raj, M. Dutta, and J. Saini, "A systematic literature review on image splicing detection and localization using emerging technologies," *Multimedia Tools and Applications*, vol. 84, no. 19, pp. 20721–20756, 2025, doi: 10.1007/s11042-024-19843-z.

[6] A. K. Venugopalan and G. Gopakumar, "Copy-Move Forgery Detection-A Study and the Survey," in *2022 Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICICT)*, IEEE, Aug. 2022, pp. 1327–1334, doi: 10.1109/ICICICT54557.2022.9917647.

[7] M. H. Farhan, K. Shaker, and S. Al-Janabi, "Copy–move forgery detection in digital image forensics: A survey," *Multimedia Tools and Applications*, vol. 83, no. 28, pp. 70603–70635, 2024, doi: 10.1007/s11042-024-18399-2.

[8] M. Verma and D. Singh, "Survey on image copy-move forgery detection," *Multimedia Tools and Applications*, vol. 83, no. 8, pp. 23761–23797, 2023.

[9] G. Kaur, N. Singh, and M. Kumar, "Image forgery techniques: a review," *Artificial Intelligence Review*, vol. 56, no. 2, pp. 1577–1625, 2023, doi: 10.1007/s10462-022-10211-7.

[10] O. M. Al-Qershi and B. E. Khoo, "Evaluation of copy-move forgery detection: datasets and evaluation metrics," *Multimedia Tools and Applications*, vol. 77, no. 24, pp. 31807–31833, 2018, doi: 10.1007/s11042-018-6201-4.

[11] B. Chaitra and P. V. B. Reddy, "Digital image forgery: taxonomy, techniques, and tools–a comprehensive study," *International Journal of System Assurance Engineering and Management*, vol. 14, pp. 18–33, 2023, doi: 10.1007/s13198-022-01829-5.

[12] E. Amiri, A. Mosallanejad, and A. Sheikhahmadi, "Copy-Move forgery detection using EOA, DWT and DCT," *Pamukkale University Journal of Engineering Sciences*, vol. 30, no. 2, pp. 222–227, 2024, doi: 10.5505/pajes.2023.94395.

[13] N. B. A. Warif, M. Y. I. Idris, A. W. A. Wahab, N. S. N. Ismail, and R. Salleh, "A comprehensive evaluation procedure for copy-move forgery detection methods: results from a systematic review," *Multimedia Tools and Applications*, vol. 81, no. 11, pp. 15171–15203, 2022, doi: 10.1007/s11042-022-12010-2.

[14] Y. Aydin, "Comparison of color features on copy-move forgery detection problem using HSV color space," *Australian Journal of Forensic Sciences*, vol. 56, no. 3, pp. 294–310, 2024, doi: 10.1080/00450618.2022.2157046.

[15] S. T. Nabi, M. Kumar, P. Singh, N. Aggarwal, and K. Kumar, "A comprehensive survey of image and video forgery techniques: variants, challenges, and future directions," *Multimedia Systems*, vol. 28, no. 3, pp. 939–992, 2022, doi: 10.1007/s00530-021-00873-8.

[16] M. A. Elaskily, M. M. Dessouky, O. S. Faragallah, and A. Sedik, "A survey on traditional and deep learning copy move forgery detection (CMFD) techniques," *Multimedia Tools and Applications*, vol. 82, no. 22, pp. 34409–34435, 2023, doi: 10.1007/s11042-023-14424-y.

[17] I. Amerini, L. Ballan, R. Caldelli, A. D. Bimbo, and G. Serra, "A SIFT-based forensic method for copy-move attack detection and transformation recovery," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3 PART 2, pp. 1099–1110, 2011, doi: 10.1109/TIFS.2011.2129512.

[18] V. Christlein, C. Riess, J. Jordan, C. Riess, and E. Angelopoulou, "An evaluation of popular copy-move forgery detection approaches," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 6, pp. 1841–1854, 2012, doi: 10.1109/TIFS.2012.2218597.

[19] T. Dijana, I. Zupancic, S. Grgic, and M. Grgic, "CoMoFoD—New database for copy-move forgery detection," in *Proceedings ELMAR-2013*, Zadar, Croatia: IEEE, 2013, doi: 10.1109/ELMAR.2013.6637241.

[20] J. Dong, W. Wang, and T. Tan, "CASIA Image Tampering Detection Evaluation Database," in *2013 IEEE China Summit and International Conference on Signal and Information Processing*, IEEE, Jul. 2013, pp. 422–426, doi: 10.1109/ChinaSIP.2013.6625374.

