Copy-move forgery detection in digital images using center-symmetric Local Binary Patterns and edge-based feature matching

[1]Sandeep Kaur, [2]Er. Harpreet Kaur, [3]Er.Rajandeep Kaur

[1]Masters of Engineering Student, [2,3]Assistant Professor,
[1]Computer Science & Engineering,
[1] Bahra Group of Institutions, Punjab.

Abstract: Copy-move forgery detection is a highly challenging task in image forensics. In this work, we propose a novel two-Match detection method for the accurate localization and identification of forgery regions with high efficiency. First, we propose the edge detection on forged image so that only edge pixel blocks can be considered in matching which decrease computation time of the algorithm. Secondly, we proposed center symmetric local binary patterns (CSLBP) as feature extractors for matching process. The local binary pattern, with its simple principle, low computational complexity, grayscale invariance, and illumination insensitivity, can extract the texture features of images and fuse the over- all features of an image. The basic idea of LBP is to compare the gray value of every neighboring pixel with that of the center pixel. Taking the radius of 1 and the number of pixels of 8 as an example, taking the center point as the base point, the gray value of the central point is compared with the gray values of 8 pixels in neighborhood. If the gray value of the neighboring pixels is greater than that of the center pixel, the gray values of all neighboring pixels are set to 1; on the contrary, the gray values of all neighboring pixels are set to 0. CSLBP is an effective variant of traditional LBP method. After Pre-matching, dilation and erosion morphological operations are applied to get region of interest for post-matching. It includes whole filled area around the edge pixels detected in first-Matching stage along with some false positives. Post-matching excludes such false positives which leaves only the true forged pixels. Future work can be expanded to other types of image forgeries i.e., image slicing and rotation, in order to perfect the forgery detection solution.

***Index Terms*- Copy-move forgery detection (CMFD), CS-LBP, Feature extraction, SVD etc.**

I. Introduction

In today's world, images are stored in a digital form which carries vital information and hence the integrity of digital images is very essential. For example, digital images can be used in the day-to-day life in medical diagnosis, television news, magazines, news- papers, and as legal evidence, etc. These images can be easily accessed and their content can also be tampered with using advanced image editing software. In the past decades, the forgery has been detected using the techniques of image watermarking and digital signatures and these techniques require some external information such as watermark and hash value. On the other hand, the latest research studies focus on detecting the image forgeries without external information [11].

Based on the mechanism involved in image forgery detection, there are two categories of image forgery detection techniques [13], active forgery detection techniques and passive forgery detection techniques. In the active image forgery detection techniques, a forgery is detected on the basis of prior embedded information in the input image such as watermarking [14] or digital signature. This extra information might be inserted at the time of image acquisition or at a later stage using a suitable tool. On the contrary, the passive image forgery detection techniques do not require any kind of prior information about the input image to detect image forgery; rather these techniques detect forgery on the basis of the disturbances in the intrinsic features of the image that might have been introduced during the manipulation process of the image. The images downloaded from the Internet have no prior information, hence the active forgery detection techniques are of no use for such kind of forged images. Therefore, it is also obvious that the passive forgery detection techniques are comparatively more practical today.

There are four main categories of passive forgery detection techniques [5] : copy-move forgery detection, image splicing forgery detection, re-touching detection and re-sampling detection techniques. The copy-move image forgery involves duplication of some parts of the image within the same image. Such tampering is generally performed with the intention of concealing some useful information or to replicate the things in order to mislead the people. The latest image editing tools allow user to perform copy-move forgeries with such a sophistication that it is almost impossible to say anything about the authenticity of an image just by looking at the image. Therefore, it is required to develop a copy-move forgery detection technique that can identify and locate the forgery in digital images.

In this work, we have proposed a block-based image forgery detection technique in which forged image has been divided into overlapping blocks from which mean and CSLBP features are extracted. Mean feature is used in matching process in order to find the similar blocks in the image whereas CSLBP features are used in carrying out matching process. In order to reduce time complexity of the method, edge detection has been carried out in which sobel filter has been used for edge detection. Blocks from edge pixels are taken in pre-matching process which makes the algorithm fast in pre-detection of forged pixels. Next section describes the literature review, proposed method and evaluation of the results on a standard dataset.

Many researchers have been worked in the CMFD and proposed versatile methods of forgery detection. Some of the latest methods proposed has been briefed below.

Badal Soni et al. (2018) [9] proposed two different systems for block-based copy-move forgery detection. The proposed system-I is based on the Local Binary Pattern Histogram Fourier Features and the proposed system-II is based on the Fast Walsh Hadamard transform. Due to rotation invariant characteristic of LBP-HF, the proposed system-I is efficient in forgery detection in comparison to existing block-based methods. Experimental results show that both proposed systems are able to detect small copied regions with the minimum false match. This work can also be extended for detection of forgery in the presence of geometric transformation attacks.

Ye Zhu et al. (2019) [12] presents a novel CMFD method based on MSERs and LIOP, integrating block-based and keypoints-based methods. Our proposed method abandons the traditional block-based scheme that divides the image into overlapping blocks and maintains the superiority of the keypoints-based scheme, which is effective for rotation and scale geometric transforms. Moreover, our method performs well on the tampering factor of illumination change and is robust against Gaussian noise, Gaussian blur and JPEG compression. However, our proposed method is still not effective for some degrees of rotation and scaling, which will be improved in future research.

