Copy Move Forgery Detection Techniques: A Comprehensive Survey of Challenges and Future Directions

Ibrahim A. Zedan[1], Mona M. Soliman[2], Khaled M. Elsayed[3], Hoda M. Onsi[4]
Department of Information Technology
Faculty of Computers and Artificial Intelligence
Cairo University, Giza, Egypt

***Abstract*—Digital Image Forensics is a growing field of image processing that attempts to gain objective proof of the origin and veracity of a visual image. Copy-move forgery detection (CMFD) has currently become an active research topic in the passive/blind image forensics field. There has no doubt that conventional techniques and especially the keypoint based techniques have pushed the CMFD forward in the previous two decades. However, CMFD techniques in general and conventional techniques in particular suffer from several challenges. And thus, increasing approaches are exploiting deep learning for CMFD. In this survey, we cover the conventional and the deep learning based CMFD techniques from a new perspective. We classify the CMFD techniques into several classifications according to the detection methodology, the detection paradigm, and the detection capability. We discuss the challenges facing the CMFD techniques as well as the ways for solving them. In addition, this survey covers the evaluation metrics and datasets commonly utilized for CMFD. Also, we are debating and proposing certain plans for future research. This survey will be helpful for the researchers' as it master the recent trends of CMFD and outline some future research directions.**

Keywords—Image forensics; copy-move forgery detection (CMFD); conventional techniques; deep learning techniques

I. Introduction

Digital image forgery has already showed up in many disturbing forms and results in inestimable lose [1]. Digital image forgery is characterized as changing the original semantic meaning of an image by adding or erasing some significant features of the image for malicious aims [2], [3]. Digital image forgeries can be classified into three classes: Image Retouching, Image Splicing, and Copy-Move Forgery (CMF). Among the image forgery types, CMF is the most common and difficult forgery type. In CMF or image cloning, a part of an image (an authentic source region) is replicated and then pasted to another part of the same image (the forged region) [1], [4] in order to remove unwanted object or replicating desirable object [5]–[7]. Fig. 1 shows two examples of CMF where the cloned regions are marked with red color. The term cloned regions is commonly utilized to refer to the forged region as well as its source region.

To face the massive increase in image forgeries and its harmful effect, Digital image forensics (DIF) becomes an important area of recent research that verifies images reliability. DIF can be classified into passive and active techniques [7], [8]. Active forensic techniques require special hardware and software to embed authentication information such as digital watermark and digital signature into the image before distribution [7], [9]. To overcome such drawback, passive/blind forensic techniques are often used. Passive forensic techniques don't require any prior information about the image to be verified [10], [11]. This survey focuses on the passive forensic techniques proposed for copy-move forgery detection (CMFD) because CMF is a very challenging and popular forgery type. It is hard to differentiate between the actual and tampered images [12]. As the forged region is picked from the image itself, image properties are consistent all over the image and the forged region will be undetectable by methods that look for features inconsistencies [13], [14]. To make detecting CMF more difficult, several geometric and post-processing operations are performed [15].

As shown in Fig. 2, CMFD techniques can be classified according to its detection methodology into: visual similarity based, tampering artifacts based, and hybrid based techniques. Depending on visual similarity aims to specifically detect CMF and isn't able to detect any other forgery type. It can localize the forged region along with its source region based on assessing their similarity. Other forgery detection techniques are based on the fact that image forgery could present tampering artifacts that can be utilized to reveal the forgery. Depending on tampering artifacts is considered a general detection methodology for various forgery types. Applying such methodology for CMFD is only able to localize the forged region without its authentic source region. There are some works that combine the two detection methodologies together. Such works are able to detect and localize the cloned regions and can discriminate the forged region from its source region.

The CMFD techniques can be classified according to its detection paradigm into: conventional techniques, deep learning techniques, and the hybrid techniques. Also, CMFD techniques can be classified according to its detection capability or outcome. The outcome of a CMFD technique could be: (a) classifying an image as original or tampered, (b) localization of the cloned regions at the pixel level if the image is forged, and (c) classifying the cloned regions as source region or forged region [16][17].

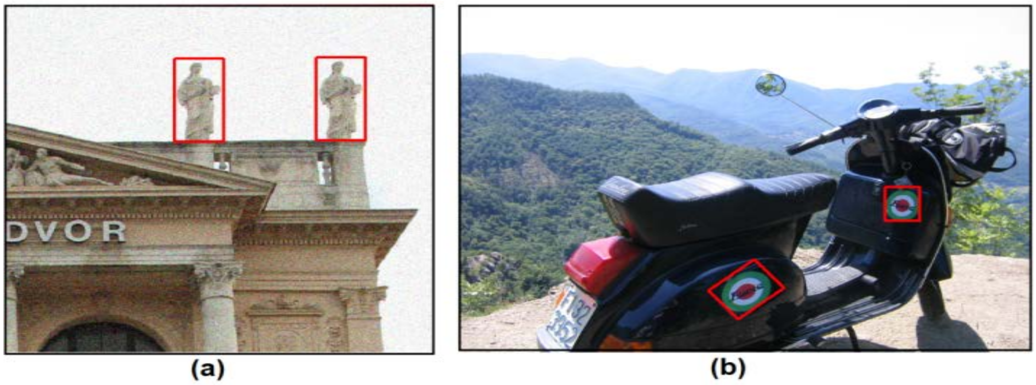

Fig. 1. Examples of CMF. (a) CMF Image Processed by Noise Addition. (b) CMF Image with Ggeometric Ttransforms.

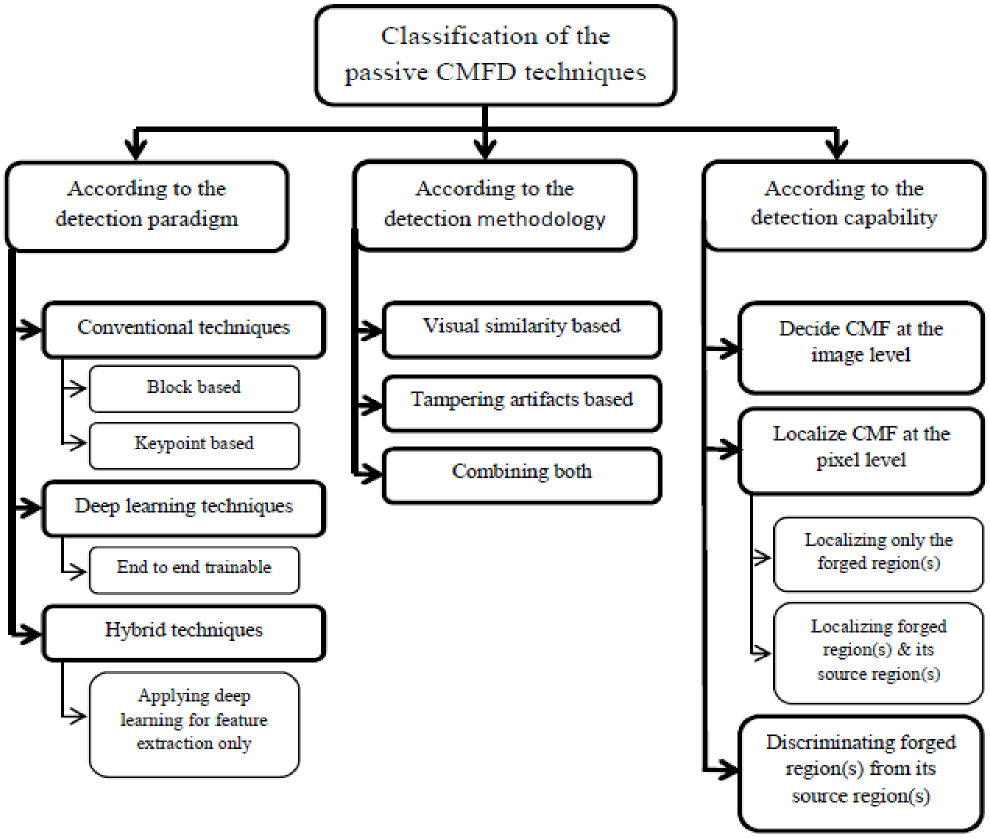
Classification of the passive CMFD techniques
According to the detection paradigm
According to the detection methodology
According to the detection capability
Conventional techniques
Block based
Keypoint based
Deep learning techniques
End to end trainable
Hybrid techniques
Applying deep learning for feature extraction only
Visual similarity based
Tampering artifacts based
Combining both
Decide CMF at the image level
Localize CMF at the pixel level
Localizing only the forged region(s)
Localizing forged region(s) & its source region(s)
Discriminating forged region(s) from its source region(s)

Fig. 2. Classification of the Passive CMFD Techniques.

Surveys such as [18]–[20] are recently proposed to summarize the CMFD techniques. Unlike previous surveys, we cover and organize the conventional techniques and the deep learning techniques for CMFD according to several aspects and with a new perspective. By analyzing the challenges facing CMFD along with the ways to solve them, a reader would be able to know the developing level of this field, and it also can inspire researchers to come up with new perspectives. In addition, we show how the performance of the CMFD techniques highly depends on the utilized dataset and the assessment mechanism. The rest of the survey is organized as follows. Section 2 presents the common procedure of the conventional CMFD techniques. Section 3 analyzes the challenges that face CMFD techniques and the solutions proposed to handle them. Section 4 presents the deep learning based CMFD techniques. Sections 5 and 6 present the standard evaluation metrics and the common datasets being utilized in literature, respectively. Section 7 highlights the important findings of the survey and outlines some future research directions. Finally, Section 8 concludes the survey.

II. Conventional CMFD Techniques

Conventional techniques are mostly adopt the visual similarity based detection methodology [13]. The majority of the conventional CMFD techniques are extract features that represent the image regions and assess the similarity between different regions to reveal the cloned regions [21]. Conventional CMFD techniques can be mainly classified into two categories: block based techniques, and keypoint based techniques [7], [17]. Conventional CMFD techniques have a common detection procedure that divided into three consecutive phases which are: feature extraction, feature matching, and forgery localization [17], [22]–[24].

A. Feature Extraction

It is common to convert RGB images into gray scale and extracting the image features from the intensity channel [22], [25]. Block based CMFD techniques divide the input image into fixed size overlapping square blocks [15], [26]. Numerous descriptors have been utilized in literature to extract the blocks features, e.g., Intensity-based features are utilized in [27], Discrete Cosine Transform (DCT) coefficients in [28]–[31], Principal component analysis (PCA) in [32], Singular Value Decomposition (SVD) [33], sector statistics [34], histogram of orientated gradients [35], Hu moments [36], [37], Local Binary Patterns (LBP) [38]–[41], zernike moments [42], [43], gaussian hermite moments [44], tetrolet features [45], tchebichef moments [46], blur moment invariants [47], polar complex exponential transform (PCET) moments [2], [48], etc.

For keypoint based CMFD techniques, the feature extraction phase consists of two steps: features detection and the description step [13], [17], [22]. Feature detection is to localize a set of keypoints/regions inside an image that are stable for geometric transformation [26]. In the description step, keypoints are described by encoding its surrounding region. SIFT, and SURF are the most popular algorithms utilized in CMFD which are able to perform the features detection and description. On the other hand, Harris corner detector, the maximally stable extremal regions (MSERs), and maximally stable color region are algorithms that only perform the features detection. Utilizing such features detectors requires using other algorithms for the features description.

