A Research on Video Forgery Detection using Machine Learning

Gaikwad Kanchan[1], Kandalkar Ambika[2], Khokale Yogita[3], Bhagwat Archana[4], Chandgude Amar[5]

[1,2,3,4]Student, Dept. of Computer Engineering, S.N.D. college of Engineering and Research Center Yeola, Maharashtra, India
[5]Professor, Dept. of Computer Engineering, S.N.D. college of Engineering and Research Center Yeola, Maharashtra, India

---***---

Abstract – *Region duplication is a very easy and effective method to create digital image forgeries, where a continuous portion of pixels in an image are copied and pasted to a different location in the same image. Now a days Video and image copy-move forgery detection is one of the major hot topic in multimedia forensics to protect digital videos and images from malicious use. Number of technique have been presented through analyzing the side effect caused by copy–move operation. In this paper, we propose a novel approach to detect copy–move image forgery. And also coarse-to-fine detection strategy based on optical flow (OF) and stable parameters is designed to detect. The detected image is initially divided into overlapping blocks. After the creation of the overlapped blocks, the feature extraction technique is applied on the image to extract the features from specific block of the image* to identify duplicate blocks of image.

Key Words- ***OPTICAL FLOW(OF); COPY MOVE; GLCM FEATURE; NOVEL APPROACH; MACHINE LEARNING; SUPPORT VECTOR MACHINE.***

1. INTRODUCTION

There is a significant role of digital images and videos in our daily life. Video is nothing but the collection of images or frames. However, image tampering has become very easy by using powerful software. Videos or images can be scanned using this software and tampered without any doubt. Now a day image authenticity is a big concern. Image forgeries may have many types- such as copy-move forgery, splicing and many more. Copy-move forgery is nothing but the copying content of another image and pasting it into same image which we want to forged. The two main types of image forensic techniques are to verify the integrity and authenticity of manipulated image. One is active forensic method and another is passive forensic method. In active methods Watermarking and steganography are two techniques which are used to insert authentic information into the image. In the authenticity of an image, then prior embedded authentication information is recalled to prove the authenticity of that image. However, embedding authentication of information to an image is very reliable. Only an authenticated users are allowed to do it or at the time of creating the image, authentic information could be embedded as well. But requirement of special cameras and multiple steps processing of the digital image are two main limitations which made this technique less efficient. To avoid these limitations, passive forensic techniques are utilize image forgery without requiring detailed previous Information. The most widely used method to make forged image is copy-move forgery. It refers to copy one part from another image, and paste it inside the same image. Sometime before pasting the copied regions, various processing operations like scale, rotation, blurring, intensifying or JPEG compression may be applied.

2. LITERATURE SURVEY

A. Block-based Image Forgery Detection

In block-based method, input image size of M x N is segmented into overlapping blocks size of z x z resulting into overlapping blocks, L = (M-z+1) x (N-z+1). A few features are extricated from each block. Distinctive features are extracted by applying different feature extraction technique such as DCT (Discrete Cosine Transform) [9], DWT (Discrete Wavelet Transform) [10], DFT(Discrete Fourier Transform) [10], PCA (Principal Component Analysis) [12] [13], SVD (Singular Value Decomposition) [14][15], and ZMs (Zernike Moments) [16]. Then, a comparison is done based on blocks features similarity and distance. After finding the most matched or similar features of block, copy-move region is identified and this region is localized. Sheng et al. [9] proposed forgery detection algorithm using block-based method. This uses DCT and circle blocking technique for extracting features of the image. Finally, the image which contains singularities within lines is presented by computing ridgelet transformation. Robustness against JPEG compression is the most significant feature of this method. Cao et al. [17] followed block-based method to detect tampered region where DCT feature extraction technique is applied. DCT is used to divide sub blocks to extract key features by producing quantized coefficients. Threshold values are set to match features between closest similar image blocks. This method shows less computational complexity compared to existing methods [23-[18] because of reducing dimensions of feature vector. Later, similar method and feature extraction technique used by Huang et al. [19]. The big difference in the result with Cao et al.'s DCT-based method [17], because of reducing the false matches rate. Due to low false matches rate, this method becomes powerful against noise and blurring. However, it is not robust for rotation attack and cannot detect multiple forgery. M. Bashar et al.,[20] developed more efficient forgery detection method based on DWT and kernel PCA (KPCA) features. Actuals images have been used as a dataset in this method. As a consequence of quantitative analysis considering noiseless and uncompressed factor, it is found that the DWT performs well than KPCA in terms of features. On the other hand, in noise and JPEG compression domain, KPCA-based features perform better than DWT. The method shows robustness against noise and JPEG compression attack hence this method takes too much time and not robust against scaling. It cannot detect multiple forgery images.