Rajeev Rajkumar et al. (2019) [8] proposed FMZM method is useful for detecting copy-move forgery in digital images. This method does not need signatures information or any metadata. The proposed work not only identifies the forgery region otherwise it eliminates the false matches presented at the images. FMZM Transform identifies the tampered portion with maximum processing speed and low computational time complexity. Here, marker-controlled watershed management is used for the segmentation process in which projected feature extraction identifies the duplication region in the flat surface.

Maryam Jaberi et al. (2013) [4] considered the problem of copy–move image forgery detection. Our emphasis was on detecting and extracting duplicated regions with higher accuracy and robustness. The proposed methodology employs a new set of keypoint-based features, called MIFT, for finding similar regions in an image. To estimate the affine transformation between similar regions more accurately, we have proposed an iterative scheme which refines the affine transformation parameter by finding more keypoint matches incrementally. To reduce false positives and negatives when extracting the duplicated region, we have proposed using dense MIFT features in conjunction with hysteresis thresholding and morphological operations.

Reza Davarzani et al. (2013) [1] proposed an efficient forensic method based on Multiresolution Local Binary Patterns (MLBP) for detecting copy-move forgery in digital images. The method does not need digital watermarks, signatures information or any metadata. The proposed method not only detects duplicated regions but determines the geometric transformations applied to the forged regions. Experimental results demonstrated that the proposed approach could even detect duplicated regions with common postprocessing operations including: scaling, JPEG compression, Gaussian blurring and AWGN.

Jobin Varghese et al. (2019) [10] proposes a robust CMFD algorithm for identifying the location of forgery in a tampered image. The proposed algorithm uses SVD and DOST for extracting the features This paper proposes a robust CMFD algorithm for identifying the location of forgery in a tampered image. The proposed algorithm uses SVD and DOST for extracting the features superior performance with high DAR and lower FPR values with lower computational complexity in comparison with other state-of-the-art methodologies.

Choudhary Shyam Prakash et al. (2018) [6] presents an effective and robust algorithm for copy-move manipulation detection in an image. It is a type of non-intrusive (passive) technique for image manipulation detection, which means that a priori knowledge of the altered image is not required. In experiments, we observed that the DAR exceeded 70% on average, depicts the efficiency of the proposed scheme. To test the robustness of the feature vectors, AWGN for various signal to noise ratio (SNR) levels, Gaussian blurring and JPEG compression on different quality factor is applied and obtained a correlation coefficient of 1 which depict the robustness of the proposed method.

Jun Young Park et al. (2020) [7] introduced a new CMFD algorithm by adding the reduced LBP histogram-based descriptor. 256-level LBP features were first generated for a 16 × 16 window centered at a single keypoint. Next, the 256-level LBP values were reduced to 10 values to prevent an increase in the descriptor dimension. The histogram of the reduced LBP features was used as the additional descriptor to detect the CMF. In total, the proposed descriptor for a keypoint had 138 dimensions. We evaluated four types of test datasets. Additionally, the performance of the proposed method was compared with that of the existing CMFD algorithms.

Yuenan Li et al. (2012) [3] presented a robust copy-move forgery detection algorithm using the PCT and approximate nearest neighbor searching. The feature extraction scheme has been developed using the orthogonal PCT. The rotation invariance in PCT has been exploited to enable rotation-resistant CMFD. In addition, the problem of similar patch identification has been formulated as approximate nearest neighbor searching and solved by utilizing LSH. To further enhance the accuracy of CMFD, a set of post verification criteria have been developed to filter outfalse matches.

Granty Regina Elwin J et al. (2016) [2] discussed an efficient method for image retrieval and detection of copy-move forgery. Images were retrieved by sparsification of graph Laplacian using spectral hashing and copy move forgery was detected by spectral-hashing-based Polar Cosine Transform. The use of graph Laplacian in the proposed method helps to capture the intrinsic structure of the images, resulting in precise image retrieval. Sparsified graph Laplacian reduces the computational time, thereby improving the search process. The spectral-hashing-based PCT method detects the copy-move forged regions of an image in an effective manner. The rotation invariant property of PCT ensures the identification of tampered regions even if the copied portion was subjected to angular rotations. The efficiency of the proposed work has been proved using various performance evaluators, including FAR, FRR, Recall and Error Rate features.

3. The proposed algorithm

In this section, details of the proposed method for duplicated and distorted region detection are presented. There are three main steps in our algorithm: feature extraction, block matching, and estimation of the geometric transformation parameters and remove the false matches. Fig. 3 shows the block diagram for our copy-move forgery detection scheme.

