Digital Image Forensic Analysis using Hybrid Features

Conference Paper · April 2021

DOI: 10.1109/ICAI52203.2021.9445228

6 authors, including:

Tahira Nazir
Riphah International University
61 PUBLICATIONS **2,283** CITATIONS

SEE PROFILE

Aun Irtaza
University of Engineering and Technology Taxila
121 PUBLICATIONS **3,775** CITATIONS

SEE PROFILE

Ali Javed
University of Engineering and Technology Taxila
186 PUBLICATIONS **4,595** CITATIONS

SEE PROFILE

Hafiz Malik
University of Michigan–Dearborn
122 PUBLICATIONS **2,548** CITATIONS

SEE PROFILE

2021 International Conference on
Artificial Intelligence (ICAI)
Islamabad, Pakistan, April 05-07, 2021

Digital Image Forensic Analysis using Hybrid Features

Tahira Nazir
Department of Computer Science, UET Taxila
Taxila, Pakistan
tahira.nazir77@gmail.com

Aun Irtaza
Department of Computer Science, UET Taxila
Taxila, Pakistan
aun.irtaza@uettaxila.edu.pk

Ali Javed
Department of Software Engineering, UET Taxila
Taxila, Pakistan
ali.javed@uettaxila.edu.pk

Hafiz Malik
Department of Electrical and Computer Engineering,
University of Michigan-Dearborn, MI, USA.
hafiz@umich.edu

Awais Mehmood
Department of Computer Science, UET Taxila
Taxila, Pakistan
awais.mehmood@uettaxila.edu.pk

Marriam Nawaz
Department of Computer Science, UET Taxila
Taxila, Pakistan
marriam.nawaz@uettaxila.edu.pk

2021 International Conference on Artificial Intelligence (ICAI) | 978-1-6654-3293-1/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICAI52203.2021.9445228

***Abstract*—Copy Move Forgery (CMF) is an exceptional sort of image control technique in which any section of the input sample is copied into another piece of a similar image. Recently, several researchers proposed different techniques to deal with the authentication of images. However, still, there is a need for improvement both in terms of localization performance and time efficiency. In this paper, we presented a passive approach that is based on LDP with DCT for the CMF recognition. Initially, the image is transformed to YCbCr and then distributes into overlapping blocks. The features are extracted through LDP with DCT and after that similarity of the blocks is measured through Euclidean distance. The evaluation of our presented framework is performed on a challenging dataset namely CoMoFoD_Small_V2. The proposed method achieved a higher precision of 99% and a recall of 96% as compared to latest methods. Experiments show that our approach successfully detects the CMF parts and is more robust to post-processing operations.**

Keywords*— *Copy-Move forgery, LDP, DCT, Image Forensics.

I. INTRODUCTION

Recently, the availability of low-cost smart devices has enabled people to keep their data in digital format (i.e. images, videos). However, the easier access to apps and tools allow them to alter this visual content and change the information convey through it. This manipulation put a question mark on the authenticity of contents especially, in those cases where these images and videos can be used in investigating a criminal case or processing other legal claims [1]. The technique of manipulating the image data is known as digital image forgery. While the methods which are employed to detect the forensic manipulations made within digital samples are known as forgery detection (FD) techniques. These approaches are classified into two type's namely active and passive methods. Active methods are focused on those cases where the information of source samples (i.e. watermarks or digital signatures) are available. So, typically passive FD methods are applied which are capable to locate two categories of image manipulations namely CMF and splicing. For splicing based image forgery, the data of various samples are combined [2]. While the CMF is an exceptional sort of image control technique in which some area of the image itself is duplicated and inserted into another piece of a similar image. CMF is more challenging as the overall properties of the image are not altered. Two types of techniques are used for detecting CMF namely block-based and key point-based approaches. The block-based methods [3] convert an image into several overlapping blocks and then perform lexicographical sorting of each block and detects matching areas to locate forgery. In key points based technique [4], the feature vector is computed from each image on which a matching process is performed to locate forensic changes. Recently, extensive work has been performed for Digital Forensic detection, however, still, there is room for performance improvement.

