Novel Method for Copy-Move Forgery Detection

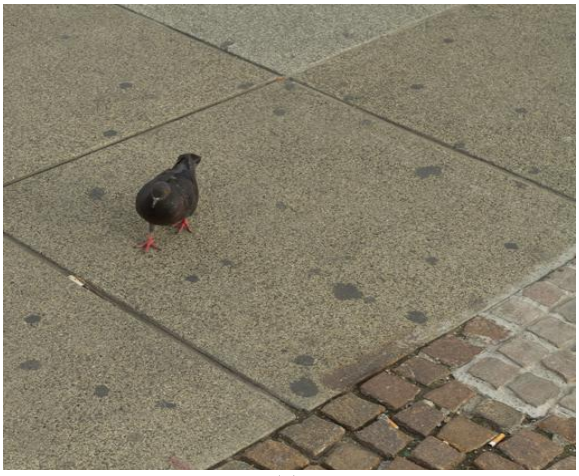
Sharanjit Kaur
Baba Banda Singh Bahadur Engineering College

Manpreet Kaur
Baba Banda Singh Bahadur Engineering College

ABSTRACT
A technique named forgery detection is used for detecting alterations from the image. This approach comprises several methods. This research study is depending on identifying theforgery in copy-move.The PCA algorithm will mark the principle component analysis for incompatible pixels from the picture. The GLCM algorithm is applied with the PCA algorithm in this research study for detecting forgery. The proposed algorithm is executed in MATLAB and outcomes are scrutinized on the basis of PSNR and MSE values. It is scrutinized that the proposed algorithm gives good performance in comparison with existing algorithm.

Keywords
PCA, GLCM, Copy-move forgery

1 INTRODUCTION
In recent times, images have become a primary source of information and an important information carrier due to the rapid development of technology. An image can be generated or copied and its storing can be done in the electronic form. The vector graphics or raster graphics are considered to define an image. Nevertheless, editing picture content has become increasingly convenient because of diverse multimedia editing software such as GIMP and Adobe Photoshop. The multimedia whose tempering has done is capable of making the lives of human more interesting. A threat is launched in various fields due to this kind of tempering [1]. A portion of an image is spliced onto some other image that leads to generate image forgeries. The tampered region is frequently scaled or rotated for creating it proportional to the neighbor un-tampered region under the splicing or object elimination. An image grid is re-sampled and thisre-sampling is detected for authenticating the manipulation of an image.Copy Move forgery is a special kind of forgery that contained the replicated parts of an image. Afterward, these replicated parts are pasted into similar image.



(a)

(b)

Fig 1: Example of copy-move forgery (a) Original image (b) Forged image

In a Copy-Move forgery, the copying and pasting of a part of the image is done itself into another part of the same image. This is often carried out with objective of making an object disappear from the image for which it is covered with a segment copied from another part of the image. The noise component and most other important properties become well-suited to the rest of the image as the duplicated parts come from the similar image [2]. Therefore, these properties are not detected with the utilization of techniques which seek inaptness in statistical measures within diverse parts of the image.

Copy move forgery also leads to overlapping of one object in an image by another object of the similar picture. This can be easily indicated in the above figure 1. Copy-Move forgery results to misleading the viewer.The feathered crop or the retouch tool is useful for masking any traces of the copied-and-moved segments further so that the forgery can be made even complicated for the detection.Over the past few decades, a number of methods have been introduced for detecting the copy-move forgery. These techniques are typically grouped into two groups, namely block-based methods and key point-based techniques [3]. Almost all CMF detection methods presented so far, are based on image processing. The CMFD framework includes 4 main phases. These phases are pre-processing, feature extraction, matching, and post-processing. In pre-processing many operations including conversion, transform, or decomposition are applied on the image. The main objective of the pre-processing phase is to develop and characterize data in a fashion, capable of making the successive feature extraction phase more efficiently. In this phase, various decomposition methods, such as wavelet decomposition or PCA are applied. Feature extraction stage aims to determine the overall precision of the CMFD framework. Creating a group of quick yet expressive data

