ISSN2394-3777 (Print)
ISSN2394-3785 (Online)
Available online atwww.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 12, April 2018

Robust Image Forgery Detection and Classification in Copy-Move Using SVM

Anitha Susan Varkey, Lekha R Nair

Electronics and Communication, College of Engineering,
*KalloopparaPathanamthitta,India*anithasusan888@gmail.com lekharnair@gmail.com

***Abstract*— An effective copy-move forgery detection scheme and classification is proposed in this paper. The proposed scheme integrates both block-based and keypoint-based forgery detection methods. First, the Simple Linear Iterative Clustering is used to segment the image into non-overlapping and irregular blocks called image blocks. Then, Scale Invariant Feature Transform is applied in each block to extract the feature points as block features. Subsequently, feature matching is performed to detect the copy-move forged region. The experimental results show that the proposed forgery detection scheme is effective in detecting the tampered region and Support Vector Machine classifier provides better classification result than Artificial Neural Network with an accuracy of 95.33%.**

Key words: Copy-Move Forgery Detection, SLIC, SIFT, SVM classifier.

I.INTRODUCTION

With the advancement and availability of powerful image processing software tools and computer technology, it is very easy to manipulate the digital images. So, it is very essential to determine the authenticity, integrity, reliability and the origin of digital images.

Image forgery may be performed by a either to hide the truth or to enhance the visual effect of the image. Normal people might neglect this malicious operation when the forger deliberately covers the tampering trace. So, an effective Copy Move Forgery Detection (CMFD) method to automatically point out the duplicate regions in the image is needed. Currently, CMFD is becoming one of the most important and popular digital forensic techniques.

As images can be used in very important fields such as medicine, astronomy, surveillance, etc. it is necessary to recognize this type of doctored images. In recent years, various algorithms are proposed to detect forgery in images. They can be classified into two categories: active and passive or blind algorithms. In active approach, watermark or digital signature is embedded into the image. In contrast, the passive techniques do not need to embed any watermark in the image or no digital signature is required to be generated. In passive approaches detection of tampered objects is done in forged

images without need of original image watermark. Digital image forgery can be classified mainly into two types:Copy-move (cloning) forgery and Copy-paste (splicing) forgery: In Copy-move type of forgery, image is duplicated by copying a part of image and pasting it into another part of the same image. There are at least two similar regions in a tampered region due to region manipulation whereas in splicing type of forgery fragments of same or different images are combined to produce another single forged image without further post processing such as smoothing of boundaries among different fragments.

Of the existing types, the most commonly used forgery method is copy –move forgery, in which a part of an image is copied to another location in the same image. The purpose of this kind of forgery is to hide an undesired object or to increase the number of objects in an image. According to the existing methods, the copy move forgery detection methods can be categorized into two algorithms: block-based algorithms and feature keypoint based algorithms. Fridich et al.[1]divided the input image into overlapping segments and then features are extracted from the blocks using DCT .Then, these blocks were sorted lexicographically to find the similarity. Later, Popescu and Farid[2] used Principal Component Analysis(PCA)to yield a reduced dimensions. Luo et al. [3] proposed an approach based on the idea of using RGB color components and direction information as block features. Li et al. [4] used Discrete Wavelet Transform (DWT) to decompose the image into four sub bands then Singular Value Decomposition (SVD) was applied to obtain a reduced dimension representation. The matrix of SV vectors is then lexicographically sorted to detect the tampered region. Bayram et al. [5]

ISSN2394-3777 (Print)
ISSN2394-3785 (Online)
Available online atwww.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 12, April 2018

presented the Fourier- Mellin Transform (FMT) to obtain features. Ryu et al.[6]applied Zernike moments as block features. BravoSolorio and nandi [7] used information entropy to extract the image features.

An alternative to the block-based methods are keypoint based forgery detection methods. H.Huang et al. [8] proposed a method to detect duplicated region using Scale-Invariant Feature Transform (SIFT). The SIFT was applied to the input images to extract feature points, which were then matched to one another. The sets of corresponding SIFT feature points were defined as the forgery region, when the value of the shift vector exceeded the threshold. In [9, 10] the Speeded up Robust Features (SURF) is used to extract features. Although, all these above algorithms are effective in forgery detection, they have drawbacks: 1) the host image is segmented into overlapping rectangular blocks, where it is computationally expensive as the size of the image increases: 2) the methods cannot detect geometrical transformations of the forgery region.To address the above-mentioned problems, a novel copymove forgery detection scheme is proposed.

