DETECTING COPY MOVE FORGERY USING DCT AND MEAN VALUE OF BLOCKS

Gurpreet Kaur[1*], Dr. R.K. Bathla[2#]
[1*]Research Scholar: Department of Computer Science, Desh Bhagat University, Mandi Gobindgarh, India
[2#] Professor: Department of Computer Science, Desh Bhagat University, Mandi Gobindgarh, India.

ABSTRACT: One of the most common ways of image forgery is to embed a duplicate. In other words, this attack is called a copy-move. The embedding process consists of three stages: copying a fragment, adding changes to this fragment, and inserting a fragment into that area of the image whose contents are supposed to be hidden from the end user.In this paper, block based feature extraction and matching process is used. At first edge detection is carried out to get the high entropy pixels in the image so that matching process is carried out only for high entropy pixel blocks. Then feature extraction is carried out by converting image into overlapped blocks and mean and DCT features are extracted. Then mean values are put into a matrix and corresponding blocks are noted. Then mean value matrix is sorted in order to match the blocks with similar mean values. In matching process, variance of DCT features is used for similarity measure and forgery detection. Experimental results show that proposed method has high accuracy of forgery detection which comes in range of 97 to 99% along with least computation time.

Keywords—Copy Move Forgery, Block Based Forgery Detection, Mean, Variance, DCT, LBP etc.

I. INTRODUCTION

Nowadays, digital images are used widely in numerousareas in our lifecyclefor instanceforensics sciences, news reports, online marketing,surveillance services and medical diagnosis. Furthermore, these could be used as evidence in courts, and in the media to transform the sense of imageswith the purpose of affecting the readers' points of observations. Therefore, theregion of digital image forensics [1] to state the originality of digital image has come to beasignificant area of investigation to regain belief in digital image [2]. The forensic examination for digital images services in providing information to support security,law enforcement, and intelligence agencies. Numerousmethodsare introduced to examine the digital image's content. The image forgery detection is explored to passive and active methods [3]. At the present time,it is simple to generate image forgeries bycommandingpresent digital image processing software packages. Image Forgery is of two types: copy-move forgery and splicing forgery [4]. In copy-move forgery, portions of one image are copied and then pasted into the image itself, whereas in splicing forgery;portions of one or more images are copied and then pasted into a different image. Recognition of copy-move forgery has been extensively investigated [4]. Established approaches for copy-move forgery detection can be regarded askeypoint-based and block-based methods. Keypoint-based methods embrace scanning of the entire image with the target of verdict points of attention (for example, point with high entropy). Those opinions are then examined to select only point with the identical possessions and distinguish analogous zones in the image. Various prevalent instances of keypoint-based methods are SIFT (Scale-invariant feature transform) [5] and SURF (Speeded Up Robust Features) [6]. Block-based approaches comprise separating an image into insignificant overlying blocks as a leading phase of the process. A set of features is then intended for each definite block, and those features are castoff for detection of analogous blocks in the image. Diverse sets of features, for instance DCT (Discrete Cosine Transform) [4] / DWT (Discrete Wavelet Transform) [7] factors, Zernike moments [9] or PCA (Principal Component Analysis) [8], have been projected for practice in block-based methods.

- Block-Based Method for CMFD

In general all block-based copy move forgery detection approaches track analogous phases:

1. First the image is pre-processed since most algorithms necessitate only the luminance component evidence, and so it is required to alter images to grayscale space. From time to time Gaussian pyramid decomposition is also smeared (as, in [10]).

2. Afterward pre-processing, an image is alienated into overlying blocks by gliding a predefined window by one pixel through the whole image. The size of the window is frequently insignificant (for illustration, 8×8, 16×16, 24×24 pixels) to guarantee recognition of zones of all magnitudes. Distributing an N×M image into overlying blocks of size b×b leads to a very bulky numeral of altered blocks affording to equation (3) (for illustration: distributing a 512×512 image by means of a 8×8 window yields 255,025 dissimilar blocks).

$$N_b = (N - b+1)\times(M-b+1) \qquad (3)$$

3. For each definite block a feature vector f is intended by identical process. The feature vector is castoff as a condensed depiction of a block since it comprehends evidence about texture, shape, orientation or certain other assets of a block. The scale of the feature vector hinges on a selection of way for its deviousness.