B. Key point-based Image Forgery Detection Method

It is different from block-based methods, features are extracted in key point-based method from the image without any type of segmentation. Extracted features from every key point are compared to find similarities between them. Finally, based on the calculation of matched features, image forgery is detected. SIFT and SURF (Speeded Up Robust Features) are two main key points based feature extraction methods. Somayeh Sadeghi et al., [21] and Diaa M. uliyan et al., [22] worked on key points based technique (e.g. SIFT). Sadeghi et al., proposed SIFT to extract features and searched for similar features based on their Euclidean distance. Both methods are robust against several post-processing attacks; including scale, noise, rotation and JPEG compression. However, inability to detect small forged areas and performance of detection and localization for those forged areas are also questionable. In [22], primary approach of Uliyan et al., was to detect image regions by using Statistical Region Merging (SRM) Segmentation algorithm. Then, the experiment proceeded with applying Angular Radial Partitioning (ARP) and Harris Corner detection method on the image region. Finally, forged regions were detected based on matched key points. The method showed less robustness against forged regions with blurring and illumination attacks. Moreover, it shows different result for same image with different resolution. The major drawbacks of the previously mentioned conventional techniques are either not powerful against all post processing attacks or have high computation time. Therefore, keeping up the low computational time is the most important robustness challenge. To tackle this issue, a new copy-move forgery detection method is proposed where region wise image segmentation is done. Gabor filters are used to extract image features. Afterwards, K Means clustering, and Euclidean distance calculation facilitated to detect forged region from the suspicious image. Reducing the false matching rate is the most significant task to exhibit the proposed method as more viable compared to conventional methods.

3. PROPOSED SYSTEM

A. Video Input:

Video forensic has become an important area of research in the last decade. System will accept video as an input. Justified format of video should be given as an input to get processed.

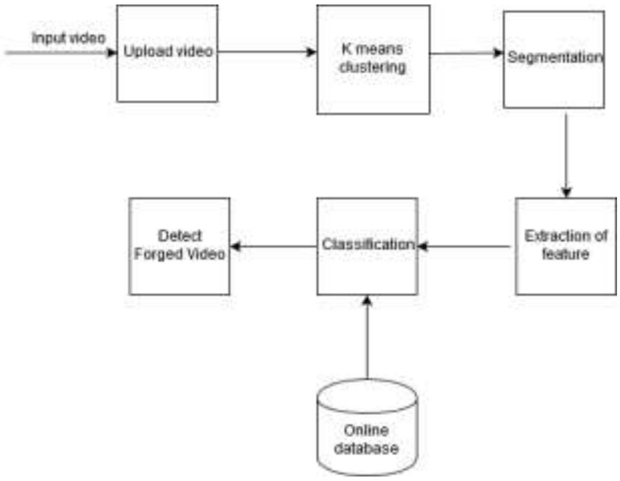

FIG 1: BLOCK DIAGRAM

B. Video Parsing/Segmentation:

Input video is been accepted and done parsing based on fps. These frames will be temporary stored in backend for further processing and feature extraction.

C. K-Means Clustering

K-means clustering is a technique for quantizing vectors. This method divides the image into k segments, each containing mutually exclave data. This is a common method when it comes to pattern recognition and machine learning. One of the segmented images is chosen on the basis of the information contained in it. To determine this, the features of each segment are calculated and the segment with the highest mean is chosen. The features of the segmented image are then compared with the original image using cross-validation, which gives another array, which is studied to determine whether an image is morphed or not, and function for the final result is added on the basis of that.