3.1. Feature extraction

In this work, we are concerned with gray-level images, so RGB images are first converted to grayscale images using standard color space conversion. The input image is first divided in overlapping blocks of B × B pixels. The blocks are assumed to be smaller than the size of the duplicated regions which have to be detected. The sliding block is moved from right to left and from top to bottom. The total number of overlapping blocks, TB, for an image of M *N pixels could be determined by Eq. (7):

$$TB = (\lfloor (M-B)/Space \rfloor + 1) \times (\lfloor (N-B)/\text{Space} \rfloor + 1) \qquad (7)$$

Local binary patterns (LBP) and its extensions

Among the feature descriptors, Local Binary Patterns (LBP) is one of the most famous and powerful ones. It has gained increasing attention in many image analyses applications due to its low computational complexity, invariance to monotonic gray-scale changes and texture description ability [15]. In practice, the LBP operator combines characteristics of statistical and structural texture analysis: it describes the texture with micro-primitives, often called textons, and their statistical placement rules. The idea of LBP is originally proposed by Ojala et al. [15] for texture classification. The original version of the LBP operator considers only a 3*3 neighboring block around each pixel. These eight neighbors are labeled by thresholding with the central pixel value, weighted with powers of two and then summed to obtain a new value assigned to the central pixel. Fig. 1 shows how the original LBP is calculated. Given a center pixel in an image, the LBP value is computed by comparing it with those of its neighborhoods. LBP can be easily extended to include all circular neighborhoods with any number of pixels as shown in Fig. 2.

Assume the central pixel at position (x_c, x_y). Having P equally spaced neighborhood pixels on a circle of radius R, LBP is obtained by Eq. (1):

$$LBP_{P,R}(x_c, y_c) = \sum_{n=0}^{P-1} s(g_n - g_c)2^n, \;\; s(x) = \begin{cases} 1 : x \geq 0 \\ 0 : x < 0 \end{cases} \qquad (1)$$

where g_c and g_n correspond to the gray value of the central pixel and neighboring pixel, respectively

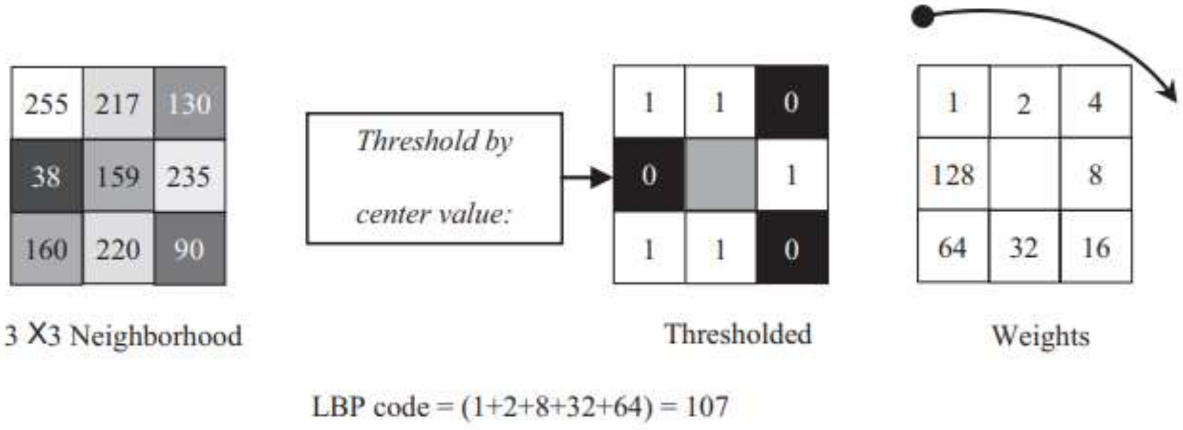


Fig. 1. The basic LBP operator.

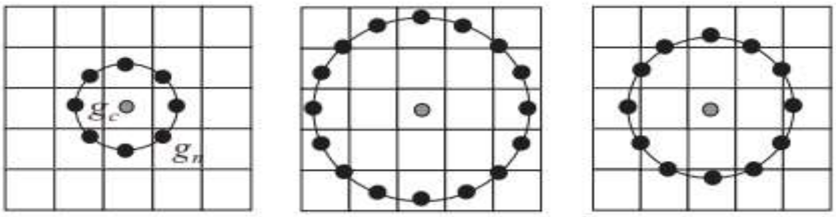


Fig. 2. Circularly symmetric neighborhoods for different P and R, (P = 8, R = 1),

(P = 16, R = 2), (P = 12, R = 1.5).

For a given M × N image, after identifying the LBP pattern of each pixel (i, j), the normalized histogram of LBP codes is computed over the image and it is used as a feature vector, Eq. (2):

$$H(t) = \left(\frac{1}{M \times N}\right) \sum_{i=1}^{M} \sum_{j=1}^{N} f(LBP_{P,\mathrm{R}}(i,j),t), t \in [0,\mathrm{T}]$$

