Image Copy-Move Forgery Detection Algorithms Based on Spatial Feature Domain

Ismail Taha Ahmed
College of Computer Sciences and Information Technology
University of Anbar
Anbar, Iraq
ismail.taha@uoanbar.edu.iq

Baraa Tareq Hammad
College of Computer Sciences and Information Technology
University of Anbar
Anbar, Iraq
baraa.tareq@uoanbar.edu.iq

Norziana Jamil
College of Computing and Informatics,
Universiti Tenaga Nasional
Malaysia
norziana@uniten.edu.my

2021 IEEE 17th International Colloquium on Signal Processing & Its Applications (CSPA) | 978-1-6654-1484-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/CSPA52141.2021.9377272

***Abstract*— Currently, digital image forgery (DIF) become more active due to the advent of powerful image processing tools. On a daily, many images are exchanged through the internet, which makes them susceptible to such effects. One of the most popular of the passive image forgery techniques is copy-move forgery. In the Copy-move forgery, the basic process is copy/paste from one area to another in the same image. In this paper, the proposed image copy-move forgery detection (IC-MFDs) involves five stages: image pre-processing, dividing the image into overlapping blocks, calculating the mean and standard deviation of each block, feature vectors are then sorted lexicographically, then feeding the feature vector to the Support Vector Machine (SVM) classifier to identify the image as authentic or forged. Experiments are performed on a standard dataset of copy move forged images MICC-F220 to evaluate the proposed technique. The findings indicate that the proposed IC-MFDs can be extremely accurate in terms of Detection Accuracy (98.44). We also compare some state-of-the-art approaches with our proposed IC-MFDs. It's noted that the findings obtained are better than these approaches.**

Keywords*— *Image copy-move forgery detection algorithms (IC-MFDs), Mean, Standard Deviation, SVM classifier.

I. INTRODUCTION

Images have an essential role as effective carriers of information in technology. Through using various advanced imaging devices like smartphones, huge amount of high-resolution digital images are taken and exchanged through people daily. There are different types of manipulations that cannot be easy to observed which effect on the images. The manipulation may either add, delete or alter any of the properties of the image, without offering any hint as to the update. Therefore, recently, digital image forgery detection has become an active research field.

In general, DIF detection methods can be divided into two groups, Based on the availability of original image information: The active methods are closely associated with information that belongs to the original image. Watermarking and steganographic are considered as one of these method. However, the absence of this information can limit active methods applications. Therefore, to assess its authenticity, passive methods may not depend on previous details concerning the original image. The general taxonomy of DIF detection techniques demonstrates in figure 1.

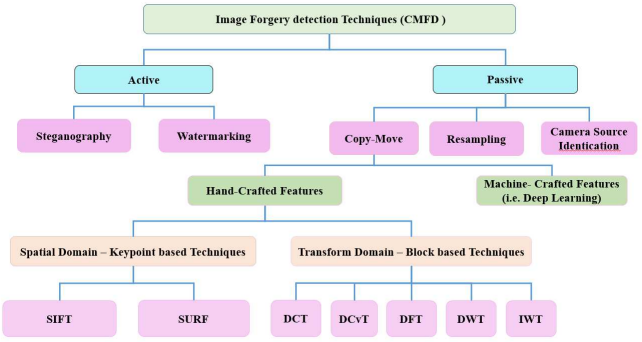

Fig. 1. DIF Detection Techniques Taxonomy.

Passive DIF detection could be narrowly classified into a few groups such as Copy-move, Resampling, and Camera Source Identification. A copy-move forgery has been one of the essential ways of passive forgery detection where it has been copied and pasted into the same image in one or more regions. When the source and target regions properties of the same image are well matched. Copy-move forgery is easily can be executed and reasonably successful in manipulating images. Figure 2 shows the original and copy-move image forgery.