D. Feature Extraction:

Extraction of Features: Out of all the methods to analyze an image, extraction of GLCM features has proven to be efficient time and time again. The gray level co-variance matrix is a tabulation that provides with statistical measures for texture analysis. This method takes into account the spatial relationship between the intensities of pixels in a gray-level image. In this paper, the GLCM features were calculated to study the differences in the original image and the digitally forged image. This gave 22 texture values (for each image) to work with, most of which were similar when it came to an image and its fraudulent counterpart. In practice, this would lead to redundancy and would also increase the time to run the algorithm. Also, the histogram of oriented gradient (HOG) features was calculated which gave another set of features for the original and the morphed image. The HOG values of the original and the morphed images were reasonably apart from each other, which meant that these values will be useful in differentiating the original document from the morphed one. However, the order of matrix generated by HOG algorithm is too large to be successfully fed into an SVM so it could also not be of practical use.

E. Online Database

Also, the Features values were computed but since the order of the matrix produced were very large to be trained by using ANN machine learning algorithm so as to enhance accuracy.

F. Classification

Initially, the classifier used for classification of dataset into two parts as original or morphed was linear kernel SVM. A linear kernel SVM is the most suitable classifier for two-class classification problems. It finds an equivalent hyperplane which separates the whole data by a specific criterion that depends on the algorithm applied. It tries to find out a hyper-plane which is far from the closest samples on the other side of the hyper plane while still classifying samples. It gives the best generalization techniques because of the larger margin.

G. Detection of the Forged region

After the identification of duplicate blocks, the further step is to highlight the duplicate blocks on the digital image, which also gives an indication of forged regions. Hence, system finally detects forged areas in the digital image. The corresponding forgered regions are being highlighted by the system.

4. PROPOSED ALGORITHM

i. The standard database consists of original, forged and processed images is considered in the performance analysis.
ii. The images in the database are converted to gray scale.
iii. The statistical features are computed on GLCMs which is developed from the gray scale images.
iv. The Support Vector Machine(SVM) is trained with statistical features for every image in the database using RBF kernel.
v. Statistical features of the testing image are obtained in similar process using steps ii and iii.
vi. Then the SVM classifier classifies the image either to be authenticated or forged one.

5. OVERVIEW OF PROJECT MODULES

A. Creation Of Non-Overlapped Blocks

In this approach, the detected image is initially divided into overlapping blocks. The basic approach here is to detect connected blocks that has been copied and moved. The copied area consists of many overlapping blocks. The further step would be extracting features from these blocks.

B. Feature Extraction Technique

After the creation of the overlapped blocks, the feature extraction technique is applied on the image to extract the features from specific block of the image. In this work approximation image local binary pattern features method is applied on the block region for extracting the features. AILBP (Approximation image Local Binary Pattern) Initially on the face images, bi-level wavelet decomposition method has been applied, which has been transformed face images into approximation images. Then, on approximation images, local binary patterns (LBP) have been used to extract local features of the face images. AILBP method is a combination of wavelets decomposition along with LBP method which is effective in terms of accuracy and it reduce time computation.

Wavelet decomposition Wavelet breakdown method is a occurrence of time and signal analysis method. It can be applied to decompose a forged image into many sub-band images with variation in spatial resolution, characteristic of frequency and directional features [21]. In this method, the approximation and details coefficients are computed by decomposing the face image up to two levels. Approximation coefficients pick the lowest frequency components and details coefficients contain the highest frequency component of an image. Only approximation coefficients are taken for further procedure. Approximation coefficients contain low frequency components of forged image, which contain whole information of the image. The variation of expression and small scale obstruct does not alter low frequency part but the high frequency portion of the image only. More than two steps decomposition of forged image result in information loss and hence, not involved in this work.