4. Smearing brute-force exploration to catch analogous blocks by communal evaluation of all pairs of blocks entails a proportion of computational time and assets. Consequently, altogether feature vectors are warehoused in one matrix that is organized by particular procedure (for sample, lexicography categorization) to undertake assemblage of analogous blocks. Alongside categorization, several supplementary ways and means for vindicating analogous blocks can be pragmatic, for instance, kd-tree.

5. Neighbor feature vectors in the organized matrix are than paralleled by scrutinizing the correspondence among them, via

the Euclidean distances concerning feature vector elements rendering to equation (4). All pairs of blocks with remoteness v advanced than certain predefined threshold Ts are detached from the set of probable outcomes. Assortment of threshold Ts contingent on the category of forgery, for specimen, it can be agreed to zero for plain CMF, or it has to be attuned to specific higher values if any transformations/post-processing procedures are smeared. Afterward this phase only analogous pairs of blocks are held in reserve as probable outcomes.

$$v=\sqrt{\sum_{i=1}^{size(f)}(f_1(i)-f_2(i))^2} \qquad \textbf{(4)}$$

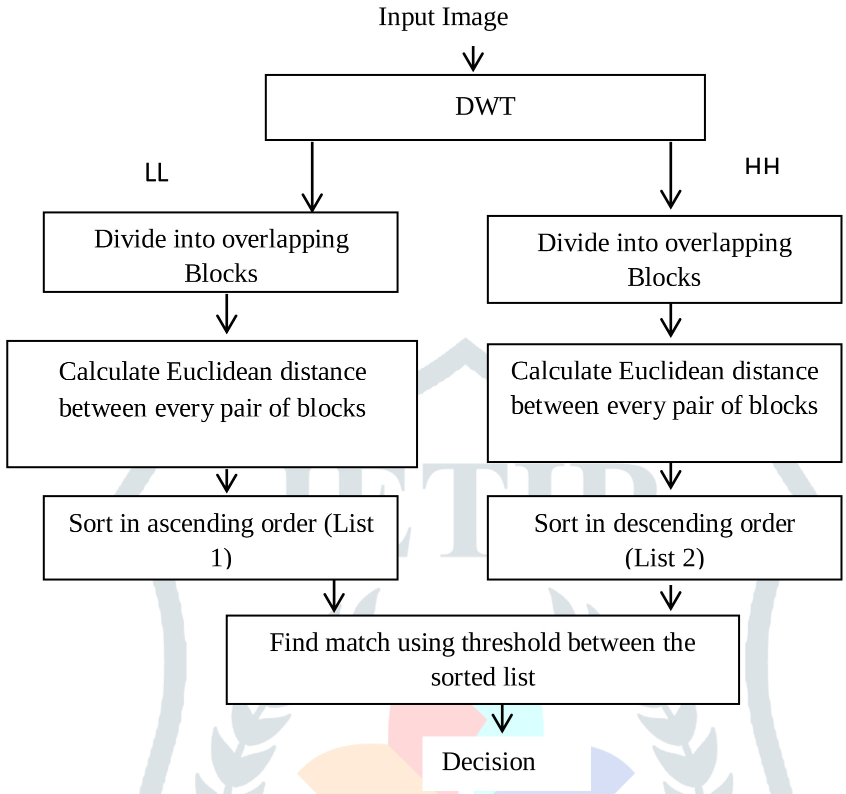

Input Image
DWT
LL
HH
Divide into overlapping Blocks
Divide into overlapping Blocks
Calculate Euclidean distance between every pair of blocks
Calculate Euclidean distance between every pair of blocks
Sort in ascending order (List 1)
Sort in descending order (List 2)
Find match using threshold between the sorted list
Decision

Figure 1: Block based method for CMFD

6. The set of probable grades is scrutinized another time and Euclidean distance d is intended among coordinates of blocks of every pair conferring to equation (5). Altogether pairs with distance d lesser than predefined threshold T_d are unconcerned from the set of potential consequences. Threshold T_d is frequently demarcated conferring to a selection of block dimensions (for specimen, k×b, where k is certain slight positive constant) to eradicate all close by blocks (it can be presumed that a block is progressed more than Td pixels). Subsequently these pace only alike pairs of blocks that are not close by to each other are retained as potential matches.

$$d=\sqrt{(x_{f1}-x_{f2})^2+(y_{f1}-y_{f2})^2}$$
(5)

7. The recognition image is engendered by coloration all enduring pairs of blocks. Some meek post-processing can be pragmatic to take away insignificant, deceitfully perceived zones in the image (for specimen, morphological opening).