B. Feature Matching

Image blocks or keypoints with similar descriptors should be matched [17], [23]. The regions of matched pairs are possibly cloned regions [22]. One way to match image features is to apply a global threshold on the distance between descriptors as in [29], [31], [33], [35], [43], [45]. Two blocks/keypoints are matched if the distance between their feature vectors is smaller than a threshold. This threshold can fluctuate from zero to one. A threshold closer to zero yields fewer but more accurate matches [13]. However, this matching method obtains a low accuracy [24], [26]. So, the two nearest neighbor (2NN) test is a widely utilized matching method in keypoint based CMFD techniques as in [10], [13], [17], [21]–[23], [49]. In the 2NN test, if the distance ratio between the nearest distance to the second nearest distance is less than a threshold, then the two keypoints are matched [26].

The 2NN test works well when a region is duplicated one time. As a result, the generalized nearest neighbor (g2NN) matching method is proposed to work when the region is cloned several times. The g2NN matching method iterates the 2NN test until the distance ratio become greater than the specified threshold [4], [5], [9], [14], [25], [26]. However, some matched keypoints still can't be recognized by the g2NN matching method. Accordingly, the transitive matching is utilized in [50] to enhance the matching relationship.

C. Forgery Localization

In the forgery localization phase, the geometric transform between the matched pairs is usually modeled. Such modeling is helpful to eliminate any mismatched pairs as cloned regions are commonly localized when detecting certain number of matched pairs with the same geometric transform [44]. Then, a forgery decision process is performed at the image level to decide if the given image is tampered with CMF or not. Finally, a localization process at the pixel level is performed to locate the cloned regions within tampered images.

In [30]–[32] the geometric transform between matched pairs is modeled by the shift vector between their coordinates. But, the shift vector concept is not suitable in case of rotation or scaling [44]. So, the geometric transform between the replicated regions is usually modeled by an affine homography [23]. Random Sample Consensus (RANSAC) is a widely used technique for accurate estimation of the affine homography that leads to the minimum error even when high number of mismatched pairs are exist [4]. So, RANSAC is utilized to estimate the affine transform and to filter some mismatched pairs in [51]–[53]. In [10] the Helmert transformation is utilized instead of the affine transformation because of its low degree of freedom and low computational complexity.

Several decision rules are utilized in literature to decide at the image level whether an image is tampered with CMF or not. In [27] The task of cloning detection is that of detecting two large similar regions bigger than an area threshold corresponds to certain percentage of the image size. The area threshold determines the smallest size of the cloned region to be detected. Large area threshold increase the misses while small values increase false alarms [28]. In [42] the forgery decision is specified if there are more than a particular number of matched pairs that meet the estimated affine transform and the similar regions are bigger than an area threshold.

Some works such as [34], [54]–[56] are just localize the cloned regions by depicting the matched pairs as lines. While other works reveal the cloned regions at the pixel level [24], [57]. Block based CMFD techniques are commonly localize CMF at the pixel level by relatively simple steps. First, a black image is created with the same size as the suspected image. Then, the blocks correspond to the matched pairs are simply assigned other color [30], [46], [47]. But, keypoint based CMFD techniques don’t have good localization power of cloned regions [10] as matched keypoints don’t cover completely the cloned regions [52]. To solve this issue, Cloned regions are commonly localized by the following steps [17]. First, the transformation matrices between the cloned regions are estimated. Then, all image pixels are transformed forward and backward. Next, the similarity between the original image and the transformed images is assessed with the correlation coefficients that are invariant to illumination changes. After that, correlation maps are smoothed to reduce the noise and transformed to binary images with a fixed correlation threshold [1], [26], [58], [59]. Finally, small isolated regions are eliminated and holes are filled [13], [16], [24], [36]–[38], [41], [60], [61]. Such final post processing can be accomplished by morphological operations, specially designed filter as in [39], or an area threshold as in [37], [62].

In [7], [8], [10], [49], [63] the cloned regions are localized by different methods rather than the common work flow described previously. In [63] to localize the cloned regions; image registration through bi-cubic interpolation is utilized. In [8] the cloned regions are localized by region growing technique through Hu's moments. In [49] cloned regions are localized based on multi-scale analysis and a voting process. Some works such as [7], [10] utilized the superpixels segmentation algorithms to localize the cloned regions.

III. Formulation of Challenges

This section analyzes the challenges that face the CMFD techniques in general and provides a comprehensive survey on different strategies adopted by the conventional CMFD techniques to handle these challenges. Table I shows the phases of the conventional CMFD techniques and the challenges that face each phase along with the solutions.

A. Geometric Transforms

Geometric transforms such as scaling, and rotation are usually applied to the forged regions to fit the scene and to mislead the human eye. Scaling or rotating an image region introduces some changes in the pixels values due to the interpolation error. Block based techniques fail when large rotation and scale are operated on the cloned regions [26] because of the de-synchronization in searching of matched blocks [60]. Dividing the image into overlapping square blocks with static size isn't able to detect CMF with large scaling and rotation regardless of whether the utilized features are scaling and rotation invariant. Utilizing circular blocks solves to some extent CMF with rotation. To handle CMF with scaling, adopting a pyramid model and performing the matching process across several scales are needed [42].

The majority of the keypoint based techniques are robust against geometric transformations, including large rotation and scale. But when utilizing a keypoints detector that wasn't robust by nature to certain geometric transforms, it is essential to make it invariant to geometric transforms such as [64], [65]. In [64], [65] to make Harris corners invariant to scaling, stable points across a scale space are only identified.

Conventional CMFD techniques such as [6], [13], [34], [62] tried to enhance its robustness against reflection because reflection needs special handling. In [34], [62] To detect CMF with flipping, a matching process between the feature vectors of the original image and the flipped image is performed. In [6], [13] a flip invariant SIFT descriptor is utilized in which each image keypoint is represented by two descriptors that reorganize the SIFT descriptor with both clockwise order and anticlockwise order. Among all the geometric transforms, deformation affect greatly the performance of the conventional CMFD techniques as it is a nonlinear transformation that can't be modeled well by an affine model [13]. Dealing with nonlinear geometric transformations still needs to be explored.

TABLE I. Conventional CMFD Techniques: Phases, Challenges and Proposed Solutions

Phases	Challenges	Solutions
Feature Extraction	Geometric Transforms	Utilizing invariant features
		Multi scale analysis & matching
		Utilizing circular blocks
	Post Processing Operations	Performing image enhancement before extracting features
	Dealing with Small or Smooth Cloned Regions	Increasing the image contrast and resolution
		Utilizing hybrid keypoints detectors
		Lowering the contrast threshold
		Adopting small block size
		Combining keypoint based and block based techniques
Feature Matching	Image Continuity	Avoid matching of neighboring features
	Handling Image Self-Similarity and Similar But Genuine Objects	Enhancing discrimination power of the descriptors
		Eliminating outlier matches
		Accurate estimation of the thresholds
		Accurate estimation and validation of the geometric transformations
	The Matching High Computational Complexity	Decreasing the image dimension and number of features
		Utilizing low dimensional and binary descriptors
		Sorting and organizing the image features before matching
		Matching search space reduction
		Searching for approximate matching
	Un Consistent Matching Order	Utilizing clustering or segmentation based algorithms
Forgery Localization	Dealing With Multiple Cloned Regions	Performing clustering of the matched pairs
		Performing iterative localization
	Discriminating Forged Region from its Source Region	Utilizing hybrid detection methodology

B. Post Processing Operations

Post processing operations are generally applied to make the detection of CMF harder to detect. The most utilized post processing operations are JPEG compression, image blurring, and noise addition [1]. CMFD techniques achieve better detection accuracy when the intensity of the post processing operation is minimal [16]. When handling low quality images or images with high noise, the performance of the CMFD algorithms is decreased because the pixel values are disturbed that result in less number of correct matched pairs, more false positives, and more false negatives [1], [34], [66].

Numerous conventional CMFD techniques tried to face image post-processing attacks. In [38], [39], [60] image is filtered by Gaussian low pass filter because the low frequencies are more steady to post processing operations. In [42] each image block is filtered by an adaptive wiener filter which can remove noise while preserving edges. In [67] the input image is enhanced before extracting the SIFT features. First, high pass filter (HPF) is applied. Then, Butterworth low pass filter (BLPF) is utilized for noise reduction. In [16], [29], [41] the stationary wavelet transform (SWT) is utilized to reduce the noise effect. Image blurring effect specially the performance of the keypoint based CMFD techniques [13] as a lot of keypoints are lost due to blurring. Dealing with blurring is similar to dealing with smooth or small cloned regions. Several solutions are reviewed in the next sub-section.

C. Dealing with Small or Smooth Cloned Regions

With small cloned regions, the performance of the CMFD techniques is low because of insufficient number of correct matched pairs [34]. Block-based CMFD techniques usually adopt small block size for revealing small cloned regions. But, small block size can’t yield robust features [3] and results in large number of blocks that increases the computational cost. On the other hand, large block size decreases the computational cost but can’t detect small cloned regions [3], [44]. So, block-based CMFD techniques face a difficulty in selecting the suitable block size [3]. It is worth noting that block based CMFD techniques work well in smooth regions.

Keypoint based CMFD techniques fail to detect the forgery if insufficient keypoints are identified that results in insufficient correct matched pairs, and that is the situation when dealing with smooth or small cloned regions or when the input image is of low resolution [4], [68]. One way to extract more keypoints is to utilize hybrid/multiple detectors such as in [5], [9], [26]. Other works such as [15], [25] applied the keypoints detectors on the opponent color space rather than the intensity channel to get an adequate number of keypoints.

Image keypoints are generally detected by applying certain contrast threshold [58]. Several works increase the keypoints in the whole image by simply lowering the contrast threshold of all images under investigation such as [9], [24]. As the suitable contrast threshold could varies from one image to another, other works tried to choose the suitable contrast threshold separately for each test image. In [17], [22], [23], [55], [69] particle swarm optimization (PSO) algorithm is utilized to generate customized parameter values for each image. Several works such as [1], [4], [56], [66], [70], [71] increase the entire image contrast or resolution instead of decreasing the contrast threshold. In [70] single image super resolution algorithm is utilized to increase the image resolution. Similarly in [4] the image is up-sampled. In [1], [56], [66] the contrast limited adaptive histogram equalization algorithm is utilized to increase the image contrast. Similarly in [71] the dynamic histogram equalization method is utilized.

Increasing the keypoints in the whole image by adopting a small contrast threshold has several drawbacks. Keypoints in the rough regions will increase quicker than in smooth regions [68] which is pointless. This phenomenon is called the non-uniform distribution of the image keypoints [58]. Also adopting a small contrast threshold will trigger numerous unstable keypoints, and expand false matching possibility. More redundant keypoints are located at nearby locations and its corresponding descriptors are similar [24].