C. Identification of the duplicate Rows in a feature matrix

In the feature matrix, each rows represents a specific blocks. In order to detect the duplicate rows, first from feature matrix, system finds out number of rows which original feature matrix is being compared with filtered out resultant rows which are duplicates. Hence, such comparison gives blocks which are duplicates in the feature matrix.

D. Detection of the Forged region

After the identification of duplicate blocks, the further step is to highlight the duplicate blocks on the digital image, which also gives an indication of forged regions. Hence, system finally detects forged areas in the digital image. The corresponding forged regions are being highlighted by the system.

EXPERIMENTAL RESULTS

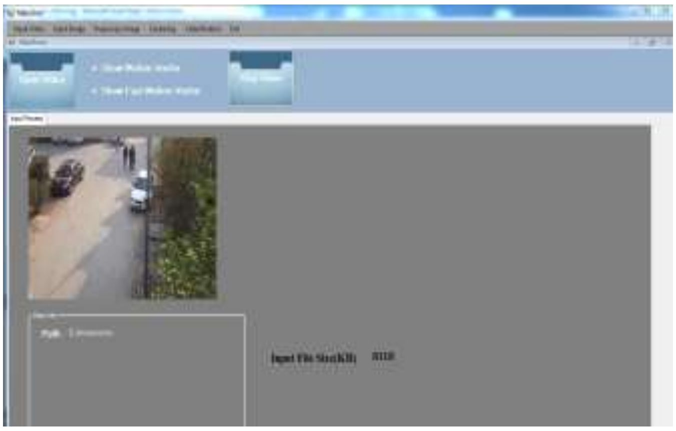

FIG 1: INPUT VIDEO

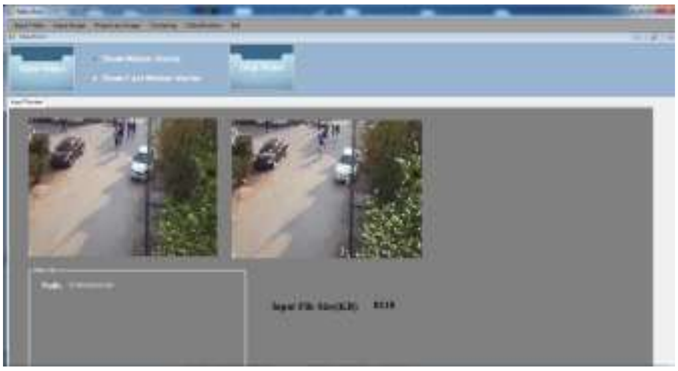

FIG 2: CAPTURE DUPLICATE REGION

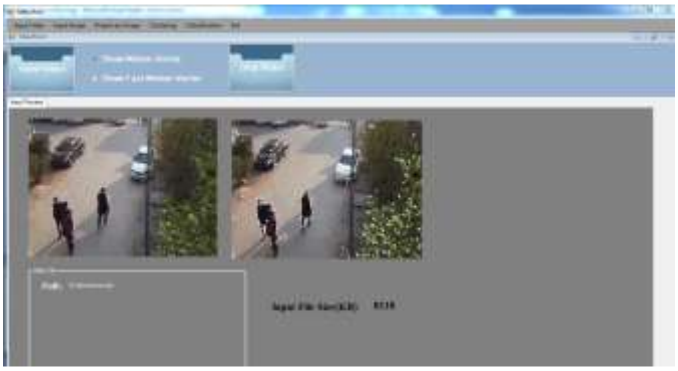

FIG 3: MOTION VECTOR SEGMENT

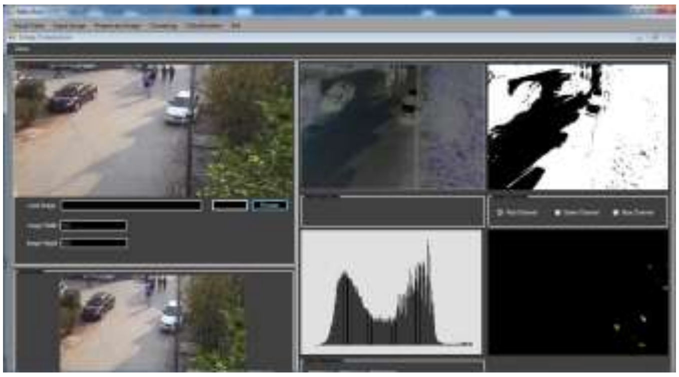

FIG 4: PROCESS IMAGE

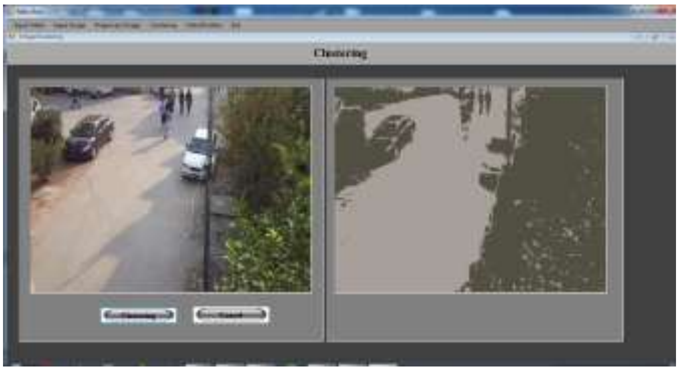

FIG 5: CLUSTERING

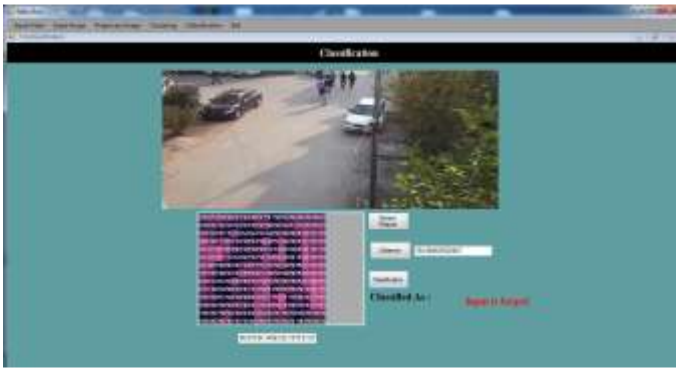

FIG 6: CLASSIFICATION

CONCLUSION

In this way, we are implemented the video forgery system for image and video forgery detection using ANN features and Machine Learning algorithm. By using GLCM feature this will classify the frame into different cluster and by using SVM method this system will generate the output as the video is forged or not.