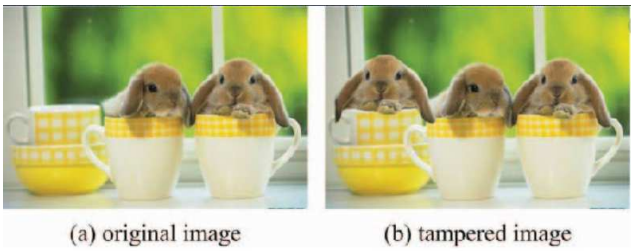

Fig. 2. Original and copy-move image forgery.

There are several copy-move image forgery detection (CMFD) techniques based on transform domain. Fridrich et al. [1] suggested an blocking approach rely on DCT coefficients, and a lexicographic sort. Khizar et al., [2], proposed CMFD by using two kinds of transform DWT and DCT. Toqeer et al., [3], suggested CMFD by using stationary wavelet transform (SWT) and DCT. Kunj et al., [4], suggested CMFD by using Tetrolet transform. These algorithms are suffers from drawback such as high computational complexity and some of them no robust to the post-processing operations such as blurring, lossy compression and a combination of these operations.

However, the main points of IC-MFDs using spatial domain starts with converting the colored image into gray scale, After that, the produced image is split into overlapping or non-overlapping blocks of the same size. For each block of these produced image, some of the features or the statistical information are obtained. These features compared with the other blocks in order to obtained any matching. Finally, potential copy_move forgery in the image could be easily detected when the match block is found. Every block is processed in the feature-based method to extract the image features such as SIFT, SURF, etc. that are then matched to identify the copy_move forgery. Compared to non-feature_based technology, the feature-based technique in the block-based approach is comparatively more feasible. Diaa et al., [5] suggested CMFD based on Hessian features and CSLBP. Guzin et al. [6] suggested CMFD by using AKAZE features and nonlinear scale space. Fan et al. [7] suggested a CMFD rely on hybrid KAZE & SIFT Features. Priya et al., [8] suggested CMFD based on combining the traditional block-based and keypoint-based techniques. Vaishnavi et al., [9], 2019 suggested CMFD based on means of symmetry based local features. Therefore, the paper proposed IC-MFDs rely on five stages: Image pre-processing, dividing the image, computing the mean and standard deviation statistical features of each block, Sorting the feature into matrix, then feeding the feature vector to the SVM classifier to identify the image as authentic or forged.

The structured of this paper is arranged in the following manner. Section 2 described the our proposed. Section 3 addresses the experimental findings. Lastly, conclusions can be made in Section 4.

II. The Proposed Methods

The main phases of the proposed methods can be classified into five stages. Figure 3 illustrates the block diagram of the suggested work. The steps details are explained as below:

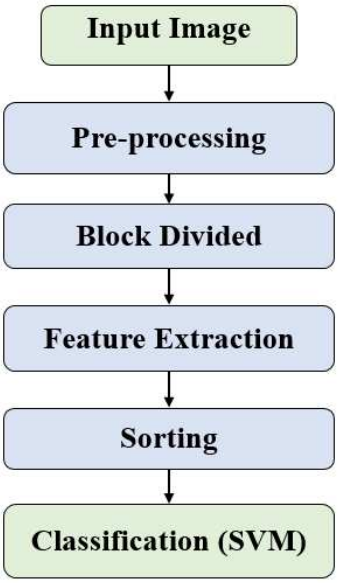

Fig. 3. The Suggested IC-MFD Block Diagram.

A. Pre-processing Stage

Initially, according to the formula given in (1), we convert the colored input image into grayscale to decrease the overall computational complexity.

$$Igray = 0.228R + 0.587G + 0.114B \quad (1)$$

Each color component (R, G, B) in above Eq (1) are equivalent to red, green, and blue, as well as *Igray* is the gray-level values.