- **LBP Operator**

LBP operator is an operative texture depiction operator. It has been efficaciously smeared in image processing zones these ages. Subsequent, familiarize how to evaluate the LBP value. In 3 × 3 window, the gray value of the midpoint point of the frame as a threshold value, supplementary pixels in the frame do binarized handling, engenders an 8-bit binary string. Then, conferring to the dissimilar locations of the pixels, acquire the LBP value of the frame by weighted summing. It can be figured by

$$LBP=\sum_{i=0}^{7}s(g_i-g_c)2^i, where\, s(x)=\begin{cases}1 & x\ge 0\\ 0 & x<0\end{cases} \qquad (6)$$

Here g_c is the center pixel of the frame, g_i symbolizes adjoining pixels. In general the direction of the neighboring pixels is underway by the pixel to the right of the center pixel, counterclockwise patent. The LBP value can imitate the texture evidence for the province [11]. LBP can be extended to a circular neighborhood. Expending (P, R) to designate the neighborhood, where P symbolizes the number of sampling points, R is the radius of the neighborhood. The gray values of neighbors which do not fall accurately in the center of pixels are projected by exclamation.

II. LITERATURE REVIEW

A procedure to perceive Copy-Move Forgery to provision image forgery detection has been proposed by Pandey et al. (2014) [12]. The outcomes were recorded utilizing three distinctive picture includes to be specific SURF, HOG and SIFT among which SIFT gave best outcomes as exactness and accuracy. By applying same technique on various highlights they have demonstrated that how one component gives better outcomes in contrast with others. In the wake of considering half and half highlights (SURF-HOG or SIFT-HOG), they are showing signs of improvement result for CMFD in contrast with SIFT or SURF or when HOG is utilized alone. Looking forward, Xiamu et al. (2016) [13] propose a feature point-based copy-move forgery detection method that is equipped for managing the imitations occurred at smooth, particularly little smooth districts. For highlight location, they exhibit a two-arrange include point identification plan to get adequate component point scope for both finished and smooth locales in a suspicious picture. They utilize the MROGH descriptor as highlight descriptor for customary areas in the picture, for the little smooth locales, they abuse include combination to upgrade the discriminative energy of the component descriptor. Their technique separates the highlights in a denser way, in this way the running time of our strategy is substantially higher than of the SIFT and SURF-based strategies. As far as identification capacity, their technique beats the cutting edge strategies for plain duplicate move recognition; moreover, the power against jpeg pressure and pivot are likewise tasteful. Their strategy can oppose direct level of scaling, added substance clamor and joined impacts, however the execution decreases quickly when these assaults are solid, because of the shakiness of the Harris Corner Detector under these conditions. The use of thick intrigue focuses or relative covariant element indicators may help. Moreover, Emam et al. (2016) [14] planned an effectual scheme meant for copy-move forgery detection that can distinguish tampering and localize the disagreed region in a digital image. Rather than utilizing the thorough piece coordinating technique, ANNs is gathered by territory touchy hashing LSH. To show signs of improvement recognition comes about, morphological activities are connected to evacuate little openings and dispose of detached pixels. Our technique can identify the copied locales of altered pictures even affected by geometric changes, for example, pivot, scaling, commotion expansion, and JPEG pressure. In the work of, Qingxiao et al. (2017) [15] propose a copy-move forgery detection technique based on Convolutional Kernel Network. The fundamental commitments can be closed as takes after: the CKN appropriation in duplicate move phony discovery and GPU-based CKN remaking, the division based keypoint dispersion (SKPD) technique and GPU-based versatile over division (COB). Accordingly, XiuLi et al. (2018) [16] propose an innovative multi-scale feature extraction and adaptive matching method to notice the copymove image forgery. In the proposed plot, to begin with, they section the host image by SLIC in multiscale, to create multi-scale patches; at that point they apply SIFT to patches in every one of the scales, to remove highlight focuses. Next, the Adaptive Patch Matching calculation is in this manner proposed for finding the coordinating which can demonstrate the suspicious fashioned locales in each scale. Lastly, the suspicious districts in all scales are combined and some morphological activities are connected to create the recognized imitation locales. As a rule, they have four fundamental commitments in the proposed conspire: 1) they supplant the covering squares of normal shape in conventional fraud location calculations, with singular unpredictable patches, which can better parcel the host images into non-covering pieces. 2) They fragment the host image into patches in different scales, from which the component focuses are separated individually. The proposed multi-scale include extraction strategy can separate more precise component focuses. 3) Instead of falsely setting the fix coordinating limit ahead of time, they propose to adaptively ascertain the coordinating edge for better component acknowledgment. What's more, 4) amid the post-preparing, they propose to utilize the predefined little superpixels to supplant the coordinated keypoints and they apply some morphology tasks into the consolidated locales to produce all the more precisely identified fabrication districts.