Several works such as [12], [52], [54], [58], [68] aim to overcome the non-uniform distribution of the image keypoints. In [58] the non-maximum value suppression algorithm is utilized. First, all possible keypoints are initially selected. Then, redundant keypoints are filtered out. In [68] the image is segmented into smooth and rough layers. Swarm intelligence algorithm is applied for each layer separately to find its customized parameter values. In [12], [52], [54] image is segmented into non-overlapping superpixels. The way of localizing the keypoints varies from smooth regions to texture regions to make keypoints uniformly covering the entire image. Other works such as [5], [71] process specific regions within the image to extract more keypoints and more matched pairs. In [5] any suspicious region that contains insufficient number of matched keypoints is up-sampled and re-examined. In [71] matched keypoints are grouped into regions. The obtained regions are scaled up instead of scaling up the entire image.

As block based CMFD techniques work well in case of smooth regions, a combination of keypoint based and block based methods is proposed for effective CMFD as in [11], [61], [72]. In [11], [61] SIFT based method is utilized to detect forgery in rough regions. To detect forgery in smooth regions, Zernike moments are utilized in [61] while the Fourier Mellin transform (FMT) is utilized in [11]. In [72] to handle cloned regions with insufficient number of matched pairs, two regions centered on the keypoints location of each matched pair are obtained. These regions are examined by Zernike moments.

D. Image Continuity

Because of the continuity of images, the similarities of neighboring blocks/keypoints are high and hence are wrongly matched. Also in block based CMFD techniques, the image is usually divided into overlapping blocks. Blocks with an overlapping ratio are highly similar and wrongly matched. So, in [2], [28], [30], [32], [42], [47], [53], [60] matched pairs are removed if their spatial separation is below a threshold. The spatial separation threshold defines the smallest spatial distance between the cloned regions to be detected [47]. The choice of the spatial separation threshold should consider its relationship with the image content and size [61].

Other solution avoids the selection of the spatial separation threshold by segmenting the image into non-overlapping superpixels and requires that two image features are comparable if they are belonging to different superpixels [4], [58], [61], [73]. But, this solution isn't able to detect CMF in case of two cloned regions are in the same superpixel [73]. Similarly in [3] the image blobs are detected utilizing DoG and BRISK keypoints in different blobs are only matched.

E. Handling Image Self-Similarity and Similar But Genuine Objects

The intrinsic self-similarity of natural images is considered the other reason of wrong matching in addition to the image continuity [17]. In addition, images might have similar but genuine objects (SGO). CMFD techniques which are based on a simple hypothesis that similar regions in an image are often made by CMF are commonly produce false positives in images having SGO [64], [65], [74]. The ability to distinguish cloned regions from SGO is essential for a successful CMFD technique [64], [65]. In the next paragraphs, we discuss how the selected features, the matching method, and the choice of the thresholds can play a vital role to deal with image self-similarity and to distinguish cloned regions from SGO.

The majority of the conventional CMFD techniques extract its features from gray scale images. But, some CMFD techniques such as [58], [75] perform the feature extraction phase in a certain color space to enhance its discrimination power and hence its performance. In [75] each color channel is considered separately. Matched blocks that are common in all color channels are considered as forged. In [58] OpponentSIFT is utilized for feature extraction. OpponentSIFT describes all the channels of the opponent color space utilizing SIFT.

Texture descriptors are useful to differentiate between really cloned regions and SGO. Also, high dimensional descriptors are generally more distinctive. As a result, several works such as [52], [53], [57], [64], [65], [76] enhance its discrimination power and hence its matching performance by utilizing texture features or utilizing high dimensional descriptors. In [53] Image blocks are described by multiple LBP operators. In [76] two regions are verified as cloned if their GLCM contrast difference is below a threshold. In [57] SIFT descriptor is combined with the histogram of the reduced LBP. In [64], [65] LBP as well as DCT and SVD are utilized to describe the detected Harris keypoints. In [52] PCET is utilized to extract descriptors of the detected SURF keypoints.

Cloned regions are commonly chosen from meaningful objects [26]. As a result, false matched pairs of intrinsic self-similar regions are usually isolated and much more scattered. Based on such idea, several conventional CMFD techniques such as [47], [77]–[79] reduce the false matching rate basically by its matching process. In [47] two blocks are matched if its neighboring blocks are also matched to each other. The number of neighboring blocks to be checked for similarity defines the smallest size of the cloned region to be detected. In [77] a match is decided when the keypoints from an image and its k nearest neighbors are matched to that of the suspicious area. In [78] the matching process is done among objects rather than single point matching. In [79] clusters of keypoints are matched instead of single point matching.

Other way to reduce the false matching rate is to adopt an outlier removal process of wrong matches after performing the matching process as in [45], [66]. In [66] the outlier matches are eliminated by combining the guaranteed outlier removal algorithm with the RANSAC algorithm. In [45] Fast outliers filtering method is utilized instead of RANSAC. But, such few outlier matches might correspond to a CMF with small cloned regions and should be further verified.

Several conventional CMFD techniques utilize the segmentation or the clustering methods to eliminate false matches [24]. The regions/clusters that contain a few matched pairs are discarded [14], [26], [50], [59]. The segmentation and the clustering based algorithms suffer from high time and space complexity [66]. Also it is hard to decide a segmentation or clustering algorithm and its associated parameters that are suitable for all images [24]. The superpixel segmentation algorithms are commonly utilized. The initial superpixel size has significant impact on the forgery detection performance. An appropriate initial superpixel size should consider the image size and content. In case of textured images, an initial superpixel size of low value should be utilized. While, a high value should be adopted as an initial superpixel size when dealing with simple images [54]. Many works have taken into account the image size and content when selecting the initial superpixel size. But, no one has dealt with the fact that a single image could contain both a smooth part and a texture part and they should be segmented differently.

Clustering based algorithms such as [10], [13], [49] aim to filter out false matches by adopting the geometric inconsistency idea. In [13] the slope of all lines connecting matched pairs is grouped in different clusters. Within each cluster, outlier matched pairs are removed if its locations are far from the cluster centroid location. In [10], [49] matched pairs are grouped into clusters dependent on the spatial separation among them and the angle of the line that connects them relative to the x-axis. Furthermore, in [10] the Helmert transformation is utilized to merge clusters with similar transformation parameters. Despite the fact that the hierarchical agglomerative clustering (HAC) is the common clustering algorithm utilized in CMFD, it is sensitive to outliers and noise. So, in [21], [51], [66], [71] the DBSCAN (density-based spatial clustering of applications with noise) is utilized.

The thresholds related to the matching process acquire special significance in handling image self-similarity and SGO. To decide an appropriate value of the matching threshold, a training phase is needed. But, the matching threshold may change from one image to another [53]. So, in [2] an appropriate matching threshold is estimated for each image utilizing PSO and the histogram of block similarities. After localizing suspected regions within an image, it is common to compute the correlation coefficient between the suspected regions. Then, a correlation threshold is utilized to differentiate really cloned regions from SGO. High value of the correlation threshold increases the misses' rate while a low value increases the false alarm rate. Many works such as [17] have focused on selecting an appropriate value of the correlation threshold. In [17] customized correlation threshold is utilized to detect each image rather than utilizing a fixed threshold for all images.

It is essential to assess the accuracy of the estimated geometric transformation [5] as inaccurate estimation of the geometric transformation results in wrong localization of cloned regions. So, many CMFD techniques have focused on the accurate estimation and validation of the geometric transformation such as [5], [13], [24]. In [24] a homography validation and inlier selection technique is proposed. For each correctly matched keypoints, the difference of the dominant orientations should be consistent with the estimated homography. In [5] inaccurate affine transformations are filtered by utilizing the Bag of Word idea. In [13] the affine transformation parameters are refined iteratively.

F. The Matching High Computational Complexity

Feature matching is the main phase that consumes time [2], [60] because of the huge number of image features and their high dimensional descriptors [80], [81]. Keypoint based techniques have a lower computation cost compared to the block based techniques because the number of keypoints for an image is generally smaller than the number of blocks [26], [82]. However, several keypoint based CMFD techniques try to increase the number of keypoints inside an image to handle small or smooth cloned regions. In this case, the computational complexity also needs to be reduced.

One way to reduce the matching time is to decrease the number of image features to be extracted as in [15], [48]. In [48] the features are computed for only the fundamental objects rather than all the overlapping blocks of the image. In [15] Image is divided into MSERs. SIFT keypoints that aren’t belong to any MSER are excluded to reduce the matching cost. Decreasing image dimension results in a reduction of the number of features. In [42] high resolution images are scaled down. In [27], [37] the image is decreased in dimension by Gaussian pyramid. In [21], [32], [36], [51], [53] the wavelet transform is utilized. However, decreasing the number of image features reduces the performance of the CMFD because the high details in the image have been lost [21].

Low dimensional descriptors and binary descriptors are more desired for fast matching. As a result, several works such as [2], [21], [28], [46], [47] tried to decrease the matching time by reducing the descriptor length through SVD or PCA. As SURF descriptor has low dimensional space compared to SIFT, so matching SURF descriptors is fast [81]. Also, binary descriptors are favored for fast matching as they are matched quickly by simple XOR operation through the hamming distance [83]. As a result, the BRISK binary descriptors are utilized in [83]. Similarly in [84] The SIFT descriptors are binarized and matched to reduce the matching time.

Several works tried to reduce the matching search space and decrease the number of comparisons needed by means of segmentation or clustering such as [3], [52], [80]. In [52] image regions are separated into texture regions and smooth regions. The image features are matched separately in smooth regions and in texture regions. In [3] the image background is eliminated prior to matching image features to speed up the matching process. In [80] image keypoints are grouped into clusters using the Fuzzy C means clustering technique. Each cluster center and its close keypoint are matched only to other clusters instead of matching all the image keypoints.

To reduce the matching time, it is common to sort or organize the image features before matching [43]. For block based CMFD techniques, Lexicographic sort is a widely utilized sorting method that makes comparable feature vectors closer to each other. A feature vector will be checked for similarity with just a specific number of neighboring vectors [40]. For computational efficiency, some conventional CMFD techniques such as [7], [60] utilized approximate matching instead of exact matching. In [1], [55], [62], [66], [81] the best bin first search algorithm is utilized which is based on a variant of the KD tree to search for approximate nearest neighbor. It is common to use the Euclidian distance to calculate the distance between image descriptors. But for computational efficiency, the cosine similarity is utilized in [25], [73], [79].

G. Inconsistent Matching Order

It is common to use RANSAC to estimate the geometric transform from the authentic source region to its forged region or vice versa. The estimation of the geometric transformation is order-dependent. If the geometric transform estimated from the source region to its forged region is T, then the geometric transform estimated from the forged region to its source region is T^{-1}. As a result, the matched pairs fed into RANSAC should have consistent matching order; otherwise they could result in erroneous estimation [24]. But in keypoint based CMFD techniques, keypoints are detected from the image without any spatial order. So, the matching process can't guarantee a consistent matching direction [24]. To solve this problem, the segmentation and the clustering based algorithms are utilized to facilitate a consistent matching direction from one region/cluster to the other region/cluster [24]. On the other hand, block based CMFD techniques aren’t suffered from this problem at all.

H. Dealing with Multiple Cloned Regions

Some conventional CMFD techniques such as [22], [23] aren’t able to handle images with multiple cloning. To handle images with multiple cloning, two issues should be considered. First, the adopted matching method should able to perform multiple matching if exist of the same block/keypoint. As mentioned before, matching methods such as G2NN and transitive matching are able to perform multiple matching. Second, the matched pairs may follow diverse geometric transformations in case of multiple cloned regions [24].