vectors (so-called "feature descriptors") to make all parts of the target digital image highlighted is the main objective of this phase. A great number of methods are available for extracting feature vectors from the digital image. All feature extraction methods are mainly categorized into two categories of key-point and block-based methods [4]. Once the feature descriptors are extracted, feature matching methods are applied. In this phase, matched patches or segments of the target image are searched using identical feature descriptors. The matching procedure is one of the important phases which determine how much the general detection speed of the CMFD system is. The post processing stage is concerned with the filtering or processing ofunprocessed matched detection results for enhancing and producing the ultimate detection results of maximum quality.Copy-Move Forgery detection methods are generally categorized into block based and key point-based methods. In Block-based technique, the categorization of an image is done into one or many blocks of square or circle for scrutiny in the pre-processing stage. These blocks can either overlap or not overlap with each other. In the feature extraction stage, the extraction of features is carried out from these blocks. Then, the extracted features are compared against each other so that the similarity between blocks within the image can be determined. These blocks refer to the manipulation of copy-move forgery carried out in the image after the detection of the matched blocks [5]. Key point-basedtechniques are not based on block because the elimination of block division is performed in pre-processing set of descriptors which is generated in a region around the features. The descriptor is assisted in maximizing the reliability of the attributes in the affine transformation. Subsequently, the attributes and descriptors are matched and determined in an image for investigating the forged regions.

2 LITERATURE SURVEY

Gupta and Girdhar [2] introduced a novel technique to uncover the CMFD falsification. There was not any requirement of information related to the actual image in this technique. The overlapping blocks are initiated from the grayscale image first of all. In addition, the attributes were extracted using the hybrid schemes. These schemes made the utilization of Principal Component Analysis (PCA) and histogram of oriented gradients [6]. The last stage aimed at storing the attributes lexicographically that assisted in carrying out the matching of fake regions in easy way. The introduced method was compared with the traditional technique using this investigating study. The results of comparative analysis validated that the presented offset threshold value assisted in alleviating the fake matches that was the main issue occurred in the previous approach. The evaluation results revealed that the introduced approach was capable of enhancing the results and improved security for two different attacks.

Yeap et.al [10] stated that the major intend was to implement thepassive forgery detection on tampered images. For this purpose, the copy move method recognized as CMFD was deployed. There were Oriented FAST and rotated BRIEF included in this method. These schemes were utilized as technique of extracting the attributes. The 2NN was suggested along with the HAC. The images whichwere gone through different geometrical attacks had utilized to compute the presented method. In the evaluation, the precision rate was found 84.33% on the MICC-F600 database and 82.79% for MICC-F2000 database. The presented technique provided the TPR above 91% for tampered images while detecting the forgery.

Dhanya and Selvi [7] suggested a brief review on the CMFD technique. This technique was utilized on the digital images in case a specific slice of the image was fixed to another portion of the similar image for concealing the offensive objects. The suggested technique emphasized on investigating the forged portions. Thus, the tampering detection practices were found complex for any such operations [8]. The inventive experiments were reviewed in this research paper on the techniques of identifying the forgery that were planned on the basis of resemblance and association amid the pasted part and the real image and their comparison was also carried out.

Wu et.al [11] investigated a novel E2E DNN predicting forgery masks to the deal with issue of detecting the forgery in copy-move. In general, block-like attributes were extracted from an image; the self-correlations were quantified among diverse blocks; a point-wise feature extractor was deployed for locating the matching points and a forgery mask was reconstructed using de-convolutional network. To achieve this, the CNN was implemented [9]. The investigated solution was fully trainable and its optimization was easily performed together for the forgery mask reconstruction loss. The outcomes of experiment depicted that the investigated technique performed more efficiently for detecting the forgery as compared to the conventional techniques on the basis of various attributes and matching techniques. Additionally, this technique had more robustness to tackle several attacks such as JPEG, compression and so on.

Shabanian and mashhadi [4] recommended a novel block-based technique in order to detect the forgery in copy-movewithin digital images.The structural similarity index was carried out as the technique for similarity matching phase in this technique [10]. Furthermore, Gaussian pyramid decomposition technique assisted in enhancing the run-time speed of the recommended approach in significant manner. Thus, there was not any necessity of feature extraction. However, it was useful for reducing the problem of time-consumption. The recommended technique was proved efficient in specification as it had simplicity for the computation and analysis. The experiments were carried out for quantifying that whether this technique was robust and had sensitivity for addressing some post-processing operations. It was indicated that the recommended technique was effective.

Dixit et.al [8] intended a scheme with the objective of discovering the replicated areas in a picture. The statistical attributes of an image were utilized in this scheme [11]. The image was divided into the pixel blocks for employing themean and variance in order to achieve this. The contribution of every individual block concerning the pixel intensity of the entire image was investigated using mean. The variation of every pixel from its neighbors was figured out in a block with the help of variance. The intended algorithm was computed and its comparison was done with other CMFD methodologies. The outcomes obtained in the experiment demonstrated that the superior performance was achieved through the intended algorithm in contrast to traditional methods.