II.PROPOSED CMFD SYSTEM

Fig.1 shows the framework of the proposed image forgery detection scheme.

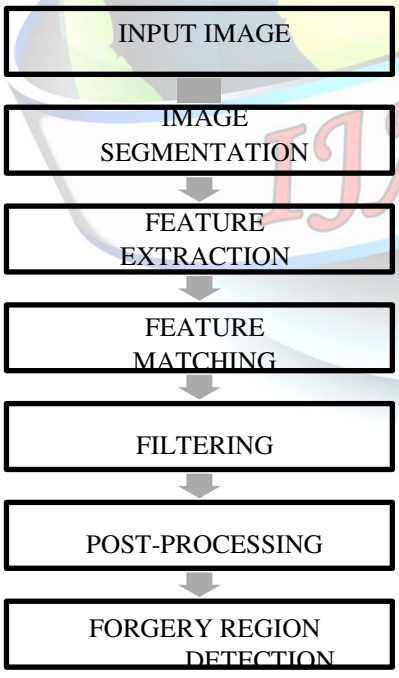

Figure.1 Framework of the proposed copy -move forgery detection scheme

First, Simple Linear Iterative Clustering (SLIC) [11] is proposed to segment the host image into non-overlapping and irregular blocks called Image Blocks (IB). Then, the Scale Invariant Feature Transform (SIFT) [10] is applied in each block to extract the SIFT feature points as Block Features (BF). Subsequently, the block features are matched using lexicographical sorting.

A.Image Segmentation

Of the existing block-based forgery detection schemes, the host image was usually divided into overlapping regular blocks, with the block size being defined and fixed.Also,the detected regions are always composed of regular blocks, which cannot represent the accurate forgery region well; as a consequence, the recall rate of the block-based methods is always very low. Moreover, when the size of the host images increases, the matching computation of the overlapping blocks will be much more expensive. To address these problems, the host image is segmented into non-overlapping regions of irregular shape as image blocks, afterward, the forgery regions can be detected by matching those non-overlapping and irregular regions. The host image is segmented into nonoverlapping regions of irregular shape and because the superpixels are perceptually meaningful atomic regions that can be obtained by over-segmentation, the simple linear iterative clustering (SLIC) algorithm [11] is employed to segment the host image into meaningful irregular superpixels, as individual blocks. SLIC is a simple and efficient method to decompose an image in visually homogeneous regions.

SLIC takes two parameters: the nominal size of the regions (superpixels) *regionsize* and the strength of the spatial regularization *regularizer*. The image is first divided into a grid with step *regionsize*.The center of each grid tile is then used to initialize a corresponding k-means (up to a small shift to avoid image edges). Finally, the k-means centers and clusters are refined by using the Lloyd algorithm, yielding segmenting the image. The parameter *regularizer* sets the trade-off between clustering appearance and spatial regularization.

The SLIC algorithm adapts a k-means clustering approach to efficiently generate the superpixels, and it adheres to the boundaries very well. Using the SLIC segmentation method, the non-overlapping segmentation can decrease the computational expenses compared with the overlapping blocking; furthermore, in most cases, the irregular and meaningful regions can represent the forgery region better than the regular blocks. vlFeatsoftware [12] is used to segment the input image. *regularizer* is used to control the regularization of the patches. It is set to 0.8 for all the images. The other parameter *region*

ISSN2394-3777 (Print)
ISSN2394-3785 (Online)
Available online atwww.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 12, April 2018

size is related to the number of segmentation patches. Its value hence should be adaptive to the image size.

B.Feature Extraction

The traditional block-based forgery detection methods extracted features of the same length as the block features or directly used the pixels of the image block as the block features; however, those features, leaves out the location information.. Therefore, feature points are extracted from each image block as block features, and the feature points should be robust to various distortions, such as image scaling, rotation, and JPEG compression.. The feature points extracted by SIFT [14] and SURF [13] were proven to be robust against common image processing operations such as rotation, scale, blurring, and compression. The SIFT features are invariant to image scale and rotation. vlFeat(version 0.9.18) [12] software is used to detect and describe the keypoints. The default setting of vlFeat for keypoints detection and description, namely SIFT is employed in the implementation.

C.Feature Matching

High similarity between two feature points is indicated as a clue for a tampered region. Feature points are sorted in ascending order for identifying similar feature vectors. A matrix of feature vectors is formed so that every feature vector becomes a row in the matrix. This matrix is then row-wise sorted and the most similar features appear in consecutive rows.