Multiple cloning is commonly handled through clustering the matched pairs and iterating the localization task [24], [26]. Clustering of matched pairs aims to group pairs that follow the same affine transformation [58], [70], [81]. In [24] the localization task runs in an iterative manner. In each iteration, RANSAC algorithm is utilized to estimate one affine homography using all the matched pairs from two contiguous local patches. In [26] the RANSAC algorithm is executed iteratively to estimate the transformation matrices. After each iteration, the inliers satisfying the previously estimated transformation are excluded from the next iteration.

Clustering based algorithms such as [1], [14], [61], [70], [81] clustered the matched pairs by their location utilizing HAC. Especially in keypoint based CMFD techniques, clustering of the matched pairs based on their location has two drawbacks: (i) the inability to separate the cloned region when cloned region is close to its source region and (ii) the difficulty to identify the cloned region as a single region, when it contains scattered keypoints [59]. To handle these drawbacks, in [58], [59] matched pairs are clustered using the J-Linkage algorithm based on a transformation domain rather than the spatial domain. In J-Linkage clustering, a number of affine transformation hypotheses are generated randomly. Each matched pair is assigned to an initial cluster. HAC process is operated on the clusters. To reduce the computational cost of J-Linkage clustering, image is segmented into superpixels in [58]. Then, the matched pairs are grouped based on the correspondence between the superpixels to produce a small number of initial clusters of the J-Linkage algorithm [58].

I. Discriminating Forged Region from its Source Region

The majority of the CMFD techniques lack the ability to classify the cloned regions into source and forged regions. Providing this capability enriches the CMFD technique. Depending on visual similarity to reveal the cloned region can't discriminate between the forged region and its source region. To provide such ability, detection of tampering artifacts should be integrated with the visual similarity based detection methodology as in [85]. In [85] a resampling based method is combined with SIFT based CMFD technique. The resampling based method takes as input the matched pairs that highlight any cloned regions. If the resampling factor of a certain region is different from its neighborhood, it is considered as forged region. Otherwise, it is considered as source region. The resampling based method fails to classify the cloned regions into source and forged regions if the forged region hasn't been modified geometrically.

IV. Deep Learning based CMFD Techniques

Conventional CMFD techniques with handcrafted features experience three limitations [86]. First, these techniques have an identical structure comprises of three phases. Each phase is trained separately and has numerous parameters that are manually tuned [86]–[88]. Second, these techniques are mostly tuned to accomplish great performance on specific dataset(s) yet fail in other datasets [86]. Third, handcrafted features are usually have restricted discrimination power [89], [90].

Deep learning exhibits a major achievement in various recognition tasks. As a result, numerous researchers attempt to adopt deep learning for CMFD [91]. Selection of the appropriate parameters/thresholds is the most important problem that conventional CMFD techniques faced, but deep learning models are able to automatically learn the suitable features for CMFD [87], [92]. To build a CMFD system using deep learning, a large number of training samples is required. But, existing CMF image datasets are limited in size [93]. One way to handle this problem is adopting transfer learning or utilizing deep learning methods to only extract the image features within a block or keypoint based CMFD techniques.

Transfer learning utilizes a pre-trained model for certain task and slightly re-train the model parameters with little training samples for other task [93]. These pre-trained models such as AlexNet, VGGNet, GoogLeNet, and ResNet are pre-trained from enormous training dataset and have a powerful generalization power [93]. Several works such as [92]–[94] utilized AlexNet for CMFD because of its simple construction. In [92][93] Alex Net is utilized for CMF classification at the image level. The proposed method in [93] fails to handle realistic CMF because its model is trained with simple CMF samples. In [92] the SVM classifier is trained with the features obtained from the fully connected layer of Alex Net. In [94] a modified AlexNet architecture is proposed for classifying various forgery types at the image level.

Several CMFD techniques such as [89], [95] utilized deep learning methods to only extract the image features within a block or keypoint based CMFD techniques. In [95] A block based CMFD technique in which Alex Net is utilized to describe the image blocks. In [89] GPU-based convolutional kernel network (CKN) is utilized to obtain local descriptors of the image keypoints. CKN is a deep convolutional network that combines neural networks with kernel methods. CKN aims to produce local descriptors that are invariant to various transformations. Also, in [89] the image is adaptively over-segmented into superpixels utilizing an efficient CNN based technique which is called the convolutional oriented boundaries (COB).

Adopting transfer learning or utilizing deep learning methods to only extract the image features have some drawbacks. CMFD techniques which utilize transfer learning are commonly deciding the forgery at the image level only. Also, CMFD techniques which utilize deep learning methods to only extract the image features aren't end to end trainable. So, several synthesized datasets that incorporates an enormous number of images with its localization binary masks are proposed to manage an end to end training process of the CMF localization task [90], [96]. In the next subsections, several deep learning based CMFD techniques are reviewed and organized according to its detection methodology. All the presented techniques are end to end deep learning based systems which are trained using huge synthesized datasets.

A. Visual Similarity Based

Deep learning based CMFD techniques which reveal the CMF on the basis of visual similarity are commonly mimic the same phases of the conventional CMFD techniques. However, each phase is accomplished by deep neural network (DNN) layers. A customized DNN layers are utilized to perform the feature matching phase. As a result of relying on visual similarity to locate CMF, similar but genuine regions may be treated as forged regions by mistake [87], [97].

Deep learning models such as [86]–[88], [97], [98] aim to localize CMF on the basis of visual similarity. In [87], [88], [97], [98] feature extraction is performed through the VGG16 architecture. BusterNet [87], [88] is considered the first end-to-end DNN model that aims to localize CMF at the pixel level. Other works such as [97], [98] aim to enhance BusterNet. For feature extraction, BusterNet utilized 4 blocks of the VGG16 network along with four pooling layers while in [98] the 4th pooling layer is removed to obtain features with higher resolution. Additionally, atrous convolution is utilized in [97], [98] to increase the filters field of views. As contextual information isn't captured well in BusterNet, the attention module is utilized in [97], [98] to capture contextual information, and enrich features. For feature matching, BusterNet performed single level feature matching while in [97], [98] hierarchical feature matching is enabled by considering features of multiple levels. To produce the CMF localization mask, BusterNet utilized a deconvolutional network which incorporates bilinear up-sampling layers and inception modules. Atrous spatial pyramid pooling (ASPP) is utilized in [97], [98] to localize CMF with scaling operation by exploiting image features in multiple scales. The CMF localization mask is refined in [97] using a residual refinement network. In [86] three dense inception blocks are combined to extract multi scale highly rich features. Three matching maps are obtained to yield a coarser to fine feature matching and then they are integrated by the loss layer to localize the CMF.

B. *Tampering Artifacts Based*

Several deep learning based CMFD techniques depend on the detection of tampering artifacts to reveal forgery in general. One example of such tampering traces is the unnatural characteristics that may appear at the forged regions boundary [91], [96]. To achieve visual consistency in a forged image, it is common to edit the forged region by various operations. The presence of multiple editing operations inside an image is also considered as a forgery indicator [99]. Additionally, geometric transformations and interpolation are commonly applied to the forged regions which results in periodic correlation artifacts. The resampling features can be utilized to catch such periodic correlation in the frequency domain [96].

Utilizing deep neural networks for detecting tampering artifacts is a difficult task. Deep neural networks (DNNs) are usually provides feature maps that describe image content rather than the forgery traces. So, utilizing DNNs for forgery detection needs some sort of adaptation to learn richer features correspond to the tampering artifacts [93], [99]. To be specific, the training of DNNs for the forgery detection purpose should be based on an information that describe the local relationship between adjacent pixels [99]. Such information could be captured through fixed spatial rich model (SRM) filters as in [100], [101], or a constrained convolutional layer that learns the filters weights as in [98], [99].

Realistic forged images show high similarity between its authentic and forged regions. So, depending only on CNN based architecture or single feature type for forgery localization isn’t enough. In [91], [96] a hybrid model consists of LSTM network and CNN is utilized to localize three image forgery types: copy-move, splicing, and removal. As LSTM network is able to handle sequential and contextual information, in [91], [96] LSTM is utilized to learn the transition between the authentic and forged regions. In [91] the proposed model comprises of an LSTM network and 5 convolutional layers. Image patches are gone through the first 2 convolutional layers to output a low-level feature map which is divided into blocks. Then, these blocks are gone through the LSTM network. The later 3 convolutional layers will get the LSTM feature map and produce the forgery localization mask. In [96] the proposed model comprises of an LSTM network and encoder-decoder network. Image is divided into blocks which are described by the resampling features. The resampling features go through the LSTM network. The encoder consists of residual units accepts the whole image as input and produce the spatial feature maps with global context. The encoder features and the LSTM features are fused and taken as input to the decoder to produce the forgery localization mask.

CNN networks are usually aim to classify an image into one of various classes by learning class-specific features. On the other hand, the Siamese network aims to discriminate various classes by learning more generic features along with a distance metric. The Siamese network comprises of twin sub-networks processing two images in parallel to decide whether the two images are similar or not [99]. In [99] a deep Siamese network is utilized to detect several types of image level post-processing operations that are usually aim to hide the forgery traces. Moreover, a forgery localization method is proposed in [99] by dividing an image into overlapping regions which are compared with each other through the Siamese network to decide whether the image regions are similarly processed or not. Image is considered as forged if its regions have different processing operations.

Several works such as [90], [100], [101] adapted object detection or segmentation networks to localize three image forgery types: copy-move, splicing, and removal. In [100] Faster R-CNN network with two parallel streams: RGB stream, and noise stream is proposed. The RGB stream models the global visual tampering artifacts. SRM filters are applied to the image to extract local noise features which go through the noise stream to figure out any noise inconsistency. A bilinear pooling layer is utilized to fuse the two streams features and enrich the network training. Object detection networks such as R-CNN, and Faster R-CNN are able to localize the forgery using bounding boxes. For this reason, object segmentation networks such as Mask R-CNN and U-net are preferred to precisely localize the forgery at the pixel level. In [90] an improved Mask R-CNN network is proposed. For precise forgery segmentation, a sobel based edge agreement head is joined to the mask prediction branch of the Mask R-CNN. In [101] a dense U-net based architecture is utilized. The image residual obtained by SRM filters is concatenated with the image pixels to enhance the learning process of the dense U-net. Through multi-scale up-sampling and concatenation, the features in the convolutional network are moved to the deconvolutional network to exploit the contextual features intersection for improving the forgery localization.

DNNs can easily localize splicing forgery utilizing the tampering traces [97]. But, this isn't the case in localizing CMF because almost all image properties are highly consistent [97]. So, in [100] the model's performance in detecting CMF is the worst compared to other forgery types. To handle CMF, a comparison mechanism between the image objects is needed. Also in [99] the experimental results provided for CMF localization isn’t enough.