III. IMPLEMENTATION

In this work, block based feature extraction and matching process is used. At first edge detection is carried out to get the high entropy pixels in the image so that matching process is carried out only for high entropy pixel blocks.

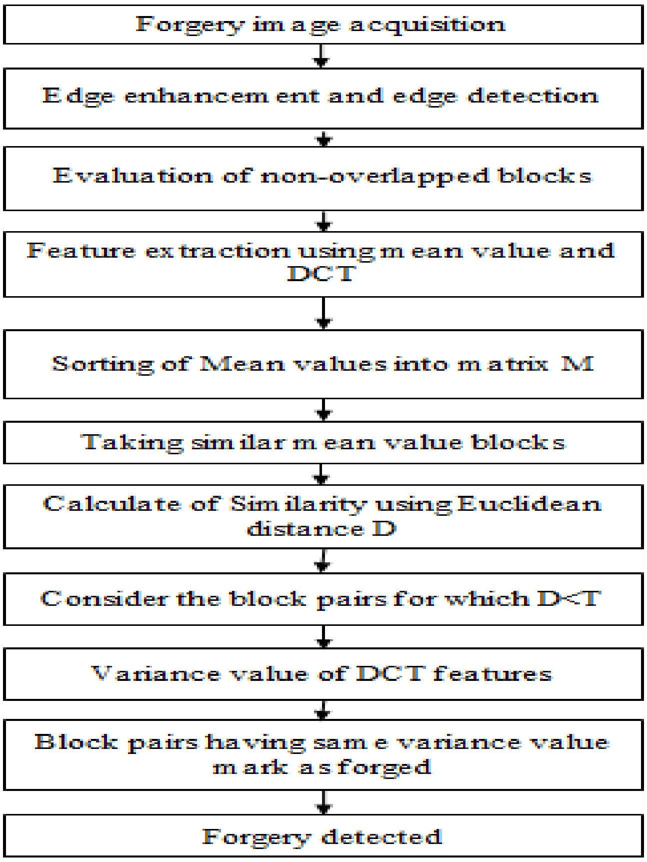

Forgery image acquisition
Edge enhancement and edge detection
Evaluation of non-overlapped blocks
Feature extraction using mean value and DCT
Sorting of Mean values into matrix M
Taking similar mean value blocks
Calculate of Similarity using Euclidean distance D
Consider the block pairs for which D<T
Variance value of DCT features
Block pairs having same variance value mark as forged
Forgery detected

Figure 2:Flowchart of the proposed method

Then feature extraction is carried out by converting image into overlapped blocks and mean and DCT features are extracted. Then mean values are put into a matrix and corresponding blocks are noted. Then mean value matrix is sorted in order to match the blocks with similar mean values. In matching process, variance of DCT features is used for similarity measure and forgery detection. The flowchart of the methodology has been shown in Figure 2.