Bilal et al. [5] presented a methodology to locate the manipulations of digital images. Then SURF descriptor was utilized to calculate the feature vector of the adjusted image. Finally, the mDBSCAN clustering algorithm was used to localize the forged area in each image. The approach [5] is robust to CMF detection, however, unable to localize the forensic changes made in flat areas. In this paper [6], the author introduced a technique to detect the modifications made in digital images. First, the SIFT descriptor is employed in the given image to extract the key-points from it. Then the two-level clustering approach is used to extract the clusters from the obtained set of features. A block-based technique to detect the forensic changes made within digital images is introduced in [7]. In the first step, multi-radius PCET is used for feature extraction. The radius ratio and position information are utilized to obtain the output. An approach named CMFD-PSO is presented in [8], that integrated the Particle Swarm Optimization (PSO) technique in the SIFT-based network. This

978-1-6654-3293-1/21/$31.00 ©2021 IEEE

technique exhibits good performance and improves a true positive rate. However, may not detect the forgery from the uniform texture region areas.

In the presented work, we introduce a novel method for CMF detection. After performing preprocessing, the image is converted into overlapped blocks. Then from each block, the keypoints are computed by applying LDP together with DCT. In the next step, Euclidean distance is computed among the calculated features to measure the similarity. Finally, morphological approaches are applied for displaying the forged areas. Following are the main contributions of our work:

- The presented framework compute LDP features over blocks to generate the rotation-invariant feature vector. LDP features present the more detailed image information by using the direction-based relation of central pixel to its neighborhoods, therefore, it helps in identifying those features which are robust to scaling and rotation changes and assists in effectively detecting the CMF.
- The presented framework decreased the dimensions of the keypoint vector for block representation that minimized the economical cost for CMFD.
- The proposed method is robust in locating forensic changes under the occurance of post-processing attacks like noise, rotation, scaling, and blurring.

The remaining menuscript is prepared as: Section 2 gives the detail regarding the proposed work which includes preprocessing, feature extraction, and similarity measures. Section 3 covers the results section which describes the results, and comparison. Finally, section 4 reprensts the conclusion of our presented work.

II. Proposed Methodology

This section describes the introduced methodology which is employed to identify the forensic alterations made within digital samples by using the idea of LDP [10] approach together with DCT [11] feature descriptor on overlapped blocks of the suspected sample. The matching of features is calculated by using the Euclidean distance formula. The detailed flow of the presented approach is shown in Fig. 1. The experimental evaluations of the presented framework confirmed that the LDP along with the DCT provides robust performance because of their effective identification power and low computational cost.

A. Preprocessing

In the first step, the RGB sample is converted into YCbCr. The luminance channel is more sensitive than chrominance channels by the human eye. As forgery is hard to perceive in the naked eye, chrominance is required to be more appropriate for fake detection. YCbCr reserves the connection between several channels, which is almost ignored in RGB. Identifying duplicate regions in an image reveals a broad search of local region matches. Our main idea after conversion of the image is that partitioned the converted sample into overlying blocks having a dimension of 8x8.

B. Feature Extraction

Feature selection is essential because of the high dimensionality and complex distribution of data. Removing such unrelated features can decrease the complexity of systems such as data analysis and processing time. Using LDP with DCT, the proposed strategy extracts exclusive features of the image blocks decreases the dimension of structures space, and increases the resistance of noise.

LDP operator gives a more informative description of the digital image as compared to LBP. The LDP [13] reflects the LBP as a non-directional pattern of 1st order derivative and achieves its directional development by using derivatives of higher order to get further discriminative structures. To get the LDP of nth order, the (n-1)th order derivatives (represented as $I_{\theta}^{n-1}(f_c)$) are calculated along with different directions θ:

$$I_{0^0}^{n-1}(f_c) = I_{0^0}^{n-2}(f_1) - I_{0^0}^{n-2}(f_c) \quad (1)$$

$$I_{45^0}^{n-1}(f_c) = I_{45^0}^{n-2}(f_2) - I_{45^0}^{n-2}(f_c) \quad (2)$$

$$I_{90^0}^{n-1}(f_c) = I_{90^0}^{n-2}(f_3) - I_{90^0}^{n-2}(f_c) \quad (3)$$

$$I_{135^0}^{n-1}(f_c) = I_{135^0}^{n-2}(f_4) - I_{135^0}^{n-2}(f_c) \quad (4)$$

The n^{th} order in LDP θ derivative function for f_c is described as

$$LDP_{P,\theta}^{n}(f_c) = \left\{ f_2(I_{\theta}^{n-1}(f_p), I_{\theta}^{n-1}(f_c)) \right\} | p = 1, 2, ..., P \quad (5)$$