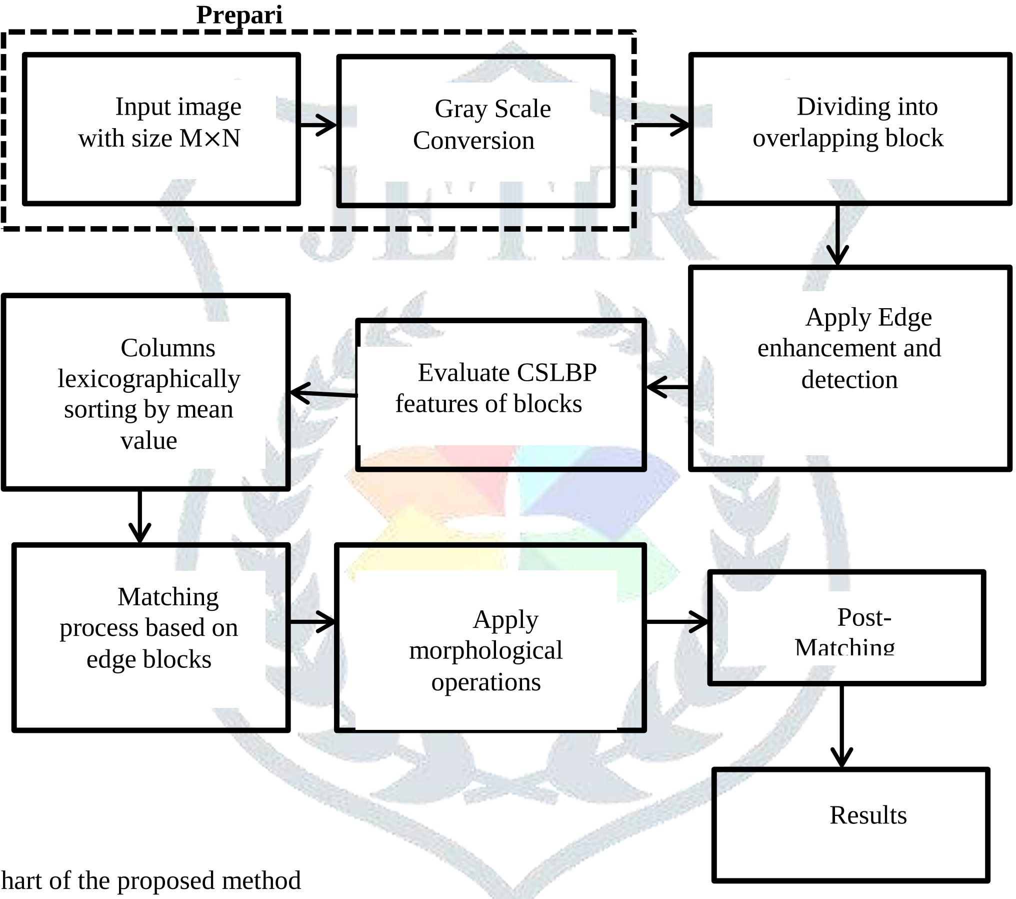
$$f(x,y) = \begin{Bmatrix} 1, & \mathrm{x=y} \\ 0, & \mathrm{otherwise} \end{Bmatrix} \quad (2)$$

where T is the maximal LBP pattern value.

3.2 Significance of different methods used

The conversion of RGB image to gray scale is useful as only single channel is needed to detect the forged regions which can later be find in colored image.

The algorithm is described using given flowchart below:



Prepari
Input image with size M×N
Gray Scale Conversion
Dividing into overlapping block
Apply Edge enhancement and detection
Evaluate CSLBP features of blocks
Columns lexicographically sorting by mean value
Matching process based on edge blocks
Apply morphological operations
Post-Matching
Results

Figure 3: Flowchart of the proposed method

- The Gabor filter is used to enhance the gradient image and to reduce the intensity unevenness further.

- The Sobel operator was originally proposed by Irwin Sobel in 1968 to estimate the gradient of a digital image. Sobel proposed to estimate the gradient of the image at a given point by performing the vector summation of the simple central gradient estimates along the 4 main directions in a 3 by 3 neighborhood. According to Sobel, each of the 4 simple central gradient estimates can be expressed as a vector sum of a pair of orthogonal vectors. Each vector is a directional derivative estimate multiplied by a unit vector specifying the derivative's direction. Sobel filter has been applied to get the binary edge image. This image is further used in matching process where matching is carried out for only white pixels which represent the edges in the original image

- In the second part of post matching, morphological operations re applied in which the iterative erosion is used to detect the cell seeds when the segmented edges are connected as a whole. The area-constrained ultimate erosion method is used to separate the overlapping cells and to detect the cell seeds when the segmented edges are not connected as a whole. To erode the binary blob one by one, we assign the sequential labeling number (i = 1,2,…,n) to each binary blob with the morphological labeling operation . Then, we could select each binary blob based on their sequential number one by one. Each labeled blob is eroded until it vanishes or its eroded-off parts vanish. Then the erosion results of the previous step are selected as the seeds. A morphological dilation is conducted to merge the possible small blobs.

- SVD and CSLBP are feature extraction algorithm's used in matching process as they provide texture properties of the image. CSLBP produces four bit binary conversion which is used in decimal form to get the content properties in the image.

4. Results and discussions

The presented edge-cslbp based methodology has been executed in MATLAB. Our test data consists of a set of 512 × 512 RGB images, taken from the CoMoFoD Database [10]. For the sake of experimentation, we have selected test images with copy–move forgery induced into them. For performance evaluation of the proposed method, recall, precision, F-measure and accuracay has been calculated for each image. First of all, Forgery detection has been extracted from whole dataset and feature extraction has been carried out using CSLBP texture algorithm and sobel edge detection. After that forged pixels has been calculated. The classification accuracy is the extent to which the classifier is able to correctly classify the exemplars and is summarized in the form of confusion matrix to the test data. This is defined as the ratio of the number of correctly classified patterns (TP and TN) to the total number of patterns (species) classified. In the following experiments, two main criteria, namely, precision and recall [10], are employed to quantitatively and qualitatively evaluate the experimental results at both the image and pixel levels. The precision and recall metrics are defined in Formulas (3) –(6) :

$$\Pr ecsion = \frac{T_p}{T_p + F_p} \qquad (3)$$

$$\operatorname{Re} call = \frac{T_p}{T_p + F_n} \qquad (4)$$

where T_p denotes true positives, F_p denotes false positives, and F_n denotes false negatives. An intuitive illustration of the relationships between T_p , F_n and F_p are shown in Tables below. To comprehensively evaluate the CMFD performances, we performed evaluations at two levels for precision, recall and F1-score the pixel level and the image level.