B. Block Dividing Stage

Here, the Igray of size M x N is subdivided into 32x32 and 64x64 non-overlapping blocks. The block division will minimize the matching process time compared to an exhaustive search in order to obtain a same feature vector in an image. Next, two kinds of features mean and standard deviation are extracted from these blocks.

C. Feature Extraction Stage

Choosing good features can have a strong effect on the identification of fraud images. Feature extraction is an significant stage for copy move forgery detection. The natural image statistics may be changes in the duplicated regions. Therefore, we used an image's statistical features, like mean and standard deviation as a feature vector for each block. The mean and standard deviation of an image Igray (M, N), is formulated as

$$\mu = \frac{1}{M}\sum_{i=1}^{M} I \quad (2)$$

$$\sigma = \sqrt{\frac{1}{M}\sum_{i=0}^{M}(I-\mu)^2}\ , \quad (3)$$

where μ and σ are the mean and standard deviation of each block and stored in a feature vector matrix (F).

D. Sorting Stage

All rows are organized by using lexicographic order, obtained feature vector matrix F of each block is lexicographically sorted. the same feature vectors of rows become neighboring by applying sorting. Lexicographic sorting thus helps to minimize computational time and makes the technique more accurate in searching for similar block pairs.

E. Classification Stage

As it is known that the DIF detection is a two-class problem, i.e. authentic vs. forged. Therefore, it is required to search for a well-known Classifier that can differentiate between the original and the fake images. Using the SVM classifier, sorted feature vector matrix is used for classifying authentic and copy moved forged images.

III. Experimental Results and Analysis

The experiments are conducted in order to assess the efficiency of our method. The proposed technique is implemented on MATLAB R2020a running on HP laptop having Intel Core i7 CPU with 2.60 GHz and 8 GB RAM with Microsoft Windows 10.

A. Data Set

Different public image databases are used to test the performance of various image CMFD algorithms. The experiment performed on an MICC-F220 [10] database. The MICC-F220 Includes 220 images, 110 original images, and 110 forgeries. In this dataset, two kinds of attacks, rotation and scaling have been used to generate the forged images, two types of attacks, rotation and scaling were used. The images range in resolution from 722 × 480 to 800 × 600 pixels. JPEG is the format of these images. This dataset includes forged images in a single copy.

B. Performance Evaluation

After the images of the datasets are divided into two classes: faked and original, the CMFD evaluated at image level. The accuracy detection measured by the Eq (4).

$$Detection\ Accuracy = \frac{(Correctly\ detected\ copy-move\ images)}{(actually\ copy-moved\ images)} \times 100\ \% \qquad (4)$$

From the Eq (4) the Accuracy can be found by calculating the percentage of authentic and forged images in MICC-F220 dataset that are correctly identified.

C. Classification

It is required to search for a well-known Classifier that can differentiate between the original and the fake images. Due to its high success in identifying original and faked images, SVM has been commonly used in most splicing and copy-move detection techniques. Thus, we selected the implementation of LIBSVM as the classifier. Feature vectors created by all the images are fed into the SVM classifier. To classify the data, SVM can use different kernels. In our work, we used the SVM kernel as the radial basis function (RBF). In order to divide the dataset into training and test sets, we used the common method, i.e. ten-fold cross-validation. The training of the classifier is accomplished by choosing the features of a random 90% images from the dataset. The rest 10% of image features have been used for testing. The procedure has been replicated 10 times, a different test set in each time.

D. Proposed performance Evaluation

Detection Accuracy values of the proposed method in Table I indicate that the detection accuracy ranges between 96.32 % , 98.44 % and 97.15% as the block size varies from 32 ×32, 64×64 to 128 ×128. The performance was compared in terms of Detection Accuracy vs. block size as shown in figure 4. It is clear from Figure 4 that the algorithm presented achieves higher detection accuracy (98.44 %) on the MICC-F220 dataset. It is obvious from table I that if the block's size increases, the detection accuracy improves marginally. Therefore, the block size of 64 × 64 is considered in the proposed algorithm.