where

$$f_2(I_{\theta}^{n-1}(f_p), I_{\theta}^{n-1}(f_c)) = \begin{Bmatrix} 0, & if\ I_{\theta}^{n-1}(f_p), I_{\theta}^{n-1}(f_c) > 0 \\ 1, & if\ I_{\theta}^{n-1}(f_p), I_{\theta}^{n-1}(f_c) \le 0 \end{Bmatrix} \quad (6)$$

The output function denotes the rotating points [14]. So, all four 8-bit directional LDPs are combined to get the nth order LDP, which is the feature vector of the image.

DCT is an important approach for CMFD and was first applied in paper [15]. After LDP, the DCT is employed to compute the keypoints [16]. Equation 7 defines the DCT of image blocks.

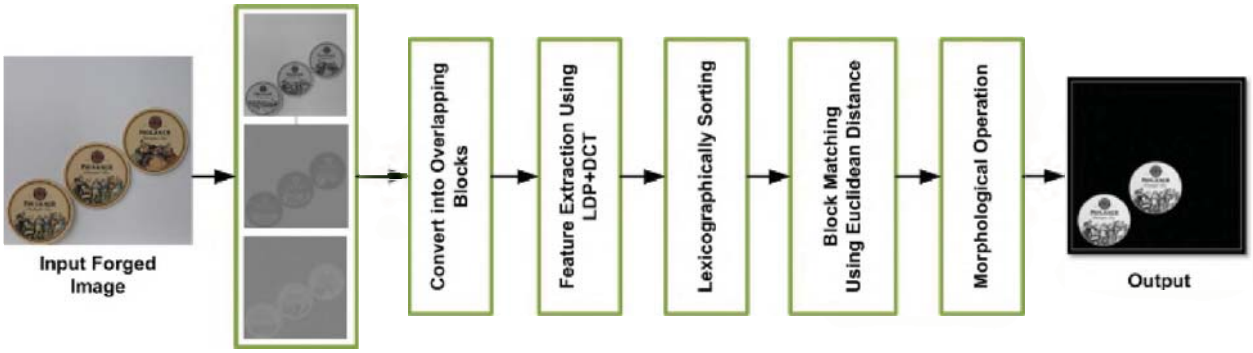

Figure 1: Flow Diagram of Proposed Method

$$f(x,y)=\frac{2}{N}.C(x).C(y)\sum_{i=0}^{7}\sum_{j=0}^{7}f(i,j)\frac{\cos(2i+1)x\pi}{2N}.\frac{\cos(2j+1)}{2N} \quad (7)$$

Where i and j are the spatial coordinates of the block of an image. Dc coefficients play an important role in feature extraction. In the DCT method, the first element is considered as DC value and all others are the AC coefficients.

C. Similarity Measures

After the feature extraction, the similarity of blocks is detected by applying the distance formula. There are different methods to measure the similarity between the blocks, and we have used the Euclidean distance [17] between two points can be found through the following equation:

$$d(X,Y)=\sqrt{(y_1-x_1)^2+(y_2-x_2)^2} \quad (8)$$

D. Morphological Operations

The morphological operations like dilation and erosion are applied for generating the binary image. To generates a binary map image, the algorithm sets the value of 1 where the regions are identical or forged and 0 for the background.

III. Experimental Results

To validate the detection accuracy of the introduced approach, extensive evaluations are conducted over the CoMoFoD dataset [9].