Pixel level:

- Precision describes the percentage of pixels that were correctly detected as a percentage of all the detected pixels.

- Recall describes the probability that the detected pixels match as a percentage of all the ground-truth forgery pixels.

Image level:

- Precision is the probability that an image in which forgery was detected is truly a forged image. A pixel-level precision above 50%, translates to a precision of 1 at the image level; otherwise, the image-level precision is 0.

- Recall is the probability that a forged image was detected as a forgery. A pixel-level recall above 50% translates to a recall of 1 at the image level; otherwise, the image-level recall is 0.

By fusing the precision and recall , we obtain an F_1 score, which comprehensively evaluates the performances of the various CMFD methods:

$$F_1 = 2 \times \frac{\Pr ecision \times recall}{precision + recall} \qquad (5)$$

$$Accuracy_1 = \frac{\left(T_p + T_N\right)}{\left(T_P + T_N + F_P + F_N\right)} \qquad (6)$$

Table 1: Forgery detection evaluation using Precision, Sensitivity, F-Score Accuracy parameters for the copy-move forged images taken from CoMoFoD Database

Parameters	TP	TN	FP	FN	Precision	Sensitivity	F-Score	Accuracy
Image1	6748	253652	1744	0	0.7946	1	0.8856	0.9933
Image2	3326	256894	1920	4	0.6340	0.9987987	0.7757	0.9927
Image3	2998	256758	2388	0	0.5566	1	0.7152	0.9909
Image4	27412	230202	4524	6	0.8583	0.9997811	0.9237	0.9827
Image5	28377	227266	6501	0	0.8136	1	0.8972	0.9752
Image6	19372	237858	4582	332	0.8087	0.9831506	0.8874	0.9813
Image7	12791	245558	1789	2006	0.8773	0.8644319	0.8708	0.9855
Image8	4916	254416	2812	0	0.6361	1	0.7776	0.9893

Table 2: Forgery detection results for the first three copy-move forged images taken from CoMoFoD Database