TABLE I. Detection Accuracy of proposed method across based on different block size under MICC-F220 Database.

Block size	Detection Accuracy (%)
32 × 32	**96.32**
64 × 64	***98.44***
128 × 128	***97.15***

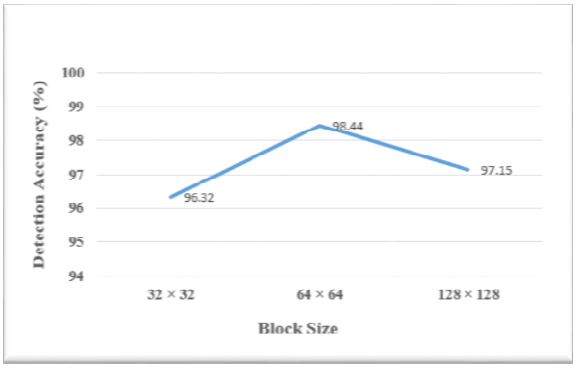

Fig. 4. Detection accuracy across various block size.

E. Comparison with Previous IC-MFD Works

The results of our proposed was compared with current CMFD techniques, which are based on the different domains. Methods[8],[7],[9],[5],[6],[11],[12],[13],[2],[14],[15],[16], [17] have recently been proposed for the period 2015-2020. In terms of detection accuracy, Table II illustrates the comparison results. The method proposed in [13] has the highest accuracy, but because of the transform domain, it consumes time. On the other hand, our technique is high accuracy and without domain transform. As shown in Table II, the proposed method showed a maximum accuracy of approximately 98.44 % compared to other current techniques based on spatial and transform domain. In terms of DA in Figure 5, the performance was compared. As seen in the Figure 5, the proposed method achieves higher detection accuracy as compared to other techniques.

TABLE II. THE COMPARATIVE RESULTS OF THE PROPOSED METHOD VERSUS VARIOUS IC-MFDS.

Techniques	***Year***	***DOMAIN***	***Methods***	***Detection Accuracy %***
Diaa et al., [5]	2015	Spatial	Hessian features	92
Guzin et al., [6]	2016	Spatial	KAZE feature	80
Fan et al., [7]	2017	Spatial	KAZE &SIFT	78.33
Umair et al., [11]	2018	Spatial	MSER &minEigen	81.33
Vaishnavi et al., [9]	2019	Spatial	Local symmetry	83.64
Gulivindala et al., [16]	2016	Transform	DWT + SIFT + PCA	64.00
Rahul et al., [13]	2017	Transform	Mean and variance DWT	***98.29***
Navdeep et al., [17]	2019	Transform	FFT+DCT+ SVM	88.62
Gulnawaz et al., [14]	2020	Transform	DCT + kd-tree	79.33
Proposed	2021	Spatial	Mean and Std and svm	***98.44***

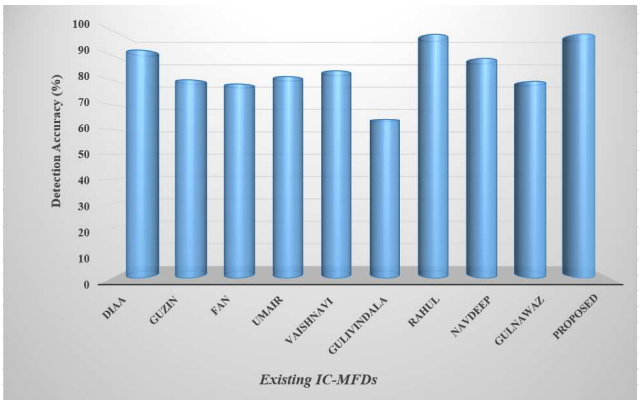

Fig. 5. Comparison of Detection accuracy of various existing IC-MFDs.