IV. PERFORMANCE EVALUATION

For performance evaluation of the proposed method, sensitivity, specificity and accuracy 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 DCT texture algorithms. After that forged pixels has been calculated. The classification accuracy is the extent to which the classifier is able to correctly classify the examplars 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. Test data for evaluation consists of a set of 512 × 512 RGB images, taken from the CoMoFoDDatabase.Accuracy is characterized as the proportionof the quantity of effectively arranged examples (TP and TN) to the aggregate number of examples (species) grouped which is given in eq (7).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \text{......(7)}$$

The sensitivity of a classifier is the fraction of the image samples correctly classified as that specific species class. It is defined by equation (2) below:

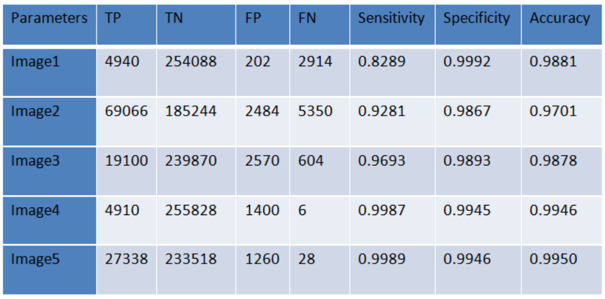
$$Se = \frac{TP}{TP+FN} \text{..........(8)}$$

The specificity is the fraction of normal pixels correctly classified as normal class. It is also called selectivity.

$$Sp = \frac{TN}{TN+FP} \text{........(9)}$$

The results for the actual pixel location using ground truth images and that of resulted outputs has been described with above parameters.

Table 1: Sensitivity, specificity and accuracy parameters for the tested images



Parameters	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy
Image1	4940	254088	202	2914	0.8289	0.9992	0.9881
Image2	69066	185244	2484	5350	0.9281	0.9867	0.9701
Image3	19100	239870	2570	604	0.9693	0.9893	0.9878
Image4	4910	255828	1400	6	0.9987	0.9945	0.9946
Image5	27338	233518	1260	28	0.9989	0.9946	0.9950