C. *Hybrid Detection Methodology*

Discriminating forged region(s) from its source region(s) is favored task in forensic investigations. CMFD techniques with hybrid detection methodology such as [88], [98] are usually aim to discriminate forged regions from its source region besides localizing them at the pixel level. BusterNet [88] is considered the first end to end DNN that is able to discriminate forged region from its source region besides localizing them at the pixel level. BusterNet consists of two parallel branches that are fused together. One branch is responsible for localizing the forged region besides its source region based on visual similarity. The other branch is responsible for localizing at the pixel level only the forged region within the entire image based on visual artifacts. It comprises of a CNN based feature extractor and a deconvolutional network. But, BusterNet fails to discriminate the source region from its forged region if any of its branches wrongly locate the regions. So, the proposed model in [98] solves this problem and proposing more faster network with less parameters than BusterNet. The proposed model in [98] consists of two serial sub-networks. The first sub-network is responsible for localizing similar regions at the pixel level which will be cropped and transferred to the second sub-network as sub-images. The second sub-network follows the same structure of the constrained CNN and is responsible for deciding the class label of each sub-image if it is source region or forged region.

V. Evaluation Metrics

The outcome of a CMFD technique could be classifying an image as authentic/tampered, or localization of the cloned regions within the image at the pixel level. Such localization of the cloned regions requires classifying each image pixel as authentic/tampered. In this way, any CMFD technique can be viewed as a classifier and its performance could be measured at the image level or at the pixel level. However, the pixel-level evaluation is the most accurate and reliable way. The standard evaluation metrics for CMFD techniques are mostly depending on some measures which could operate at the image level or the pixel level. These measures are: $\mathrm{T_P}, \mathrm{F_P}, \mathrm{T_N}, and\ \mathrm{F_N}$ [106]. True Positive ($\mathrm{T_P}$) represents the No. of tampered images/pixels correctly recognized as tampered. False Positive ($\mathrm{F_P}$) represent the No. of authentic images/pixels erroneously recognized as tampered. True Negative ($\mathrm{T_N}$) represents the No. of authentic images/pixels correctly recognized as authentic. False Negative ($\mathrm{F_N}$) represents the No. of tampered images/pixels erroneously recognized as authentic.

From the above measures, different evaluation metrics can be computed at the image / pixel level as listed below [106]:

$$precision\ (p) = \frac{T_P}{T_P+F_P} \quad (1)$$

$$Recall\ (r)/TPR/sensitivity = \frac{T_P}{T_P+F_N} \quad (2)$$

$$F1 = \frac{2.p.r}{p+r} = \frac{2.T_P}{2.T_P+F_P+F_N} \quad (3)$$

$$Accuracy\ (Acc) = \frac{T_P+T_N}{T_P+F_P+T_N+F_N} \quad (4)$$

$$False\ positive\ rate\ (FPR) = \frac{F_P}{F_P+T_N} \quad (5)$$

The performance of CMFD techniques can also be evaluated through the receiver operating characteristics (ROC). The ROC curve examines the effect of various thresholds on the prediction result by plotting the TPR against the FPR. However, it is common to convert the ROC curve into single value by computing the area under the ROC curve (AUC). The AUC value of certain classification system represents its discrimination capability and hence facilitates the performance comparison of different classification systems [88].

As mentioned before, several attacks including the geometric operations and the post-processing operations are performed to make it difficult to detect the CMF. So, it is required to test the ability of the CMFD techniques to face these attacks. Such type of test is called the robustness test. In robustness test, the evaluation metrics mentioned above are usually measured for various attacks. Here comes the role of the evaluation datasets and how it covers various attacks.

VI. The CMFD Datasets

Many datasets are available for CMFD in which they vary according to some aspects such as the dataset volume, the way for expressing its ground truth, and the dataset complexity. Table II highlights the main CMFD datasets utilized in literature along with its main characteristics.

The dataset volume is determined by the number of authentic (A) & tampered (T) images it contains, the images size, and the images format. In general, evaluating the performance of CMFD techniques using datasets with massive number of images obtains reliable measures at the expense of the time complexity. Images with high resolution could provide more details that facilitate the forgery detection task. On the other hand, low image size is preferred for fast computation. CMFD datasets with compressed images add some difficulty for the forgery detection task because of missing some details.

The CMFD datasets commonly provide its ground truth at two levels: at the image level and/or at the pixel level. At the image level, each image should have a class label to indicate if it is authentic or tampered. Evaluating the pixel-level performance requires the presence of the ground truth localization masks. Not all the CMFD datasets provide binary masks that localize the cloned regions within the tampered images. But since most of the CMFD datasets provide the authentic images and its tampered images, it is possible to indirectly obtain the CMF localization mask through images subtraction followed by thresholding and morphological operations. This idea was adopted by the authors of [13] to get the localization masks for the CAISA dataset [107].

The dataset complexity is determined by the challenges it contains, attacks involved in creating the forged images, and the intensity of such attacks. The attacks include the geometric transforms and the post processing operations which are usually carried out in forged images. Also, the shape and size of the cloned regions greatly affect the detection performance of CMFD techniques. Small and irregularly shaped cloned regions poses a great challenge for CMFD techniques. Furthermore, images with multiple CMF pose other challenge. Several datasets are designed to intensively cover certain challenge(s). For example, the COVERAGE [105] dataset is intensively introduce SGO regions. Also, the GRIP [103] dataset introduced several small and smooth cloned regions.

TABLE II. SUMMARY OF THE CMFD DATASETS

Dataset	**No. of images**	**Size of images**	**Format of images**	**Shape of cloned region(s)**	**Geometric transforms**	**Post processing operations**	**Existence of localization masks**	**Other notes**
The CMFD datasets with single cloning								
MICC-F220 [14]	110 A + 110 T	722*480 to 800*600	JPEG	Square or rectangular	Scaling (S) & Rotation (R)	Not exist	Not exist	Size of forged region = 1.2% of the entire image
MICC-F2000 [14]	1300 A + 700 T	2048*1536	JPEG	Square or rectangular	S & R	Not exist	Not exist	Size of forged region = 1.12% of the entire image
SBU-CM16 [102]	240 T	800*580	PNG/JPEG	arbitrary shapes	Only R	Noise addition (NA), JPEG compression (JC), blurring (B)	Exist	Although the cloned regions vary from smooth to texture, they are not skillfully prepared
GRIP [103]	80 A + 3440 T	768*1024	PNG	arbitrary shapes	S & R	NA & JC	Exist	Provide small / smooth CMF. The complete dataset is obtained by executing perl scripts.
CMH [49]	216 T	845*634 to 1296*972	PNG/JPEG	arbitrary shapes	S & R	JC	Exist	
CVIP [104]	70 A + 970 T	1000*700 or 700*1000	BMP	arbitrary shapes	S & R	Not exist	Exist	The forged region varies in size from small to large.
COVERAGE [105]	100 A + 100 T	400*486 (on average)	TIF	arbitrary shapes	S, R, and free form	Illumination change	Exist	Have multiple SGO objects. Large size of forged region.
The CMFD datasets containing multiple cloning								
MICC-F8multi [14]	8 T	800*532 to 2048*1536	JPEG	arbitrary shapes	S & R	Not exist	Not exist	
MICC-F600 [14]	440 A + 160 T	800*533 to 3888*2592	JPEG/PNG	arbitrary shapes	S & R	Not exist	Exist	
FAU [106]	48 A	3000*2300	PNG/JPEG	arbitrary shapes	S & R	NA, JC, downscaling	Exist	Cloned regions with varied size and texture are exist. Forgeries could be generated on demand through scripts.
CAISA ITDE v1.0 [107]	800 A + 451 images with CMF	384*256	JPEG	Regular / Arbitrary	S, R, and distortion	Not exist	Not exist	Include images with CMF as well as splicing
CAISA ITDE v2.0 [107]	7200 A + 3274 images with CMF	320*240 to 800*600	JPEG / TIF / BMP	Mostly arbitrary shaped	S, R, and distortion	B	Not exist	Include images with CMF as well as splicing
CoMoFoD small [108]	5000 A + 5000 T	512*512	PNG/JPEG	arbitrary shapes	S, R, and distortion	NA, JC, B, Brightness change, Contrast adjustment, Color reduction	Exist	Post processing operations are applied to the authentic images as well as the forged images. The forged region varies in size from small to large.
CoMoFoD large [108]	1500 A + 1500 T	3000*2000	PNG/JPEG	arbitrary shapes	S, R, and distortion	Same as CoMoFoD small	Exist	Same as CoMoFoD small

VII. DISCUSSION AND FUTURE DIRECTION

Assessing the visual similarity for revealing the CMF is the most effective and common detection methodology. Such detection methodology can be implemented through the conventional techniques or the deep leaning techniques. Regardless of the implementation paradigm, the detection system usually consists of three stages as follows: feature extraction, feature matching, and forgery localization. Each stage suffers from certain challenges. In the feature extraction phase, it is required to deal with small, smooth cloned regions and low resolution images as well as the geometric transforms and post processing operations. In the matching phase, dealing with similar but genuine objects and reducing the false matching rate are of great importance. In the forgery localization phase, it is essential to deal with multiple cloning.

Among the conventional CMFD techniques, keypoint based techniques and hybrid techniques have been proved to provide better performance than the block based techniques. To handle CMF with small or smooth cloned regions, there are two options: either integrating block based techniques with keypoint based techniques or covering the entire image by enough keypoints. There are several alternatives to acquire an adequate number of keypoints covering the entire image such as utilizing multiple keypoints detectors, lowering the contrast threshold of the keypoint detector, and increasing the image resolution or contrast. Among all these alternatives, techniques that handle the non-uniform distribution of the image keypoints phenomenon are favored.

The matching complexity is a fundamental problem with conventional CMFD techniques. Also, increasing the extracted features from an image to handle CMF with small/smooth cloned regions makes the matching complexity problem more difficult. Adopting low dimensional descriptors can decrease the matching time but reduce the CMFD performance. On the other hand, matching search space reduction or utilizing approximate matching are favored techniques for reducing the matching time.

For accurate localization of cloned regions that avoid detecting SGO as cloned regions, it is important to utilize descriptors with high discrimination power, choose appropriate values of the thresholds, validate the estimated geometric transform and estimate it accurately. Conventional CMFD techniques are commonly utilizing clustering or segmentation techniques for different reasons: to eliminate false matching, facilitate consistent matching direction between the cloned regions, and localize multiple cloned regions.

Although the conventional CMFD techniques have created many solutions to deal with different challenges, there are two problems that remain without an efficient solution. First, CMFD techniques aim to localize the cloned regions with high accuracy whatever the applied geometric and post processing operations. Also, in the same time the CMFD technique should have reasonable time complexity. There is a tradeoff between these objectives and a way to balance between them is needed. Second, several parameters are utilized in conventional CMFD techniques. A way to automatically choose customized values for these parameters that are suitable for each image is also needed. On the other hand, deep learning can find a solution to these two problems, as it is the best way to learn features as well as the classification task.

Because of its massive learning power, deep learning techniques can overcome many of the challenges facing CMFD. Deep learning models are able to extract features with high description and discrimination ability. Such extracted features could be further enriched by adopting attention modules and utilizing the contextual information. Deep learning techniques have achieved an adequate level of invariance to geometric transformations, and post processing operation through the polling units and data augmentation. In addition, deep learning techniques achieve scaling invariance by adopting either the atrous spatial pyramid pooling or the inception modules. Atrous spatial pyramid pooling requires less number of parameters than the inception module. However, invariance to rotation and especially large rotation needs to be investigated. Deep learning techniques are commonly handle small cloned regions by performing multi-level matching. In other words, the matching process is performed between the low level features of early layers as well as the high level features of subsequent layers.