ACKNOWLEDGEMENT

A very firstly we gladly thanks to our project guide Prof. Chandgude A. S. for his valuable guidance for implementation of proposed system. We will forever remain a thankful for their excellent as well as polite guidance for preparation of this report. Also we would sincerely like to thanks HOD and other staff for their helpful coordination and support in project work.

REFERENCES

[1] Shruti Ranjan, Prayati Garhwal, Anupama Bhan, Monika Arora, Anu Mehra " Framework For Image Forgery Detection And Classification Using Machine Learning", IEEE Xplore Compliant Part Number: CFP18K74-ART; ISBN:978-1-5386-2842-3.

[2] H M Shahriar Parvez, Hamid A. Jalab, and Somayeh Sadegh, "Copy-move Image Forgery Detection Based on Gabor Descriptors and K-Means Clustering" . (ICSCEE2018) ©2018 IEEE.

[3] R. Poisel and S. Tjoa, "Forensics investigations of multimedia data: A review of the state-of-the-art," in Proceedings - 6th International Conference on IT Security Incident Management and IT Forensics, IMF 2011, 2011, pp. 48–61.

[4] G. K. Birajdar and V. H. Mankar, "Digital image forgery detection using passive techniques: A survey," Digital Investigation, vol. 10, no. 3, pp. 226–245, 2013.

[5] G. Lynch, F. Y. Shih, and H. Y. M. Liao, "An efficient expanding block algorithm for image copy-move forgery detection," Inf. Sci. (Ny)., vol. 239, pp. 253–265, 2013.