D.Filtering

Filtering is used to reduce the probability of false matches. Neighboring pixels often have similar intensities, which can lead to false forgery detection. So, Euclidean distance criteria and similarity criteria using correlation coefficient are used to filter out weak matches.

.Post-Processing

This step is used to preserve matches that exhibit a common behavior. The outliers are removed by imposing standard deviation.

III.CLASSIFICATION USING SVM

For classifying original and forged images first important issue is which color model should be used. RGB, Lab, YCbCr, HSV, etc, are different color models used in which RGB is the most common model for image visualization but it is not suitable for image forgery detection. It is found that YCbCr color model gives better classification accuracy. YCbCr Contain the color information and they can be highly compressed. In YCbCr model, the Y component is luminance component; Cb and Cr are the blue difference and red difference chrominance components, respectively. So we convert RGB images into YCbCr images. LBP is a powerful feature for texture classification which has been used widely.

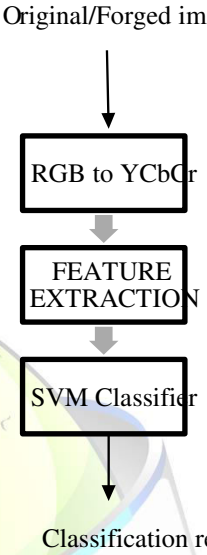

Fig.2 Block diagram of proposed copy-move forgery classification

The powerful STD filter is used to highlight the abrupt changes in the image. This filter segments the image and highlights the important details of the image. The STD filter replaces each pixel value in an image with the standard deviation of its neighbors, including itself. Further, RI-LBP operator is applied to extract the highlighted internal statistical information. The RI-LBP operator is invariant to image rotations and have high descriptive ability. Fig. 2 shows proposed copy-move forgery classification. For feature extraction, Rotation Invariant-Local Binary Pattern (RI-LBP) is utilized.

A. Local Binary Pattern

LBP operator computes the LBP code by thresholding value of 8 neighbors. If the neighbor's pixel value is less than the center, it will hold binary digit '0'otherwise '1'.Then neighbors' binary digits are put together to build binary code. The LBP code will be the decimal vale of that binary code.

B.Classifier

Support Vector Machine (SVM) and Artificial Neural Network (ANN) will be trained and used as classifier to closely detect whether the given image is forged or not.

IV.RESULTSANDDISCUSSION

ISSN2394-3777 (Print)
ISSN2394-3785 (Online)
Available online atwww.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 12, April 2018

The database MICC-F220 [15] is used to test the proposed method. The database MICC F220 consists of 220 in which half of the images are tampered.

The performance of the algorithm in detecting copy-move forgery is quantified using the following definitions: True positive (TP): forged images detected as forged, False negative (FN): forged images detected as original, False positive (FP): original images detected as forged, True negative (TN): original images detected as original

Total 30 images were tested, out of which 22 were forged images (true matches) and 8 were genuine images (true nonmatches). Out of which 22 were forged images (true matches) and 8 were genuine images (true non-matches). Out of the 22 forged images, the algorithm detected 20 images correctly as forged. Out of the 8 genuine images, the algorithm falsely detected 4 images as forged.
The accuracy, precision and recall for proposed method is defined as below:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

Two characteristics *precision* and *recall* are used to evaluate the performance of the proposed forgery detection scheme.
Precision is the probability that the detected regions are relevant. *Recall* is the probability that the relevant regions are detected. In general, a higher *precision* and a higher *recall* indicate superior performance.

ABLE I

PERFORMANCE COMPARISON OF CMFD USING SVD AND SIFT

Image Features	Performance (%)		
	Accuracy	Precision	Recall
SVD	63.33	76.19	72.72
SIFT	80	83.33	90.9

Table I shows the experimental results of the proposed method. It is observed that CMFD using SIFT features provides best result.

For classification, COMOFOD database [16] is used. Images of COMOFOD databases are divided into two sets in which 70% images are used for training and remaining images are used for testing. The accuracy represents the percentage of the correctly classified images. Total 150 images were tested, out of which 100 were forged image and 50 were genuine images.

The performance of the algorithm in classifying copy-move forgery is quantified using the following definitions: True positive (TP): forged images classified as forged, False negative (FN): forged images classified as original, False positive (FP): original images classified as forged, True negative (TN): original images classified as original.