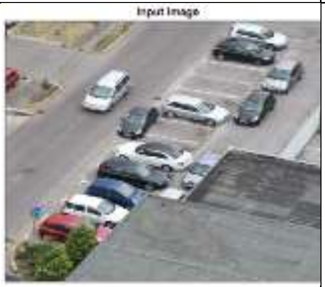
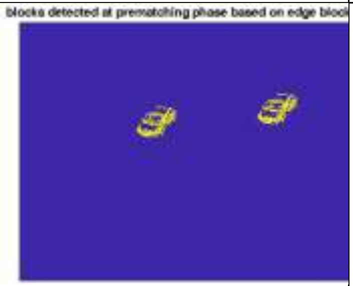
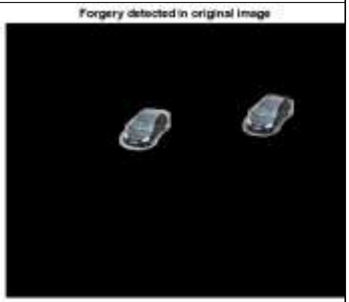
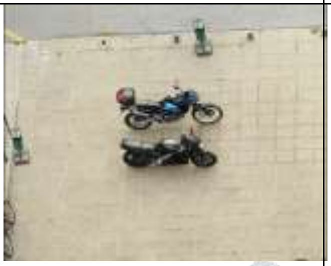
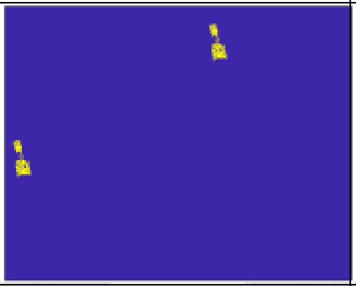
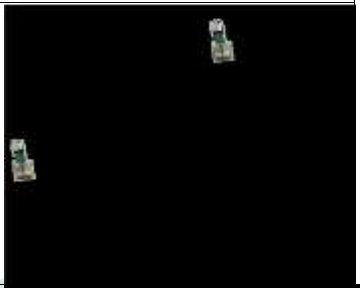
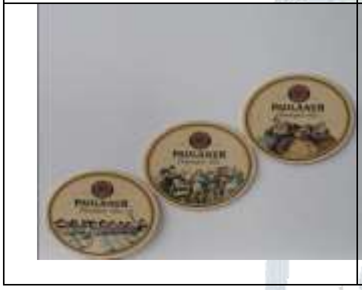
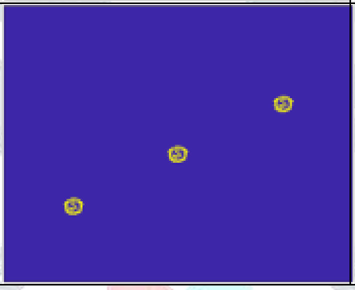
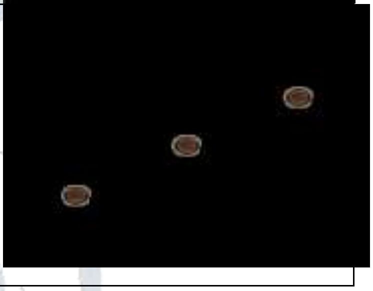
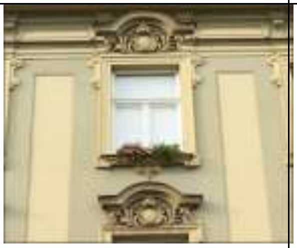
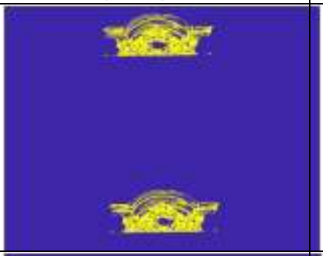
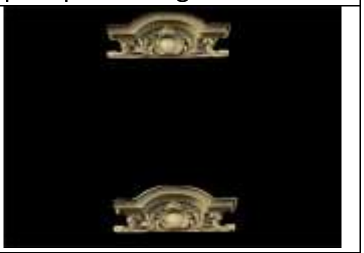
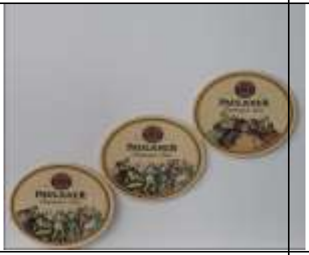
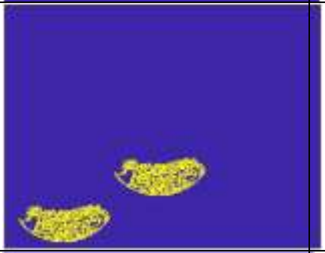
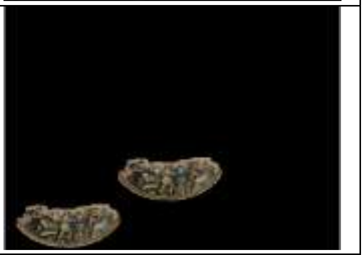
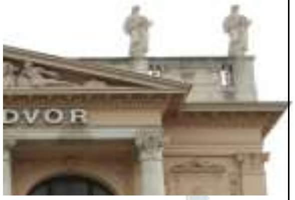
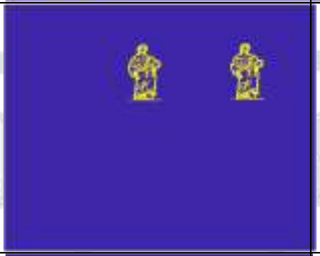
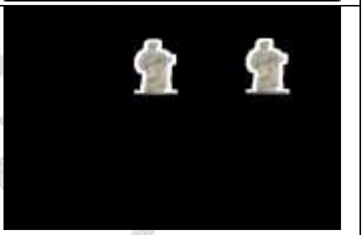
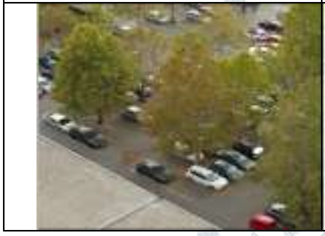
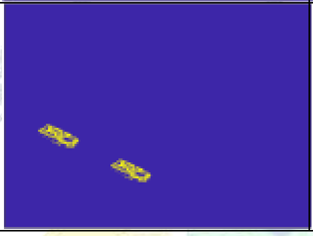
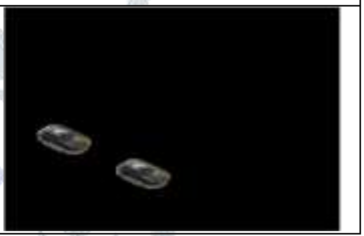
Image used	Ground Truth	Forgery area detected
		
		
		

Table 3: Forgery detection results for the next four copy-move forged images taken from CoMoFoD Database

Image used	Blocks Detected at pre-matching stage	Forgery area detected after post processing
		
		
		
		

Below are the bar graphs for Precision, Sensitivity , F-Score, Accuracy parameters