The evaluation power of the proposed system has been measured through precision, recall, and F1_Score, which are defined as follows:

$$Precision = \frac{Forged\ region \cap Detected\ region}{Detected\ region} \quad (9)$$

$$Recall = \frac{Forged\ region \cap Detected\ region}{Forged\ region} \quad (10)$$

$$F1\_score = \frac{2 * Precision * Recall}{Precision + Recall} \quad (11)$$

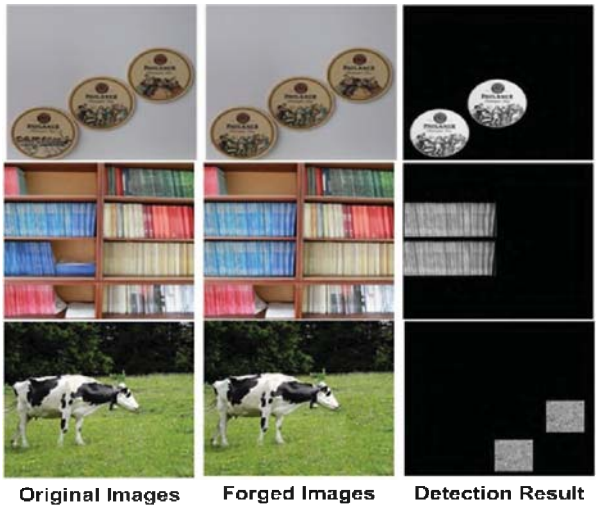

Figure 2: Detection Results of Regular Shapes

A. Detection Results

For performance evaluation, we have tested our method on images having different shapes. The outcomes demonstrate that the calculation recognizes the produced ranges precisely. The results show the ability of the calculation when the produced districts have regular shapes (as shown in Fig.2). The first column of Fig.2 represents the original images of the datasets, the second column shows the altered content where some section of the suspected sample is copied into it. While the last column shows the detected results.

TABLE I. Proposed Method Results

Image	Precision	Recall	F1_Score
Img_1	1	1	1
Img_2	1	0.9827	0.9912
Img_3	0.9986	0.8998	0.9467
Img_4	1	0.9420	0.9701
Img_5	1	0.9195	0.9580
Img_6	0.9995	0.9605	0.9796
Img_7	0.9998	0.9795	0.9896
Img_8	1	0.9547	0.9768
Img_9	1	0.9965	0.9762
Img_10	1	0.9739	0.9867

The results of our method are shown in Table.1 by using the formulas Precision, Recall, and F1_Score. Our introduced technique has attained remarkable results in terms of all employed metrics. The average accuracy of the proposed technique is 99% along with 99%, 96% and 97.7% values of precision, recall, and F1_score respectively.

B. Comparison with latest Approaches

Here, we compared our method with other competitive approaches over the CoMoFoD database. Table 2 exhibits the comparison of presented framework against the methods i.e. [18], [12], [19], [20], [21],[22], [23]. In terms of accuracy, the introduced methodology gives high performance in contrast with the other techniques. Our method has the highest precision and F1_score i.e. 99.0% and 97.71 respectively as compared to other methods. The recall value of our method is 96% which almost equals or greater than compared methods. The proposed method can calculate the key points from low-quality and noisy images and hence improved the CMFD accuracy.

TABLE II. COMPARISON WITH OTHER TECHNIQUES OVER COMOFOD.

Technique	Precision (%)	Recall (%)	F1_score (%)
DCT [18]	64.52	96.57	75.16
SWT [19]	97.17	94.08	95.47
DWT and DCT [20]	72.5	96.3	81.8
Efficient SWT [12]	98.83	95.51	97.02
DoG and ORB [21].	96.47	91.33	93.82
ACC [23].	95.65	91.67	93.62
Proposed	99.0	96.0	97.71

IV. CONCLUSION

In the introduced framework, we have presented a block-based method based on LDP together with the DCT method to locate the forensic alterations from the digital samples. Initially, the image is transformed into YCbCr which is further distributed into blocks. The LDP along with the DCT descriptor is utilized on every block to compute the final feature vector. Then Euclidean distance formula is applied over the calculated keypoints of each block to find the similarity. Finally, the morphological operation is employed to show the final output. The stated results show promising performance in comparison to other latest CMF methods under the presence of various post-processing operations like scale and rotational variations, compression, additive noise, and intensity changes. As the presented framework for CMF identification can expose the forensic manipulations under various digital image post-processing operations, therefore, it can perform a significant role in the area of image forensics-based applications.

REFERENCES

[1] J.-C. Lee, "Copy-move image forgery detection based on Gabor magnitude," J. Vis. Commun. Image Represent., vol. 31, pp. 320–334, 2015.