IV. CONCLUSION

The proposed IC-MFD consists of five steps: image pre-processing, dividing the image into overlapping blocks, determining the statistical features mean and standard deviation of each block, sorting the feature into a matrix, then feeding the feature vector to the SVM classifier to classify the image as authentic or forged. Our experimental findings show that copy-move forgery can be successfully detected by the proposed IC-MFDs with high accuracy (98.44 %). We compared the results of our proposed IC-MFD with many current CMFD methods. The findings revealed that our proposed produces higher outcomes than others. We plan to expand the application of our proposed in the future to distinguish other kinds of image forgeries, such as splicing.

ACKNOWLEDGMENT

This research is supported by Uniten Research Fund for Journal Publication 2020.

REFERENCES

[1] M. A. Qureshi and M. Deriche, “A bibliography of pixel-based blind image forgery detection techniques,” Signal Process. Image Commun., vol. 39, pp. 46–74, 2015.

[2] K. Hayat and T. Qazi, “Forgery detection in digital images via discrete wavelet and discrete cosine transforms,” Comput. Electr. Eng., vol. 62, pp. 448–458, 2017.

[3] 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,” J. Vis. Commun. Image Represent., vol. 53, pp. 202–214, 2018.

[4] K. B. Meena and V. Tyagi, “A copy-move image forgery detection technique based on tetrolet transform,” J. Inf. Secur. Appl., vol. 52, p. 102481, 2020.

[5] D. M. Uliyan, H. A. Jalab, and A. W. A. Wahab, “Copy move image forgery detection using Hessian and center symmetric local binary pattern,” in 2015 IEEE Conference on Open Systems (ICOS), 2015, pp. 7–11.

[6] G. Ulutas and G. Muzaffer, “A new copy move forgery detection method resistant to object removal with uniform background forgery,” Math. Probl. Eng., vol. 2016, 2016.

[7] F. Yang, J. Li, W. Lu, and J. Weng, “Copy-move forgery detection based on hybrid features,” Eng. Appl. Artif. Intell., vol. 59, pp. 73–83, 2017.

[8] P. M. Raju and M. S. Nair, “Copy-move forgery detection using binary discriminant features,” J. King Saud Univ. Inf. Sci., 2018.

[9] D. Vaishnavi and T. S. Subashini, “Application of local invariant symmetry features to detect and localize image copy move forgeries,” J. Inf. Secur. Appl., vol. 44, pp. 23–31, 2019.

[10] 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 Trans. Inf. forensics Secur., vol. 6, no. 3, pp. 1099–1110, 2011.

[11] U. A. Khan, M. A. Kaloi, Z. A. Shaikh, and A. A. Arain, “A hybrid technique for copy-move image forgery detection,” in 2018 3rd International Conference on Computer and Communication Systems (ICCCS), 2018, pp. 212–216.

[12] A. V Malviya and S. A. Ladhake, “Copy move forgery detection using low complexity feature extraction,” in 2015 IEEE UP Section Conference on Electrical Computer and Electronics (UPCON), 2015, pp. 1–5.

[13] R. Dixit, R. Naskar, and A. Sahoo, "Copy-move forgery detection exploiting statistical image features," in 2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), 2017, pp. 2277–2281.

[14] G. Gani and F. Qadir, "A robust copy-move forgery detection technique based on discrete cosine transform and cellular automata," J. Inf. Secur. Appl., vol. 54, p. 102510, 2020.

[15] H. Wang and H. Wang, "Perceptual hashing-based image copy-move forgery detection," Secur. Commun. Networks, vol. 2018, 2018.

[16] G. Suresh and C. S. Rao, "RST invariant image forgery detection," Indian J. Sci. Technol., vol. 9, no. 22, 2016.

[17] N. Kanwal, A. Girdhar, L. Kaur, and J. S. Bhullar, "Detection of digital image forgery using fast fourier transform and local features," in 2019 International Conference on Automation, Computational and Technology Management (ICACTM), 2019, pp. 262–267.