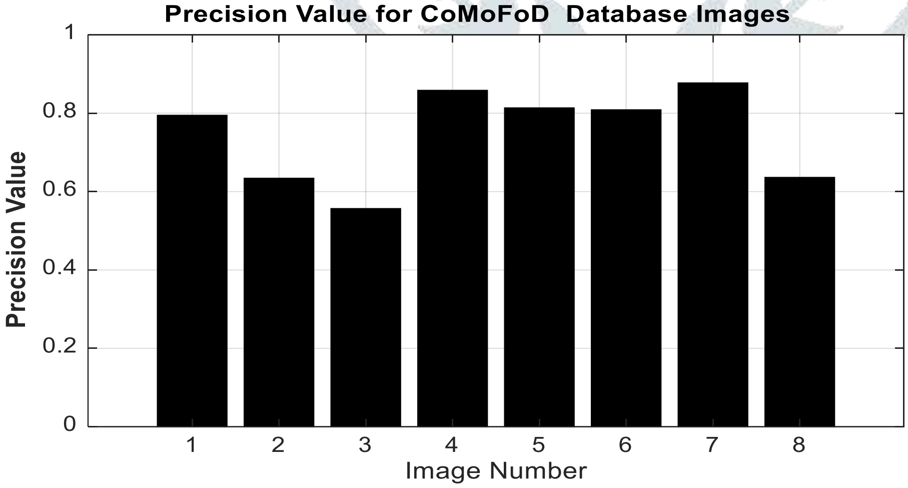


Figure 4: Precision Value for CoMoFoD Database Images

Precision quantifies the number of positive class predictions that actually belong to the positive class. Positive class is the forged region whereas negative class are background pixels in image where there is no forgery.

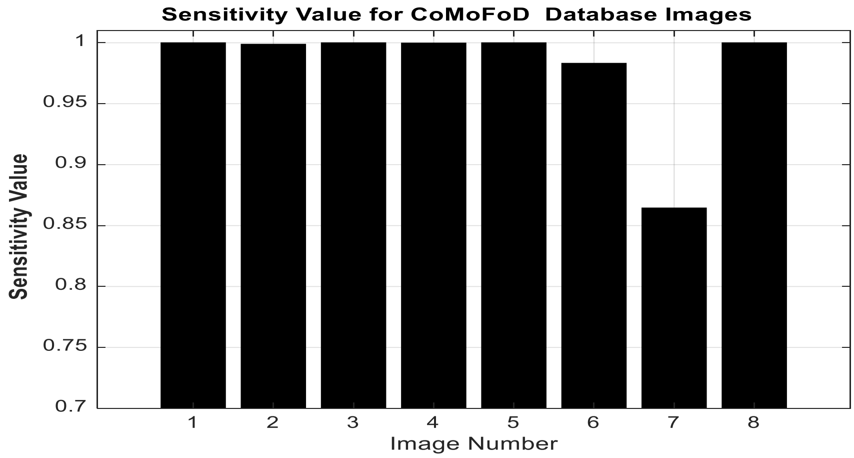


Figure 5: Sensitivity Value for CoMoFoD Database Images

Recall or sensitivity quantifies the number of positive class predictions made out of all positive examples in the dataset. It come approx. above ninety five percent for most of the images.

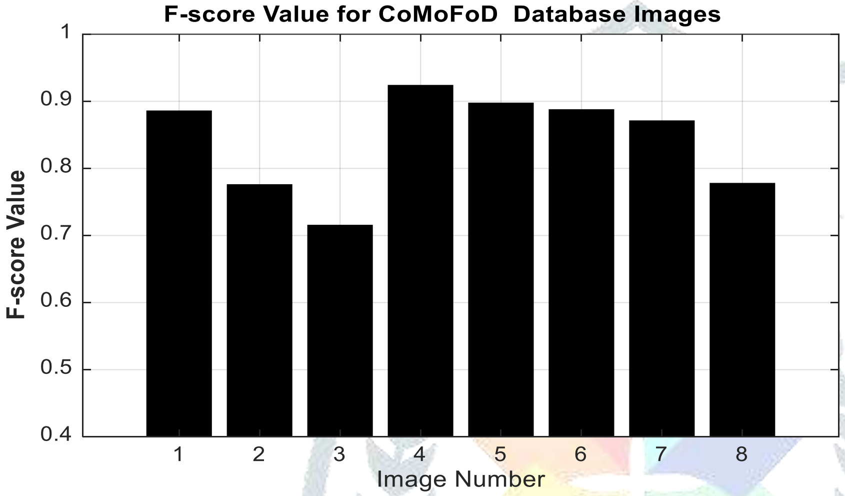


Figure 6: F-Score Value for CoMoFoD Database Images

F-Score is the weighted average of Precision and Recall. Therefore, this score takes both false positives and false negatives into account.