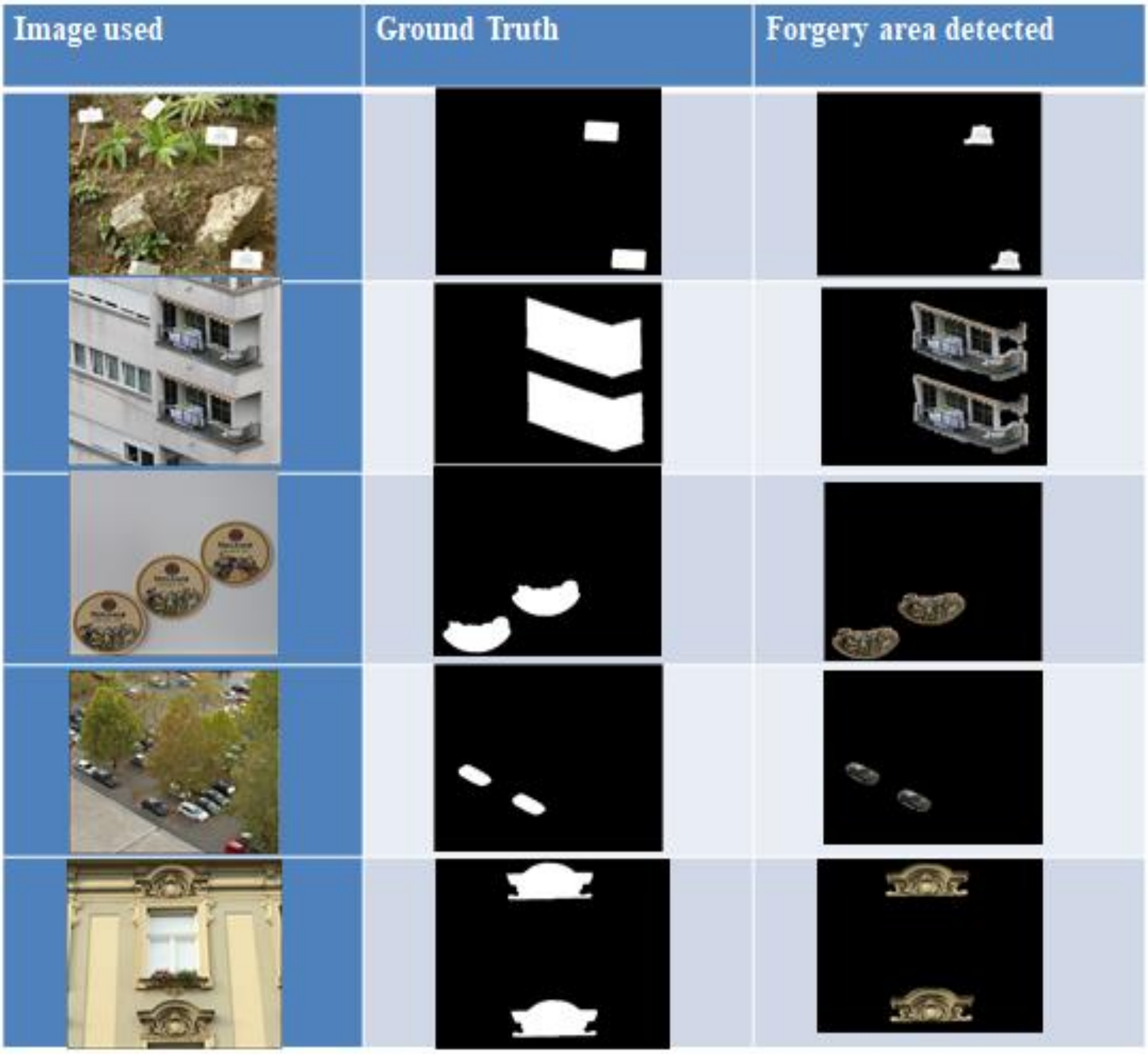


Figure 3:Results of forgery detection system

Figure 3 shows the experimental results in which column one gives original forged images. Column two is ground truth images where forgery is induced. Column 3 results show the forgery detection results by the proposed method.

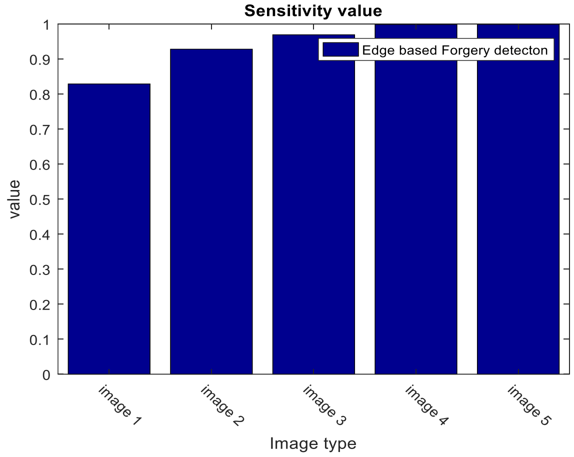


Figure 4: Sensitivity value for the copy move forged pixels detected by proposed method

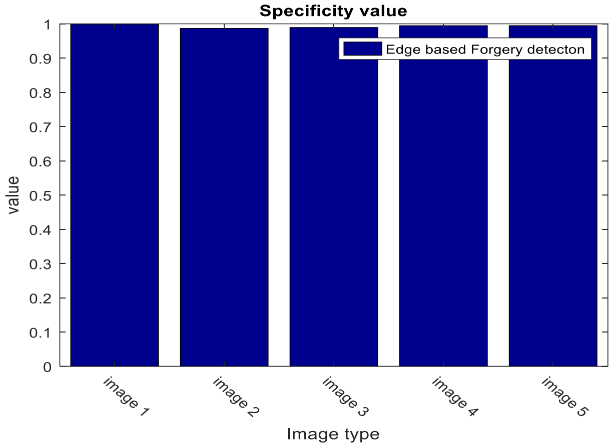


Figure 5: Specificity value for the copy move forged pixels detected by proposed method

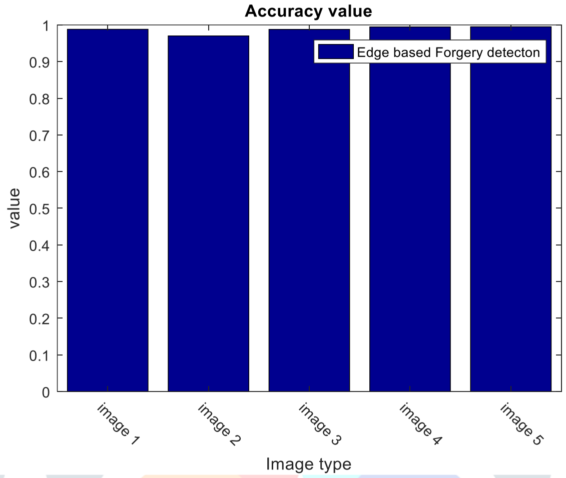


Figure 6: Accuracy value for the copy move forged pixels detected by proposed method