TABLE II

PERFORMANCE COMPARISON OF CMFD USING SVM AND ANNCLASSIFIER

Classifier	Performance (%)		
	Accuracy	Precision	Recall
SVM	95.33	97.93	95
ANN	86	93.40	85

From Table II, it can be observed that SVM classifier determines better classification result than ANN classifier. It can be easily observed that the recall of the SVM classifier is much better than that of the ANN classifier.

V.CONCLUSION

It is very challenging to detect digital forgery images created with copy-move operations. In this, a novelcopymove forgery detection scheme and classification is proposed. The Superpixel Segmentation algorithm is proposed to segment the host image into non-overlapping and irregular blocks adaptively according to the given host image. Then, in each block, the feature points are extracted as block features, and the block features are sorted in ascending order to detect the tampered region. Lastly, filtering and post-processing is applied to detect more accurately. The proposed method is evaluated on a number of original and images. Experimental results showed that the method is quite

ISSN2394-3777 (Print)
ISSN2394-3785 (Online)
Available online atwww.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 12, April 2018

attractive with an accuracy of 80% and with a smallest false negative rate, which means the proposed scheme is good at detecting the tampered images. For classification SVM classifier provides better result with an accuracy of 95.33%.

Future work could focus on applying the proposed forgery detection scheme on other types of forgery, such as splicing or other types of media, for example, video and audio.

ACKNOWLEDGMENTS

The authors would like to thank the Principal and HOD of College of Engineering, Kallooppara and coordinators for their support and assistance.

REFERENCES

[1] J. Fridrich, B. D. Soukal, and A. J. Luk, "Detection of copy-move forgery in digital images," in *Proc. Digital Forensic Research Workshop,* 2003.

[2] A. C. Popescu and H. Farid, "Exposing digital forgeries by detecting duplicated image regions," *Dept. Comput. Sci., Dartmouth College, Tech. Rep. TR2004-515*, 2004.

[3] W. Luo, J. Huang, and G. Qiu, "Robust detection of region-duplication forgery in digital image," *in Pattern Recognition, 2006. ICPR 2006. 18th International Conference on,* 2006, pp. 746-749.

[4] G. Li, Q. Wu, D. Tu, and S. Sun, "A sorted neighborhood approach for detecting duplicated regions in image forgeries based on DWT and SVD," in *Multimedia and Expo, 2007 IEEE International Conference on,* 2007, pp. 1750-1753.

[5] S. Bayram, H. T. Sencar, and N. Memon, "An efficient and robust method for detecting copy-move forgery," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on,* 2009, pp. 1053-1056.

[6] S. Ryu, M. Lee, and H. Lee, "Detection of copy-rotate-move forgery using Zernike moments*," in Information Hiding,* 2010, pp. 51-65.

[7] S. Bravo-Solorio and A. K. Nandi, "Exposing duplicated regions affected by reflection, rotation and scaling," *in Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on,* 2011, pp. 1880-1883.

[8] H. Huang, W. Guo, and Y. Zhang, "Detection of copy-move forgery in digital images using SIFT algorithm," *in Computational Intelligence and Industrial Application, 2008. PACIIA'08. Pacific-Asia Workshop on,* 2008, pp. 272-276.

[9] X. Bo, W. Junwen, L. Guangjie, and D. Yuewei, "Image copy-move forgery detection based on SURF," *in Multimedia Information Networking and Security (MINES), 2010 International Conference on,* 2010, pp. 889-892.

[10] B. Shivakumar and L. D. S. S. Baboo, "Detection of region duplication forgery in digital images using SURF*," IJCSI International Journal of Computer Science Issues*, vol. 8, 2011.

[11] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans Pattern Anal Mach Intell*, vol. 34, pp. 2274-82, Nov 2012.

[12] A. Vedaldi and B. Fulkerson.2008. "VLFeat: An open and portable library of computer vision algorithms," http://www.vlfeat.org.

[13] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *Computer Vision–ECCV 2006,*ed: Springer, 2006, pp. 404-417.

[14] D. G. Lowe, "Object recognition from local scale-invariant features," in Computer vision, 1999. The proceedings of the seventh IEEE international conference on, 1999, pp. 1150-1157.

[15] I. Amerini, L. Ballan, R. Caldelli, A. Del Bimbo, and G. Serra, "A siftbased forensic method for copy–move attack detection and transformation recovery," *Information Forensics and Security, IEEE Transactions on*, vol. 6, pp. 1099-1110, 2011. [16]http://www.vcl.fer.hr/comofod.