Although deep learning systems have several achievements in many areas, their use in the CMFD problem still needs more research to improve performance. In deep learning models, it is common to resize the training/testing images to specific size to fit the input layer. The effect of this resizing operation as well as the image resolution on the detection performance should be investigated. Some deep learning based models apply a preprocessing step to suppress the image content and highlights the relationship between image pixels to reveal the forgery. More preprocessing operations need to be investigated especially to resist against noise addition and blurring.

Depending on visual similarity to reveal CMF could result in false alarms. Conventional CMFD techniques are usually verifying the suspected regions by assessing the geometric transform between them. While in deep learning based CMFD techniques such verification step is missed. It is true that relying on tampering artifacts to reveal CMF isn’t the best choice. But, combining it with the visual similarity based detection methodology helps to enhance the performance, reduce false alarms and discriminate forged regions from its source regions. Such hybrid training can be accomplished through deep learning from two streams: the image stream and the image residual stream.

The most recent CMFD techniques are summarized in Table III. All the reported performance results are at the pixel level. In case of measuring the performance of certain CMFD technique with respect to several attacks, the reported performance result is expressed as a range. From Table III, we can find that the performance of most CMFD techniques isn’t mature enough, varies from dataset to another, and needs more enhancements. Some CMFD techniques are deceiving in terms of their efficiency as they were either evaluated with small subset of test images or evaluated under simple conditions. Detecting the forgeries in certain dataset may be more difficult than other dataset because some datasets include more challenges/attacks than other datasets. Also, it is common to have many datasets include certain challenge. But, the strength of applying such challenge could vary from one dataset to another.

TABLE III. SUMMARY OF THE RECENT CMFD TECHNIQUES

Paper	Details	Performance	Reviews
Block based CMFD techniques			
[44]	Overlapping square blocks division, Gaussian hermite moments, assess similarity through euclidian distance, RANSAC, Morphological opening.	CoMoFoD small: $F1_p$=[0.9555,0.8144] GRIP: $F1_p$= 0.9805 [Plain CMF] CVIP: $F1_p$= 0.9497 [Plain CMF]	**Pros**: less sensitive to noise, able to detect small/multiple CMF. **Cons**: high time complexity, detect scaling only in the range [80% -140%], test with simple rotations.
[45]	Overlapping square blocks division, tetrolet features, assess similarity through the absolute difference, Fast outliers filtering method, Morphological opening and closing.	CoMoFoD small: $F1_p$=[0.9564,0.8282]	**Pros**: enhanced time complexity, able to detect small /multiple CMF. **Cons**: detect CMF only where scaling in the range [80% -135%] and the quality factor is 80 or more.
[29]	SWT, overlapping square blocks division, DCT mean features, similarity through euclidian distance, Morphological opening.	CoMoFoD small: $F1_p$=[0.941,0.834]	**Pros**: adopting descriptors of reduced dimension. **Cons**: a tiny subset of the dataset has been utilized for evaluation, false alarms are exist.
[46]	Square blocks division, tchebichef moments, SVD, similarity through euclidian distance, Morphological opening.	CoMoFoD: $F1_i$= [0.9586-0.7786]	**Pros**: adopting descriptors of reduced dimension. **Cons**: robustness test against geometric transformations was missed.
[42]	Image size reduction, Pyramid model construction, overlapping circular blocks division, zernike moments, KD tree, RANSAC.	FAU: $F1_i$= 0.9717	**Pros**: able to handle large scaling & rotation. **Cons**: slow, bad performance against noise addition, and small cloned regions.
[2]	Circular blocks division, PCET+ SVD, similarity through euclidian distance, PSO optimization of the matching threshold.	374 test images from CoMoFoD & CASIA: $F1_i$= [0.9913-0.8572]	**Pros**: adopting descriptors of reduced dimension. **Cons**: fail to handle images with large smooth regions, small cloned regions, and large scaling.
Keypoint based CMFD techniques			
[8]	SIFT keypoints are described by Hu's moments, Global threshold, Region growing.	CoMoFoD small: $F1_p$=0.7672 [Plain CMF]	**Pros**: able to detect CMF with rotation/flipping. **Cons**: unable to detect small CMF.
[5]	Keypoints are detected by Harris-Laplace+Hessian-Laplace+SIFT and described by SIFT, G2NN, RANSAC, Bag of Word	FAU: $F1_p$=0.8022 [Plain CMF] MICC-F600: $F1_p$=0.7011	**Pros**: able to detect small /multiple CMF. **Cons**: false alarms are exist which reduced the precision.
[50]	SIFT+LIOP, G2NN+transitive matching, Require matching density through SLIC superpixels segmentation, RANSAC	FAU: $F1_p$=0.7442 [Plain CMF]	**Pros**: able to handle cloned regions with few keypoints, enhance the matching relationship. **Cons**: the precision needs enhancement.
[74]	Keypoints are detected by SURF and described by RLBP, G2NN through Euclidean distance, Hierarchal clustering, RANSAC	COVERAGE: Acc= [0.705 - 0.645] CoMoFoD small: Acc=0.701 FAU: Acc = 0.833	**Pros**: distinguish SGO regions from cloned regions. **Cons**: tested mostly without post processing operations. Only consider jpeg compression and blurring for the COVERAGE dataset.
[11]	Image is segmented into rough and smooth regions. In rough regions: SIFT + G2NN + HAC + RANSAC. In smooth regions: Fourier Mellin + Patchmatch. Morphological operations.	FAU: $\mathrm{F1}_\mathrm{i}$=[0.9697,0.7407] GRIP: $\mathrm{F1}_\mathrm{i}$=0.9581 [Plain CMF]	**Pros**: able to detect smooth /multiple CMF as well as CMF with large scaling/rotation. **Cons**: evaluation at the pixel level isn't provided.
[57]	SIFT, the histogram of the reduced LBP, 2NN, RANSAC, Correlation coefficient computation, standard thresholding.	MICC-F220: Acc_i=0.9682 CMH: Acc_p= [0.9772-0.9766] CVIP: Acc_p= [0.982- 0.9583] COVERAGE: Acc_i=0.675	**Pros**: combining the histogram of the reduced LBP with SIFT for enhancing the performance. **Cons**: not consider multi CMF, evaluated with datasets with minimal post-processing.
Deep learning based CMFD techniques			
[88]	BusterNet: DNN with two parallel branches. Mani-Det branch: VGG16, bilinear up-sampling layers, inception modules, binary classifier. Simi-Det branch: same as Mani-Det branch + Self-Correlation, Percentile Pooling.	CAISA ITDE v2.0: $F1_p$=0.456 COVERAGE: $F1_p$=0.618 CoMoFoD small: $F1_p$=0.493	**Pros**: discriminate forged regions from its source regions. **Cons**: contextual information is lost, difficulty in detecting small cloned regions.
[97]	DNN with single stream. VGG16, atrous convolution, attention module, hierarchical feature matching, ASPP, residual refinement network.	CAISA ITDE v2.0: $F1_p$=0.455 COVERAGE: AUC= 0.8488 CoMoFoD small: $F1_p$=0.501	**Pros**: Enrich the extracted features, the CMF localization mask is refined. **Cons**: Not distinguish well SGO regions from really cloned regions.
[98]	DNN with two serial sub-networks. CMSDNet: VGG16, atrous convolution, double level self-correlation, attention module, ASPP. STRDNet: constrained conv. layer, 4 conv. groups, fully connected classification network.	CAISA ITDE v2.0: $F1_p$=0.538 COVERAGE: $F1_p$= 0.677 CoMoFoD small: $F1_p$= 0.511	**Pros**: enhances the performance of BusterNet [88] as well as the computational time, able to discriminate forged regions from its source regions. **Cons**: The performance still needs enhancement.
[86]	End to end DNN. three dense inception blocks, hierarchical feature matching and post processing.	CAISA ITDE v2.0: $F1_p$=0.6429 CoMoFoD small: $F1_p$= 0.441 [averaged for several attacks]	**Pros**: detect unseen forged regions through learning the correlations of multi scale dense features. **Cons**: bad performance for the FAU dataset.
[89]	Adaptive over segmentation by COB, keypoints detection using DoG, keypoints description using CKN, matching through KD tree, RANSAC.	CoMoFoD small: $F1_p$= 0.6318 [with no post processing]	**Pros**: the processing time is reduced due to the GPU implementation of CKN and COB. **Cons**: scale invariance needs further enhancement.
[95]	Overlapping square blocks division, feature extraction using AlexNet, matching through global threshold, Morphological operations.	GRIP: $\mathrm{F1}_\mathrm{p}$=0.93 [Plain CMF]	**Pros**: able to handle CMF with SGO regions and smooth cloned regions. **Cons**: robustness test was missed.

VIII. Conclusion

In this survey, we have studied the CMFD problem in depth. We have categorized the CMFD techniques based on their detection methodology, their detection paradigm, and their detection capability. Different detection methodologies and paradigms have been analyzed and discussed regarding their advantages and disadvantages. Moreover, we have deeply examined the challenges that face the CMFD techniques in general and the conventional CMFD techniques in specific. Consequently, this survey gives an integrated and in-depth view of the CMFD techniques, challenges and recent trends.

The CMFD is a very challenging problem and still an open research area. The majority of the CMFD techniques aren't achieved yet good enough performance due to many conflicting challenges. In order ensure that a specific CMFD technique has achieved satisfactory results, it should be evaluated at the pixel level and evaluate its robustness against a wide range of challenges that might face the CMFD techniques. Consequently, additional work should be carried out to solve several conflicting challenges and there is a great need to further investigate and employ diverse deep learning capabilities in tackling the CMFD problem.

References

[1] Hegazi, A. Taha, and M. M. Selim, "Copy-Move Forgery Detection Based on Automatic Threshold Estimation," International Journal of Sociotechnology and Knowledge Development, vol. 12, no. 1, pp. 1–23, 2020.

[2] Y. Wang, X. Kang, and Y. Chen, "Robust and accurate detection of image copy-move forgery using PCET-SVD and histogram of block similarity measures," Journal of Information Security and Applications, vol. 54, pp. 1–11, 2020.

[3] P. Niyishaka and C. Bhagvati, "Copy-move forgery detection using image blobs and BRISK feature," Multimedia Tools and Applications, 2020.

[4] C. Lin, W. Lu, W. Sun, J. Zeng, T. X. J. Lai, and W. Lu, "Region duplication detection based on image segmentation and keypoint contexts," Multimedia Tools and Applications, vol. 77, pp. 14241–14258, 2018.

[5] C. Lin, W. Lu, X. Huang, K. Liu, W. Sun, and H. Lin, "Region duplication detection based on hybrid feature and evaluative clustering," Multimedia Tools and Applications, vol. 78, pp. 20739–20763, 2019.

[6] Y. Yu, G. Wang, and J. Zhao, "FI-SIFT Algorithm for Exposing Image Copy-Move Forgery with Reflection Attacks," International Journal of Network Security, pp. 1–8, 2019.

[7] A. Badr, A. Youssif, and M. Wafi, "A Robust Copy-Move Forgery Detection In Digital Image Forensics Using SURF," in 8th International Symposium on Digital Forensics and Security (ISDFS), 2020.