Proposed method works well for copy-move forgery only but it is not efficient when there is blurring or contrast change of the pixels of the moved pixels. In this method, Euclidian distance is used as a threshold when two blocks are compared along with matching of edge pixel blocks only which decreases computation time of the algorithm. Experimental results show the accuracy of tested images comes in the range of 97 to99 per cent.

V. CONCLUSION

In this paper, an efficient block-based method is presented for CMFD. First edge enhancement and edge detection is used to generate a binary image containing edge and non-edge areas, Purpose of edge detection is to reduce the computation time of the algorithm as most of the existed CCMFD algorithms have large computation time. It enables to match only those blocks which come as edge pixel blocks in binary image. Secondly mean and DCT features are used in which mean values of all the overlapped blocks are calculated and sorted and then similarity matching is carried out for those blocks which have similar mean values. Forged areas are marked for those blocks which have similar variance values of the DCT features. Experimental results show high accuracy of forgery detection. Proposed method can amend to include rotation invariant forgery detection as it fails for the rotated copy move blocks.

References

[1] Kirchner M (2012). Notes on digital image forensics and counter forensics. Forensic analysis of Re-sampled digital signals,1-97

[2] WarbheAD, Dharaskar RV, Thakare VM (2016) Computationally efficient digital image forensic method for image authentication. Procedia Computer Science, 78:464-470.

[3] Osamah MAQ, KhooBE (2013) Passive detection of copy-move forgery in digital images: state-of-the-art. Forensic Science International, 231:284-295.

[4] FridrichJ, SoukalD, LukasJ (2003) Detection of copy-move forgery in digital images, Proceedings of Digital Forensic Research Workshop, 3:55-61.

[5] AmeriniI,BallanL,CaldelliR, BimboAD and SerraG (2011) A SIFT-based forensic method for copy-move attack detection and transformation recovery. IEEE Transactions on Information Forensics and Security, 6:1099–1110

[6] ShivakumarBL,and BabooS (2011) Detection of region duplication forgery in digital images using surf. International Journal of Computer Science Issues, 8:199–205

[7] BasharM, NodaK, OhnishiN,and MoriK (2010) Exploring duplicated regions in natural images. IEEE Transactions on Image Processing

[8] PopescuA and FaridH (2004) Exposing digital forgeries by detecting duplicated image regions. Tech. rep. tr2004-515, Dartmouth College

[9] RyuSJ, LeeMJ and Lee,HK (2010) Detection of copy-rotate-move forgery using zernike moments. International Workshop on Information Hiding: 51–65

[10] WangJ, LiuG, LiH, DaiY and WangZ (2009) Detection of image region duplication forgery using model with circle blocks. International Conference on Multimedia Information Networking and Security: 25-29

[11] ZhengN, WangY and MingX(2013) A LBP-Based Method for Detecting Copy-Move Forgery with Rotation. Multimedia and Ubiquitous Engineering: 261-267

[12] PandeyRC,AgrawalR, SinghSK andShuklaKK (2014) Passive Copy Move Forgery Detection Using SURF, HOG and SIFT Features. Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA): 659-666

[13] YuL, HanQ,andNiuX (2016) Feature point-based copy-move forgery detection: covering the non-textured areas. Multimedia Tools and Applications, 75:1159–1176

[14] EmamM, HanQ,andNiuX (2016) PCET based copy-move forgery detection in images under geometric transforms. Multimedia Tools and Applications, 75:11513–11527

[15] LiuY, GuanQ,and ZhaoX (2017) Copy-move forgery detection based on convolutional kernel network. Multimedia Tools and Applications: 1–25

[16] XiuLB, PunCM and YuanXC (2018) Multi-scale feature extraction and adaptive matching for copy-move forgery detection. Multimedia Tools and Applications, 77:363–385