[21] E. Ardizzone, A. Bruno, and G. Mazzola, "Copy-Move Forgery Detection by Matching Triangles of Keypoints," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 10, pp. 2084–2094, 2015, doi: 10.1109/TIFS.2015.2445742.

[22] D. Cozzolino, G. Poggi, and L. Verdoliva, "Efficient Dense-Field Copy-Move Forgery Detection," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 11, pp. 2284–2297, 2015, doi: 10.1109/TIFS.2015.2455334.

[23] 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.

[24] J. Yang, Z. Liang, Y. Gan, and J. Zhong, "A novel copy-move forgery detection algorithm via two-stage filtering," *Digital Signal Processing*, vol. 113, Jun. 2021, doi: 10.1016/j.dsp.2021.103032.

[25] N. B. A. Warif *et al.*, "Copy-move forgery detection: Survey, challenges and future directions," *Journal of Network and Computer Applications*, vol. 75, pp. 259–278, 2016, doi: 10.1016/j.jnca.2016.09.008.

[26] M. Emam, Q. Han, and X. Niu, "PCET based copy-move forgery detection in images under geometric transforms," *Multimedia Tools and Applications*, vol. 75, no. 18, pp. 11513–11527, 2016, doi: 10.1007/s11042-015-2872-2.

[27] 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, 2018, doi: 10.1007/s00521-016-2663-3.