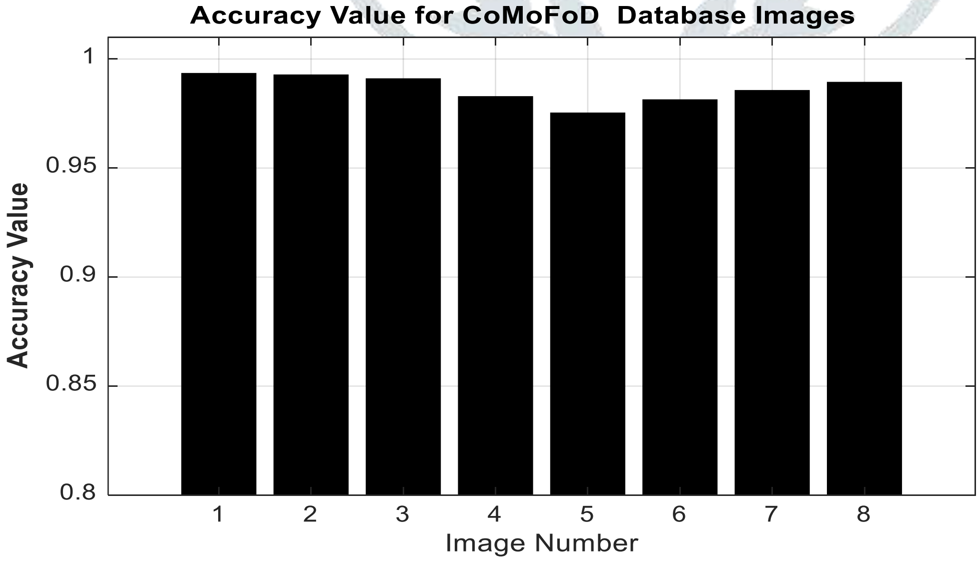


Figure 7: Accuracy Value for CoMoFoD Database Images

It has been found that proposed method of forgery detection is effective when a patch or block of an image is copy moved to another places. As Euclidian distance gives minimum error or difference between two feature sets, an exact replica of a patch of forging can be easily detected by the proposed method. Small artefacts or noisy forged detected blocks have been eliminated using morphological operations, higher accuracy has been achieved which conforms accuracy of about 98-100 per cent.

5. CONCLUSION

With better access to various image editing tools, the credibility of digital images is at stake. Images are used as a proof of reality - both in formal and informal settings and hence, its authenticity is of prime concern. In the past few years, the detection of tampered images has gathered much attention from researchers worldwide. A particular focus is given to copy-move forgery detection as it is one of the most commonly used image tampering techniques. This work presented an integrated approach for detection and localization of forged regions. Forged images are identified from standard dataset of images using texture-based center-symmetric local binary patterns with average detection accuracy above 98%. A series of experimental results showed that the proposed detector gives better detection rates in comparison to other existing algorithms in literature. Also, this work achieves relatively high detection accuracy with fewer features. Once the forged images are identified, localization of tampered region is performed using techniques based on the type of forgery. Computation time has been decreased in matching the forged pixels by applying edge detection method. This binary edge image is used in matching such that only edge blocks are matched which decreases the number of computations in pre-matching process. Then morphological operations are applied which filled the detected forged pixels. After that Post-Matching is applied to get the true forged regions in the image, Performance evaluation of the proposed method has been carried using Recall, precision, F-Measure and accuracy quality metrics.

References:

[1] Davarzani, R., Yaghmaie, K., Mozaffari, S., & Tapak, M. (2013). Copy-move forgery detection using multiresolution local binary patterns. *Forensic science international*, *231*(1-3), 61-72.

[2] Granty, R. E. J., & Kousalya, G. (2016). Spectral-hashing-based image retrieval and copy-move forgery detection. *Australian Journal of Forensic Sciences*, *48*(6), 643-658.

[3] Li, Y. (2013). Image copy-move forgery detection based on polar cosine transform and approximate nearest neighbor searching. *Forensic science international*, *224*(1-3), 59-67.

[4] Jaberi, M., Bebis, G., Hussain, M., & Muhammad, G. (2014). Accurate and robust localization of duplicated region in copy–move image forgery. *Machine vision and applications*, *25*(2), 451-475.

[5] Meena KB , Tyagi V . Image forgery detection : survey and future directions, Data, Engineering and Applications, 2. Singapore: Springer; 2019. p. 163–95

[6] Prakash, C. S., Maheshkar, S., & Maheshkar, V. (2018). Detection of copy-move image forgery with efficient block representation and discrete cosine transform. *Journal of Intelligent & Fuzzy Systems*, *35*(5), 5241-5253.

[7] Park, J. Y., Kang, T. A., Moon, Y. H., & Eom, I. K. (2020). Copy-Move Forgery Detection Using Scale Invariant Feature and Reduced Local Binary Pattern Histogram. *Symmetry*, *12*(4), 492.

[8] Rajkumar, R., Roy, S., & Manglem Singh, K. (2019). A robust and forensic transform for copy move digital image forgery detection based on dense depth block matching. *The Imaging Science Journal*, *67*(6), 343-357.

[9] Soni, B., Das, P. K., & Thounaojam, D. M. (2018). Dual System for Copy-move Forgery Detection using Block-based LBP-HF and FWHT Features. *Engineering Letters*, *26*(1).

[10] Varghese, J., & Kumar, C. S. (2019). Robust copy-move forgery detection algorithm using singular value decomposition and discrete orthonormal Stockwell transform. *Australian Journal of Forensic Sciences*, 1-17.

[11] Vaishnavi, D., & Subashini, T. S. (2019). Application of local invariant symmetry features to detect and localize image copy move forgeries. *Journal of information security and applications*, *44*, 23-31.

[12] Zhu, Y., Shen, X., & Liu, Y. (2019). Copy-move forgery detection based on local intensity order pattern and maximally stable extremal regions. *Journal of Intelligent & Fuzzy Systems*, (Preprint), 1-8.

[13] Tyagi V . Understanding digital image processing. CRC Press; 2018 .

[14] Wang S , Zheng D , Zhao J , Tam WJ , Speranza F . An image quality evaluation method based on digitalwatermarking. IEEE Trans Circuits Syst Video Technol 2007;17:98–105

[15] Ojala, T., Pietikainen, M., & Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on pattern analysis and machine intelligence, 24(7), 971-987.