[8] C. Chen, W. Lu, and C. Chou, "Rotational copy-move forgery detection using SIFT and region growing strategies," Multimedia Tools and Applications, pp. 1–16, 2019.

[9] C. Wang, Z. Zhang, and X. Zhou, "An Image Copy-Move Forgery Detection Scheme Based on A-KAZE and SURF Features," Symmetry, vol. 10, no. 12, pp. 1–20, 2018.

[10] H. Huang and A. Ciou, "Copy-move forgery detection for image forensics using the superpixel segmentation and the Helmert transformation," EURASIP Journal on Image and Video Processing, vol. 68, 2019.

[11] K. B. Meena and V. Tyagi, "A hybrid copy-move image forgery detection technique based on Fourier-Mellin and scale invariant feature transforms," Multimedia Tools and Applications, 2020.

[12] Y. Liu, H. Wang, Y. Chen, H. Wu, and H. Wang, "A passive forensic scheme for copy-move forgery based on superpixel segmentation and K-means clustering," Multimedia Tools and Applications, vol. 79, no. 1, pp. 477–500, 2020.

[13] M. Jaberi, G. Bebis, M. Hussain, and G. Muhammad, "Accurate and robust localization of duplicated region in copy – move image forgery," Machine Vision and Applications, vol. 25, no. 2, pp. 451–475, 2014.

[14] I. Amerini, L. Ballan, R. Caldelli, A. Del 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, pp. 1099–1110, 2011.

[15] R. Dixit and R. Naskar, "Region duplication detection in digital images based on Centroid Linkage Clustering of key – points and graph similarity matching," Multimedia Tools and Applications, vol. 78, pp. 13819–13840, 2019.

[16] T. Mahmood, Z. Mehmood, M. Shah, and T. Saba, "A robust technique for copy-move forgery detection and localization in digital images via stationary wavelet and discrete cosine transform," Journal of Visual Communication and Image Representation, vol. 53, no. 1, pp. 202–214, 2018.

[17] F. Zhao, W. Shi, B. Qin, and B. Liang, "A Copy-Move Forgery Detection Scheme with Improved Clone Region Estimation," in Third International Conference on Trustworthy Systems and Their Applications, 2016, pp. 8–16.

[18] R. Agarwal, O. P. Verma, A. Saini, A. Shaw, and R. Patel, "The Advent of Deep Learning-Based Image Forgery Detection Techniques," Innovative Data Communication Technologies and Application. Lecture Notes on Data Engineering and Communications Technologies, vol. 59, pp. 679–693, 2021.

[19] M. Kharanghar and A. Doegar, "Copy-Move Forgery Detection Methods : A Critique," Advances in Information Communication Technology and Computing. Lecture Notes in Networks and Systems., vol. 135, pp. 501–523, 2021.

[20] R. Thakur and R. Rohilla, "Recent Advances in Digital Image Manipulation Detection Techniques: A brief Review," Forensic Science International, vol. 312, p. 110311, 2020.

[21] M. F. M. Mursi, M. M. Salama, and M. H. Habeb, "An Improved SIFT-PCA-Based Copy-Move Image Forgery Detection Method," International Journal of Advanced Research in Computer Science and Electronics Engineering (IJARCSEE), vol. 6, no. 3, pp. 23–28, 2017.

[22] S. H. I. Wenchang, Z. Fei, Q. I. N. Bo, and L. Bin, "Improving image copy-move forgery detection with particle swarm optimization techniques," China Communications, vol. 13, no. 1, pp. 139–149, 2016.

[23] F. Zhao, W. Shi, B. Qin, and B. Liang, "Analysis of SIFT Method Based on Swarm Intelligent Algorithms for Copy-Move Forgery Detection," Lecture Notes in Computer Science, vol. 10066, no. 1, pp. 478–490, 2016.

[24] 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, 2019.

[25] G. Jiachang and G. Jichang, "Image Copy - Move Forgery Detection Using SURF in Opponent Color Space," Transactions of Tianjin University, vol. 22, no. 2, pp. 151–157, 2016.

[26] F. Yang, J. Li, W. Lu, and J. Weng, "Copy-move forgery detection based on hybrid features," Engineering Applications of Artificial Intelligence, vol. 59, no. 1, pp. 73–83, 2017.

[27] J. Wang, G. Liu, H. Li, Y. Dai, and Z. Wang, "Detection of Image Region Duplication Forgery Using Model with Circle Block," in International Conference on Multimedia Information Networking and Security, 2009, pp. 25–29.

[28] J. Zhao and J. Guo, "Passive forensics for copy-move image forgery using a method based on DCT and SVD," Forensic Science International, vol. 233, no. 1–3, pp. 158–166, 2013.

[29] F. H. Pugar, S. Muzahidin, and A. M. Arymurthy, “Copy-Move Forgery Detection Using SWT-DCT and Four Square Mean Features,” in International Conference on Electrical Engineering and Informatics (ICEEI), 2019, pp. 63–68.

[30] Y. Huang, W. Lu, W. Sun, and D. Long, “Improved DCT-based detection of copy-move forgery in images,” Forensic Science International, vol. 206, no. 1–3, pp. 178–184, 2011.

[31] M. 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.

[32] M. Zimba and S. Xingming, “DWT-PCA (EVD) Based Copy-move Image Forgery Detection,” International Journal of Digital Content Technology and its Applications, vol. 5, no. 1, pp. 251–258, 2011.

[33] X. Kang and S. Wei, “Identifying Tampered Regions Using Singular Value Decomposition in Digital Image Forensics,” in International Conference on Computer Science and Software Engineering, 2008, pp. 926–930.

[34] L. Chen, W. Lu, and J. Ni, “An Image Region Description Method Based on Step Sector Statistics and its Application in Image Copy-Rotate / Flip-Move Forgery Detection,” International Journal of Digital Crime and Forensics, vol. 4, no. 1, pp. 49–62, 2012.

[35] J. Lee, C. Chang, and W. Chen, “Detection of copy – move image forgery using histogram of orientated gradients,” INFORMATION SCIENCES, vol. 321, pp. 250–262, 2015.

[36] T. Mahmood, T. Nawaz, M. Shah, Z. Khan, R. Ashraf, and H. A. Habib, “Copy-move forgery detection technique based on DWT and Hu Moments,” International Journal of Computer Science and Information Security (IJCSIS), vol. 14, no. 5, pp. 156–161, 2016.

[37] G. Liu, J. Wang, S. Lian, and Z. Wang, “A passive image authentication scheme for detecting region-duplication forgery with rotation,” Journal of Network and Computer Applications, vol. 34, no. 5, pp. 1557–1565, 2011.

[38] P. Yang, G. Yang, and D. Zhang, “Rotation Invariant Local Binary Pattern for Blind Detection of Copy-Move Forgery with Affine Transform,” Cloud Computing and Security. ICCCS 2016. Lecture Notes in Computer Science, vol. 10040, pp. 404–416, 2016.

[39] L. Li, S. Li, and H. Zhu, “An Efficient Scheme for Detecting Copy-move Forged Images by Local Binary Patterns,” Journal of Information Hiding and Multimedia Signal Processing, vol. 4, no. 1, pp. 46–56, 2013.

[40] Y. Wang, L. Tian, and C. Li, “LBP-SVD Based Copy Move Forgery Detection Algorithm,” in IEEE International Symposium on Multimedia (ISM), 2017, pp. 553–556.

[41] T. Mahmood, A. Irtaza, Z. Mehmood, and M. T. Mahmood, “Copy - move forgery detection through stationary wavelets and local binary pattern variance for forensic analysis in digital images,” Forensic Science International, vol. 279, pp. 8–21, 2017.

[42] J. Ouyang, Y. Liu, and M. Liao, “Robust copy-move forgery detection method using pyramid model and Zernike moments,” Multimedia Tools and Applications, vol. 78, pp. 10207–10225, 2019.

[43] K. Mahmoud and A. Abu-alrukab, “Copy-Move Forgery Detection Using Zernike and Pseudo Zernike Moments,” The International Arab Journal of Information Technology, vol. 13, no. 6A, pp. 930–937, 2016.

[44] K. B. Meena and V. Tyagi, “A copy-move image forgery detection technique based on Gaussian-Hermite moments,” Multimedia Tools and Applications, vol. 78, pp. 33505–33526, 2019.

[45] K. B. Meena and V. Tyagi, “A copy-move image forgery detection technique based on tetrolet transform,” Journal of Information Security and Applications, vol. 52, pp. 1–9, 2020.

[46] 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, 2020.

[47] B. Mahdian and S. Saic, “Detection of copy – move forgery using a method based on blur moment invariants,” Forensic Science International, vol. 171, no. 2–3, pp. 180–189, 2007.

[48] K. M. Hosny, H. M. Hamza, and N. A. Lashin, “Copy-move forgery detection of duplicated objects using accurate PCET moments and morphological operators,” The Imaging Science Journal, vol. 66, no. 6, pp. 330–345, 2018.

[49] 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, no. 1, pp. 16–32, 2015.

[50] C. Lin, W. Lu, X. Huang, L. Wei, S. Hanhui, and L. Zhiyuan, “Copy-move forgery detection using combined features and transitive matching,” Multimedia Tools and Applications, vol. 78, pp. 30081–30096, 2019.

[51] M. Bilal, H. A. Habib, Z. Mehmood, T. Saba, and M. Rashid, “Single and Multiple Copy – Move Forgery Detection and Localization in Digital Images Based on the Sparsely Encoded Distinctive Features and DBSCAN Clustering,” Arabian Journal for Science and Engineering, 2019.

[52] C. Wang, Z. H. I. Zhang, Q. LI, and X. ZHOU, “An Image Copy-Move Forgery Detection Method Based on SURF and PCET,” IEEE Access, vol. 7, pp. 170032–170047, 2019.

[53] R. Davarzani, K. Yaghmaie, S. Mozaffari, and M. Tapak, “Copy-move forgery detection using multiresolution local binary patterns,” Forensic Science International, vol. 231, no. 1–3, pp. 61–72, 2013.

[54] Y. Liu, H. Wang, H. Wu, and Y. Chen, “An Efficient Copy-Move Detection Algorithm Based on Superpixel Segmentation and Harris Key-Points,” Cloud Computing and Security. ICCCS 2017. Lecture Notes in Computer Science, vol. 10602, pp. 61–73, 2017.

[55] L. Chen, W. Lu, J. Ni, W. Sun, and J. Huang, “Region duplication detection based on Harris corner points and step sector statistics,” JOURNAL OF VISUAL COMMUNICATION AND IMAGE REPRESENTATION, vol. 24, no. 3, pp. 244–254, 2013.

[56] W. Zhang, Z. Yang, S. Niu, and J. Wang, “Detection of Copy-Move Forgery in Flat Region Based on Feature Enhancement,” Digital Forensics and Watermarking. IWDW 2016. Lecture Notes in Computer Science, vol. 10082, pp. 159–171, 2017.

[57] J. Y. Park, T. A. Kang, Y. H. Moon, and I. K. Eom, “Copy-Move Forgery Detection Using Scale Invariant Feature and Reduced Local Binary Pattern Histogram,” Symmetry, vol. 12, no. 4, pp. 1–16, 2020.