[2] T. Mahmood, T. Nawaz, A. Irtaza, R. Ashraf, M. Shah, and M. T. Mahmood, "Copy-move forgery detection technique for forensic analysis in digital images," Math. Probl. Eng., vol. 2016, 2016.

[3] B. Soni, P. K. Das, and D. M. Thounaojam, "CMFD: a detailed review of block based and key feature based techniques in image copy-move forgery detection," IET Image Process., vol. 12, no. 2, pp. 167–178, 2017.

[4] X.-Y. Wang, L.-X. Jiao, X.-B. Wang, H.-Y. Yang, and P.-P. Niu, "A new keypoint-based copy-move forgery detection for color image," Appl. Intell., pp. 1–23, 2018.

[5] 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 clustering," Aust. J. Forensic Sci., pp. 1–24, 2020.

[6] M. Abdel-Basset, G. Manogaran, A. E. Fakhry, and I. El-Henawy, "2-Levels of clustering strategy to detect and locate copy-move forgery in digital images," Multimed. Tools Appl., pp. 1–19, 2018.

[7] Y. Wo, K. Yang, G. Han, H. Chen, and W. Wu, "Copy–move forgery detection based on multi-radius PCET," IET Image Process., vol. 11, no. 2, pp. 99–108, 2016.

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

[9] D. Tralic, I. Zupancic, S. Grgic, and M. Grgic, "CoMoFoD—New database for copy-move forgery detection," in ELMAR, 2013 55th international symposium, 2013, pp. 49–54.

[10] N. I. Evdokimova and A. V Kuznetsov, "Copy-move detection algorithm based on local derivative pattern," in CEUR Workshop Proceedings, 2016, vol. 1638, pp. 304–312.

[11] P. Mukherjee and S. Mitra, "A review on copy-move forgery detection techniques based on DCT and DWT," Int. J. Comput. Sci. Mob. Comput., vol. 4, no. 3, pp. 702–708, 2015.

[12] T. Mahmood, Z. Mehmood, M. Shah, and Z. Khan, "An efficient forensic technique for exposing region duplication forgery in digital images," Appl. Intell., vol. 48, no. 7, pp. 1791–1801, 2018.

[13] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the support of a high-dimensional distribution," Neural Comput., vol. 13, no. 7, pp. 1443–1471, 2001.

[14] J. B. Naik, C. Srinivasarao, and G. B. Kande, "enhanced local vector pattern with logarithmic function for content-based medical image retrieval," 2017.

[15] N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete cosine transform," IEEE Trans. Comput., vol. 100, no. 1, pp. 90–93, 1974.

[16] A. Parveen, Z. H. Khan, and S. N. Ahmad, "Block-based copy–move image forgery detection using DCT," Iran J. Comput. Sci., pp. 1–11, 2019.

[17] A. Fadlil, I. Riadi, and T. Sari, "Measurement of Copy-Move Forensic Image Similarity Using Distance Function," Adv. Sci. Lett., vol. 24, no. 12, pp. 9157–9162, 2018.

[18] M. H. Alkawaz, G. Sulong, T. Saba, and A. Rehman, "Detection of copy-move image forgery based on discrete cosine transform," Neural Comput. Appl., vol. 30, no. 1, pp. 183–192, 2018.

[19] A. K. Jaiswal and R. Srivastava, "Copy-Move Forgery Detection Using Shift-Invariant SWT and Block Division Mean Features," in Recent Trends in Communication, Computing, and Electronics, Springer, 2019, pp. 289–299.

[20] 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.

[21] P. Niyishaka and C. Bhagvati, "Digital Image Forensics Technique for Copy-Move Forgery Detection Using DoG and ORB," in International Conference on Computer Vision and Graphics, 2018, pp. 472–483.

[22] Y. Sun, R. Ni, and Y. Zhao, "Nonoverlapping Blocks Based Copy-Move Forgery Detection," Secur. Commun. Networks, vol. 2018, 2018.

[23] A. V Malviya and S. A. Ladhake, "Pixel based image forensic technique for copy-move forgery detection using auto color correlogram," Procedia Comput. Sci., vol. 79, pp. 383–390, 2016.