[28] N. B. A. Warif, M. Y. I. Idris, A. W. A. Wahab, R. Salleh, and A. Ismail, "CMF-iteMS: An automatic threshold selection for detection of copy-move forgery," *Forensic Science International*, vol. 295, pp. 83–99, 2019, doi: 10.1016/j.forsciint.2018.12.004.
[29] N. A. M. Abir, N. B. A. Warif, and N. Zainal, "An automatic enhanced filters with frequency-based copy-move forgery detection for social media images," *Multimedia Tools and Applications*, vol. 83, no. 1, pp. 1513–1538, 2024, doi: 10.1007/s11042-023-15506-7.
[30] D. M. Uliyan, H. A. Jalab, A. W. A. Wahab, P. Shivakumara, and S. Sadeghi, "A novel forged blurred region detection system for image forensic applications," *Expert Systems with Applications*, vol. 64, pp. 1–10, 2016, doi: 10.1016/j.eswa.2016.07.026.
[31] A. Kumar, K. U. Singh, C. Swarup, T. Singh, L. Raja, and A. Kumar, "Detection of Copy-Move Forgery Using Euclidean Distance and Texture Features," *Traitement du Signal*, vol. 39, no. 3, pp. 781–788, 2022, doi: 10.18280/ts.390302.
[32] S. Ganguly, S. Mandal, S. Malakar, and R. Sarkar, "Copy-move forgery detection using local tetra pattern based texture descriptor," *Multimedia Tools and Applications*, vol. 82, no. 13, pp. 19621–19642, 2023, doi: 10.1007/s11042-022-14287-9.
[33] A. Dixit and S. Bag, "Utilization of HOG-SVD based Features with Connected Component Labeling for Multiple Copy-move Image Forgery Detection," in *ISBA 2019 - 5th IEEE International Conference on Identity, Security and Behavior Analysis*, 2019, doi: 10.1109/ISBA.2019.8778494.
[34] Priyanka, G. Singh, and K. Singh, "An improved block based copy-move forgery detection technique," *Multimedia Tools and Applications*, vol. 79, no. 19–20, pp. 13011–13035, 2020, doi: 10.1007/s11042-019-08354-x.
[35] S. Kumar, S. Mukherjee, and A. K. Pal, "An improved reduced feature-based copy-move forgery detection technique," *Multimedia Tools and Applications*, vol. 82, no. 1, pp. 1431–1456, 2023, doi: 10.1007/s11042-022-12391-4.
[36] K. B. Meena and V. Tyagi, "A copy-move image forgery detection technique based on Gaussian-Hermite moments," *Multimedia Tools and Applications*, vol. 78, no. 23, pp. 33505–33526, 2019, doi: 10.1007/s11042-019-08082-2.
[37] T. Mahmood, M. Shah, J. Rashid, T. Saba, M. W. Nisar, and M. Asif, "A passive technique for detecting copy-move forgeries by image feature matching," *Multimedia Tools and Applications*, vol. 79, no. 43–44, pp. 31759–31782, 2020, doi: 10.1007/s11042-020-09655-2.
[38] A. Kashyap, K. D. Tyagi, and V. B. Tyagi, "Robust and Optimized Algorithm for Detection of Copy-Rotate-Move Tempering," *IEEE Access*, vol. 11, pp. 66626–66640, 2023, doi: 10.1109/ACCESS.2023.3291128.
[39] Z. N. Khudhair, F. Mohamed, A. Rehman, T. Saba, and S. A. Bahaj, "Detection of Copy-Move Forgery in Digital Images Using Singular Value Decomposition," *Computers, Materials and Continua*, vol. 74, no. 2, pp. 4135–4147, 2023, doi: 10.32604/cmc.2023.032315.
[40] M. R. Archana, D. N. Biradar, and J. Dayanand, "Image forgery detection in forensic science using optimization based deep learning models," *Multimedia Tools and Applications*, vol. 83, no. 15, pp. 45185–45206, 2024, doi: 10.1007/s11042-023-17316-3.
[41] A. Diwan and A. K. Roy, "CNN-Keypoint Based Two-Stage Hybrid Approach for Copy-Move Forgery Detection," *IEEE Access*, vol. 12, pp. 43809–43826, 2024, doi: 10.1109/ACCESS.2024.3380460.
[42] H. Chen, X. Yang, and Y. Lyu, "Copy-move forgery detection based on keypoint clustering and similar neighborhood search algorithm," *IEEE Access*, vol. 8, pp. 36863–36875, 2020, doi: 10.1109/ACCESS.2020.2974804.
[43] I. A. Zedan, M. M. Soliman, K. M. Elsayed, and H. M. Onsi, "A New Matching Strategy for SIFT Based Copy-Move Forgery Detection," *International Journal of Intelligent Engineering and Systems*, vol. 16, no. 4, pp. 422–435, 2023, doi: 10.22266/ijies2023.0831.34.
[44] K. Rathi and P. Singh, "Copy-move forgery detection by improved SIFT K-means algorithm," *International Journal of Computational Vision and Robotics*, vol. 14, no. 4, pp. 339–354, 2024, doi: 10.1504/IJCVR.2024.139547.
[45] A. Roy, R. Dixit, R. Naskar, and R. S. Chakraborty, "Copy-Move Forgery Detection with Similar But Genuine Objects," *Studies in Computational Intelligence*, vol. 755, pp. 65–77, 2020, doi: 10.1007/978-981-10-7644-2_5.
[46] S. K. Narasimhamurthy, V. K. Mahadevachar, and R. K. T. Narasimhamurthy, "A Copy-Move Image Forgery Detection Using Modified SURF Features and AKAZE Detector," *International Journal of Intelligent Engineering and Systems*, vol. 16, no. 4, pp. 12–24, 2023, doi: 10.22266/ijies2023.0831.02.
[47] H. S. L. Patilano and M. A. Ballera, "Modified Binary Large Object (BLOB) Algorithm Integrated with Speeded Up Robust Feature (SURF) and Oriented FAST and Rotated BRIEF (ORB) Algorithm for Improvement Feature Detection," in *7th International Conference on Sensors, Signal and Image Processing*, 2025, pp. 29–36, doi: 10.1145/3725949.3725960.
[48] Lowe D.G, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
[49] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded up robust features," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 3951, pp. 404–417, 2006, doi: 10.1007/11744023_32.
[50] G. G. Rajput, S. D. Dabhole, and Prashantha, "Modified Keypoint-Based Copy Move Area Detection," *Procedia Computer Science*, vol. 235, pp. 3389–3396, 2024, doi: 10.1016/j.procs.2024.04.319.
[51] J. L. Zhong, Y. F. Gan, and C. F. Zou, "A Keypoint-Based and Block-Based Fusion Method for Image Copy-Move Forgery Detection," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 36, no. 10, 2022, doi: 10.1142/S0218001422540167.
[52] S. S. Narayanan and G. Gopakumar, "Recursive Block Based Keypoint Matching for Copy Move Image Forgery Detection," in *2020 11th International Conference on Computing, Communication and Networking Technologies, ICCCNT 2020*, Kharagpur, India: IEEE, 2020, pp. 1-6, doi: 10.1109/ICCCNT49239.2020.9225658.
[53] R. Singh, S. Verma, S. A. Yadav, and S. Vikram Singh, "Copy-move Forgery Detection using SIFT and DWT detection Techniques," in *Proceedings of 3rd International Conference on Intelligent Engineering and Management, ICIEM 2022*, London, UK: IEEE, 2022, pp. 338–343, doi: 10.1109/ICIEM54221.2022.9853192.
[54] N. B. S. P. S. Harshith, D. Sindhuja, C. Raghava Reddy, A. Deepthi, and G. GopaKumar, "Copy-Move Forgery Detection Using K-Means and Hu's Invariant Moments," *Lecture Notes in Networks and Systems*, vol. 473, pp. 611–619, 2023, doi: 10.1007/978-981-19-2821-5_51.
[55] R. Kumari and H. Garg, "An Image Copy-Move Forgery Detection based on SURF and Fourier-Mellin Transforms," in *2023 International Conference on Artificial Intelligence and Smart Communication, AISC 2023*, Greater Noida, India: IEEE, 2023, pp. 515–519, doi: 10.1109/AISC56616.2023.10085429.
[56] X-y. Wang, X-q. Wang, P-p. Niu, and H-y. Yang, "Accurate and robust image copy-move forgery detection using adaptive keypoints and FQGPCET-GLCM feature," *Multimedia Tools and Applications*, vol. 83, no. 1, pp. 2203–2235, 2024, doi: 10.1007/s11042-023-15499-3.
[57] R. Archana and P. S. E. Jeevaraj, "Deep learning models for digital image processing: a review," *Artificial Intelligence Review*, vol. 57, no. 1, 2024, doi: 10.1007/s10462-023-10631-z.