[6] H. C. Hsu and M. S. Wang, "Detection of copy-move forgery image using Gabor descriptor," in Proceedings of the International Conference on Anti-Counterfeiting, Security and Identification, ASID, 2012.

[7] D. M. Uliyan, H. A. Jalab, A. Abuarqoub, and M. Abu-Hashem, "Segmented-Based Region Duplication Forgery Detection Using MOD Keypoints and Texture Descriptor," in Proceedings of the International Conference on Future Networks and Distributed Systems, 2017, p. 6:1--6:6.

[8] I. Amerini, L. Ballan, R. Caldelli, A. Del Bimbo, and G. Serra, "A SIFT-based forensic method for copy-move attack detection and transformation recovery," IEEE Trans. Inf. Forensics Secur., vol. 6, no. 3 PART 2, pp. 1099–1110, 2011.

[9] G. Sheng, T. Gao, Y. Cao, L. Gao, and L. Fan, "Robust algorithm for detection of copy-move forgery in digital images based on ridgelet transform," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2012, vol. 7530 LNAI, pp. 317–323.

[10] G. Muhammad, M. Hussain, and G. Bebis, "Passive copy move image forgery detection using undecimated dyadic wavelet transform," Digital Investigation, vol. 9, no. 1, pp. 49–57, 2012.

[11] B. Mahdian and S. Saic, "Blind authentication using periodic properties of interpolation," IEEE Trans. Inf. Forensics Secur., vol. 3, no. 3, pp. 529–538, 2008.

[12] A. C. Popescu and H. Farid, "Exposing digital forgeries by detecting duplicated image regions," Department of Computer Science., Dartmouth Coll. Tech. Rep. TR2004-515, no. 2000, pp. 1–11, 2004.

[13] Z. Moghaddasi, H. A. Jalab, R. Md Noor, and S. Aghabozorgi, "Improving RLRN Image Splicing Detection with the Use of PCA and Kernel PCA," Sci. World J., vol. 2014, 2014.

[14] D. Y. Huang, T. W. Lin, W. C. Hu, and C. H. Chou, "Boosting scheme for detecting region duplication forgery in digital images," in Advances in Intelligent Systems and Computing, 2014, vol. 238, pp. 125–133.

[15] W. Luo, Z. Qu, J. Huango, and G. Qiu, "A novel method for detecting cropped and recompressed image block," in ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 2007, vol. 2.

[16] S. J. Ryu, M. Kirchner, M. J. Lee, and H. K. Lee, "Rotation invariant localization of duplicated image regions based on Zernike moments," IEEE Trans. Inf. Forensics Secur., vol. 8, no. 8, pp. 1355–1370, 2013.

[17] Y. Cao, T. Gao, L. Fan, and Q. Yang, "A robust detection algorithm for copy-move forgery in digital images," Forensic Sci. Int., vol. 214, no. 1–3, pp. 33–43, 2012.

[18] D. M. Uliyan, H. A. Jalab, A. W. A. Wahab, P. Shivakumara, and S. Sadeghi, "A novel forged blurred region detection system for image forensic applications," Expert System Application, vol. 64, pp. 1–10, 2016.

[19] Y. Huang, W. Lu, W. Sun, and D. Long, "Improved DCT-based detection of copy-move forgery in images," Forensic Sci. Int., vol. 206, no. 1–3, pp. 178–184, 2011.

[20] M. Bashar, K. Noda, N. Ohnishi, and K. Mori, "Exploring duplicated regions in natural images," IEEE Trans. Image Process., no. 99, 2010.

[21] S. Sadeghi, H. A. Jalab, K. S. Wong, D. Uliyan, and S. Dadkhah, "Key point based authentication and localization of copy-move forgery in digital image," Malaysian J. Computer. Science., vol. 30, no. 2, pp. 117–133, 2017.

[22] D. M. Uliyan, H. A. Jalab, A. W. A. Wahab, and S. Sadeghi, "Image region duplication forgery detection based on angular radial partitioning and harris key-points," Symmetry (Basel)., vol. 8, no. 7, 2016.