[58] G. Jin and X. Wan, “An improved method for SIFT-based copy-move forgery detection using non-maximum value suppression and optimized J-Linkage,” Signal Processing: Image Communication, vol. 57, no. 1, pp. 113–125, 2017.

[59] I. Amerini, L. Ballan, R. Caldelli, A. Del, L. Del, 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, 2013.

[60] M. Emam, Q. Han, and X. Niu, “PCET based copy-move forgery detection in images under geometric transforms,” Multimedia Tools and Applications, vol. 75, pp. 11513–11527, 2016.

[61] J. Zheng, Y. Liu, J. Ren, T. Zhu, Y. Yan, and H. Yang, “Fusion of block and keypoints based approaches for effective copy-move image forgery detection,” Multidimensional Systems and Signal Processing, vol. 27, no. 4, pp. 989–1005, 2016.

[62] X. Pan and S. Lyu, “Region Duplication Detection Using Image Feature Matching,” IEEE Transactions on Information Forensics and Security, vol. 5, no. 4, pp. 857–867, 2010.

[63] S. Samir, E. Emary, K. Elsayed, and H. Onsi, “Copy-Move Forgeries Detection and Localization Using Two Levels of Keypoints Extraction,” Journal of Computer and Communications, vol. 07, no. 09, pp. 1–18, 2019.

[64] Y. Zhu, T.-T. Ng, X. Shen, and B. Wen, “Revisiting copy-move forgery detection by considering realistic image with similar but genuine objects,” arXiv preprint arXiv:1601.07262, 2016.

[65] Y. Zhu, T. Ng, B. Wen, X. Shen, and B. Li, “Copy-move Forgery Detection in the Presence of Similar but Genuine Objects,” in IEEE 2nd International Conference on Signal and Image Processing (ICSIP), 2017, pp. 25–29.

[66] A. Hegazi, A. Taha, and M. M. Selim, “An improved copy-move forgery detection based on density-based clustering and guaranteed outlier removal,” Journal of King Saud University - Computer and Information Sciences, pp. 1–9, 2019.

[67] M. A. Elaskily, H. K. Aslan, M. M. Dessouky, and E. Fathi, “Enhanced Filter-based SIFT Apprroach for Copy-Move Forgery Detection,”

Menoufia Journal of Electronic Engineering Research (MJEER), vol. 28, no. 1, pp. 159–182, 2019.

[68] Z. Fei, S. H. I. Wenchang, Q. I. N. Bo, and L. Bin, "Image Forgery Detection Using Segmentation and Swarm Intelligent Algorithm," Wuhan University Journal of Natural Sciences, vol. 22, no. 2, pp. 141–148, 2017.

[69] A. Gupta and I. Chawla, "An Efficient Copy-Move Forgery Detection Technique Using Nature-Inspired Optimization Algorithm," in Recent Advances on Memetic Algorithms and its Applications in Image Processing. Studies in Computational Intelligence, vol 873. Springer, Singapore, Springer Singapore, 2020, pp. 153–166.

[70] M. M. Al-hammadi and S. Emmanuel, "Improving SURF Based Copy-Move Forgery Detection Using Super Resolution," in International Symposium on Multimedia, 2016, pp. 341–344.

[71] M. Bilal, H. A. Habib, Z. Mehmood, R. M. Yousaf, T. Saba, and A. Rehman, "A robust technique for copy-move forgery detection from small and extremely smooth tampered regions based on the DHE-SURF features and mDBSCAN," Australian Journal of Forensic Sciences, pp. 1–24, 2020.

[72] K. Liu, W. Lu, Y. Xue, C. Lin, X. Huang, X. Liu, and Y. Yeung, "Copy move forgery detection based on keypoint and patch match," Multimedia Tools and Applications, vol. 78, pp. 31387–31413, 2019.

[73] K. Sachdev, M. Kaur, and S. Gupta, "A Robust and Fast Technique to Detect Copy Move Forgery in Digital Images Using SLIC Segmentation and SURF Keypoints," in Proceeding of International Conference on Intelligent Communication, Control and Devices, 2017, pp. 787–793.

[74] A. Roy, R. Dixit, R. Naskar, and R. S. Chakraborty, "Copy-Move Forgery Detection with Similar But Genuine Objects," Digital Image Forensics. Studies in Computational Intelligence, vol. 755, pp. 65–77, 2020.

[75] G. Muhammad, M. Hussain, and G. Bebis, "Passive copy move image forgery detection using undecimated dyadic wavelet transform," Digital Investigation, vol. 9, no. 1, pp. 49–57, 2012.

[76] M. Chowdhury, H. Shah, T. Kotian, N. Subbalakshmi, and S. D. S, "Copy-Move Forgery Detection using SIFT and GLCM-based Texture Analysis," in IEEE Region 10 Conference (TENCON), 2019, pp. 960–964.

[77] K. H. Paul, K. R. Akshatha, A. K. Karunakar, and S. Seshadri, "SURF Based Copy Move Forgery Detection Using kNN Mapping," in Advances in Computer Vision. CVC 2019. Advances in Intelligent Systems and Computing, vol. 944, Springer International Publishing, 2020, pp. 234–245.

[78] M. A. Elaskily, H. A. Elnemr, M. M. Dessouky, and O. S. Faragallah, "Two stages object recognition based copy-move forgery detection algorithm," Multimedia Tools and Applications, vol. 78, pp. 15353–15373, 2019.

[79] E. Ardizzone, A. Bruno, and G. Mazzola, "Detecting multiple copies in tampered images," in IEEE International Conference on Image Processing, 2010, pp. 2117–2120.

[80] H. A. Alberry, A. A. Hegazy, and G. I. Salama, "A fast SIFT based method for copy move forgery detection," Future Computing and Informatics Journal, vol. 3, no. 2, pp. 159–165, 2018.

[81] P. Mishra, N. Mishra, S. Sharma, and R. Patel, "Region Duplication Forgery Detection Technique Based on SURF and HAC," The ScientificWorld Journal, 2013.

[82] C. S. Prakash, P. P. Panzade, H. Om, and S. Maheshkar, "Detection of copy-move forgery using AKAZE and SIFT keypoint extraction," Multimedia Tools and Applications, vol. 78, pp. 23535–23558, 2019.

[83] T. Du, L. Tian, and C. Li, "Image Copy-Move Forgery Detection based on SIFT-BRISK," in International Conference on Control, Automation and Information Sciences (ICCAIS), 2018, pp. 141–145.

[84] G. Muzaffer and G. Ulutas, "A Fast and Effective Digital Image Copy Move Forgery Detection with Binarized SIFT," in 40th International Conference on Telecommunications and Signal Processing (TSP), 2017, pp. 595–598.

[85] D. V´azquez-Pad´ın and F. P´erez-Gonz´alez, "Exposing Original and Duplicated Regions Using SIFT Features and Resampling Traces," Digital Forensics and Watermarking. IWDW 2011. Lecture Notes in Computer Science, vol. 7128, pp. 306–320, 2012.

[86] J. Zhong and C. Pun, "An End-to-End Dense-InceptionNet for Image Copy-Move Forgery Detection," IEEE Transactions on Information Forensics and Security, vol. 15, no. 1, pp. 2134–2146, 2019.

[87] Y. Wu, W. Abd-almageed, and P. Natarajan, "Image Copy-Move Forgery Detection via an End-to-End Deep Neural Network," in IEEE Winter Conference on Applications of Computer Vision (WACV), 2018, pp. 1907–1915.

[88] Y. Wu, W. Abd-almageed, and P. Natarajan, "BusterNet : Detecting Copy-Move Image Forgery with Source / Target Localization," in Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 168–184.

[89] Y. Liu, Q. Guan, and X. Zhao, "Copy-move forgery detection based on convolutional kernel network," Multimedia Tools and Applications, vol. 77, no. 1, pp. 18269–18293, 2018.

[90] X. Wang, H. Wang, S. Niu, and J. Zhang, "Detection and localization of image forgeries using improved mask regional convolutional neural network," Mathematical Biosciences and Engineering, vol. 16, no. 5, pp. 4581–4593, 2019.

[91] J. H. Bappy, A. K. Roy-chowdhury, J. Bunk, L. Nataraj, and B. S. Manjunath, "Exploiting Spatial Structure for Localizing Manipulated Image Regions," in IEEE International Conference on Computer Vision (ICCV), 2017, pp. 4970–4979.

[92] A. Doegar, M. Dutta, and G. Kumar, "CNN based Image Forgery Detection using pre-trained AlexNet Model," International Journal of Computational Intelligence & IoT, vol. 2, no. 1, pp. 402–407, 2019.

[93] J. Ouyang, Y. Liu, and M. Liao, "Copy-Move Forgery Detection Based on Deep Learning," in 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 2017, pp. 1–5.

[94] S. Samir, E. Emary, K. El-sayed, and H. Onsi, "Optimization of a Pre-Trained AlexNet Model for Detecting and Localizing Image Forgeries," Information, vol. 11, no. 5, p. 275, 2020.

[95] G. Muzaffer and G. Ulutas, "A new deep learning-based method to detection of copy-move forgery in digital images," in Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), 2019, pp. 1–4.

[96] J. H. Bappy, C. Simons, L. Nataraj, B. S. Manjunath, and A. K. Roy-chowdhury, "Hybrid LSTM and Encoder-Decoder Architecture for Detection of Image Forgeries," IEEE Transactions on Image Processing, vol. 28, no. 7, pp. 3286 – 3300, 2019.

[97] Y. Zhu, C. Chen, G. Yan, Y. Guo, and Y. Dong, "AR - Net : Adaptive Attention and Residual Refinement Network for Copy - Move Forgery Detection," IEEE Transactions on Industrial Informatics, vol. 16, no. 10, pp. 6714 – 6723, 2020.

[98] B. Chen, W. Tan, G. Coatrieux, Y. Zheng, and Y.-Q. Shi, "A serial image copy-move forgery localization scheme with source / target distinguishment," IEEE Transactions on Multimedia, 2020.

[99] A. Mazumdar and P. K. Bora, "Siamese convolutional neural network-based approach towards universal image forensics," IET Image Processing, vol. 14, no. 13, pp. 3105–3116, 2020.

[100] P. Zhou, X. Han, V. I. Morariu, and L. S. Davis, "Learning Rich Features for Image Manipulation Detection," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 1053–1061.

[101] R. Zhang and J. Ni, "A dense u-net with cross-layer intersection for detection and localization of image forgery," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020, pp. 2982–2986.

[102] M. Zandi, A. Mahmoudi-aznaveh, and A. Mansouri, "Adaptive Matching for Copy-Move Forgery Detection," IEEE international workshop on information forensics and security (WIFS), pp. 119–124, 2014.

[103] 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.

[104] 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.

[105] B. Wen, Y. Zhu, R. Subramanian, and T.-T. Ng, “Coverage – A Novel Database For Copy-Move Forgery Detection,” in IEEE International Conference on Image Processing (ICIP), 2016.

[106] 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.

[107] J. Dong, W. Wang, and T. Tan, “CASIA image tampering detection evaluation database,” in IEEE China Summit and International Conference on Signal and Information Processing, 2013, pp. 422–426.

[108] D. Tralic, I. Zupancic, S. Grgic, and M. Grgic, “CoMoFoD - New Database for Copy-Move Forgery Detection,” in Proceedings ELMAR, 2013.