[58] M. A. Elaskily, M. H. Alkinani, A. Sedik, and M. M. Dessouky, "Deep learning based algorithm (ConvLSTM) for Copy Move Forgery Detection," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 3, pp. 4385–4405, 2021, doi: 10.3233/JIFS-201192.
[59] T. Nazir, M. Nawaz, M. Masood, and A. Javed, "Copy move forgery detection and segmentation using improved mask region-based convolution network (RCNN)," *Applied Soft Computing*, vol. 131, 2022, doi: 10.1016/j.asoc.2022.109778.
[60] E. Liang, K. Zhang, Z. Hua, Y. Li, and X. Jia, "TransCMFD: An adaptive transformer for copy-move forgery detection," *Neurocomputing*, vol. 638, 2025, doi: 10.1016/j.neucom.2025.130110.
[61] Z. Yu, J. Ni, Y. Lin, H. Deng, and B. Li, "DiffForensics: Leveraging Diffusion Prior to Image Forgery Detection and Localization," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA: IEEE, 2024, pp. 12765–12774, doi: 10.1109/CVPR52733.2024.01213.
[62] D. Joshi, A. Kashyap, and P. Arora, "Optimized detection and localization of copy-rotate-move forgeries using biogeography-based optimization algorithm," *Journal of Forensic Sciences*, vol. 70, no. 4, pp. 1392–1413, 2025, doi: 10.1111/1556-4029.70068.
[63] A. Akram, M. A. Jaffar, J. Rashid, K. Mahmood, and A. Ghani, "Advanced digital image forensics: A hybrid framework for copy-move forgery detection in multimedia security," *Journal of Forensic Sciences*, vol. 70, no. 5, pp. 1801–1823, 2025, doi: 10.1111/1556-4029.70076.
[64] S. Li, W. Ma, J. Guo, S. Xu, B. Li, and X. Zhang, "UnionFormer: Unified-Learning Transformer with Multi-View Representation for Image Manipulation Detection and Localization," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA: IEEE, 2024, pp. 12523–12533, doi: 10.1109/CVPR52733.2024.01190.
[65] K. Zeng, R. Cheng, W. Tan, and B. Yan, "MGQFormer: Mask-Guided Query-Based Transformer for Image Manipulation Localization," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2024, pp. 6944–6952, doi: 10.1609/aaai.v38i7.28520.

BIOGRAPHIES OF AUTHORS

Li-xian Jiao is a lecturer in Computer and Big Data, Jining Normal University, China. She received her Bachelor's degree in Engineering from Jining Normal University in China in 2016. Later, she pursued and earned her master of science degree from Liaoning Normal University in China in 2019. She is currently pursuing her Ph.D. at the Faculty of Computing and Informatics, Multimedia University (MMU), Malaysia. Her research interests include information security, digital image forensics, and deep learning. She can be contacted at email: 1211406457@student.mmu.edu.my.

Kok-Why Ng is an associate professor and the Deputy Dean (Student Affairs) at the Faculty of Computing and Informatics, Multimedia University (MMU), Malaysia. He received his bachelor's degree in mathematics from the University of Science, Malaysia, in 2001. He obtained his master's and Ph.D. degrees in information technology from MMU in 2006 and 2017, respectively. His research interests primarily lie in digital image forensics, deep learning, computer vision, recommender systems, and computer graphics. He leads several funded projects and is actively involved in industrial research collaborations. He can be contacted at email: kwng@mmu.edu.my.

Hau-Lee Tong is a senior lecturer at the Faculty of Computing and Informatics, Multimedia University (MMU), Malaysia. He is currently the chairperson of the Industrial Training Committee. His research interests include medical image processing and semantics-based image retrieval. He can be contacted at email: hltong@mmu.edu.my.