3 RESEARCH METHODOLOGY

This examination relies on the PCA method. This technique is employed for finding the dissimilar areas in a digital data. The copied region is marked with black color with the help of PCA algorithm, it is a multivariate manner which is used to analyze data table. This data table represents several interrelated quantitatively dependent variables. The major purpose of this approach is the extraction of important

information from the table for the representation of novel orthogonal variables. These variables are known as principal components.

The patterns of similarity of observations and the variables in the form of points within maps are displayed here. The data is centered primarily with respect to each variable when a given data matrix contains p variables and n samples. On the origin of principal components, the data occurs in the middle which however does not influence the spatial relations of data or the variances present along the variables. The initial principal component (Y_1) is specified through the linear combination of variables X_1, X_2, ...,X_p which is given below:

$$Y_1 = a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p \quad \ldots (1)$$

In the form of matrix notation, it can be specified as:

$$Y_1 = a_1^T X \quad \ldots (2)$$

The initial principal component is calculated for finding the greatest possible variance within the data set. Selecting large values for weights $a_{11}, a_{12}, \ldots a_{1p}$, the variance of Y_1 can be made. The weights are computed with the constraint such that the sum of squares is 1, to prevent such condition.

$$a_{11}^2 + a_{12}^2 + \cdots + a_{1p}^2 = 1 \quad \ldots (3)$$

The second principal component is computed in similar way as no correlation occurs towards the initial principal component. The next highest variance utilizes this second principal component.

$$Y_2 = a_{21}X_1 + a_{22}X_2 + \cdots + a_{2p}X_p \quad \ldots\ldots (4)$$

This process keeps on going till the computation of p principal components. These components are equal to the original number of variables. Equivalent values are obtained for the sum of variances of all principal components and the sum of variances of all variables in this point. Therefore, the alterations of all original variables to the principal components can be demonstrated as:

$$Y = XA \quad \ldots\ldots (5)$$

In this research study, the textual features of an input image are detected with the help of GLCM algorithm. By using these identified features, the region of copy-move forgery is detected from the image. The texture features are calculated at particular positions corresponding to each other using statistical texture analysis. On the basis of available intensity points within each combination, the statistics is classified in first-order, second-order and higher order. The second order statistical texture features can be extracted easily with the help of GLCM algorithm. The GLCM algorithm provides information related to the locations of pixels that include similar gray level values. A matrix which comprises equal number of rows, columns and gray levels in an image is known as Grey Level Co-occurrence Matrix (GLCM).

4 RESULT AND DISCUSSION

A tool named MATLAB is applied to calculate complex mathematical problems. This tool utilizes C programming language. Large numbers of inbuilt toolboxes remain present within MATLAB and several operations are performed easily using these tools. This tool can be used for the implementation of algorithms, plotting of graphs and the designing of various user interfaces.

Fig 2: Input image

The figure 2 represents an input image which is a cut-copy image.The initial data is a colored (RGB) picture, in the next stage this picture is further transformed into gray scale representation in order to proceed further image processing operations.

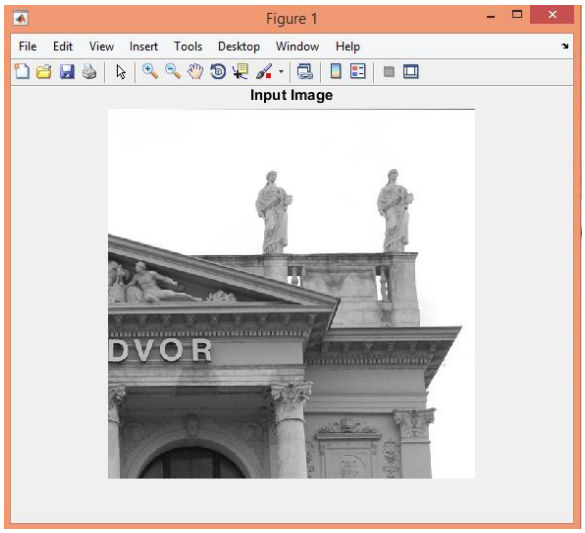


Fig 3: Grayscale image

The input image after being processed is shown in figure 3 The above image is a grayscale image. Now GLCM is applied on this image to perform the extraction of features.

GLCM applying parameter values
0.0006
0.0008
0.0002
0.0009
0.1141
0.0410
0.0071
0.0159
1.3539

0.0010
0.0027
-0.0008
0.2550

Fig 4: GLCM Parameters

The above figure 4 shows the values of the parameters obtained after applying GLCM, Further PCA algorithm is to be applied here for detecting the forged part of this image.

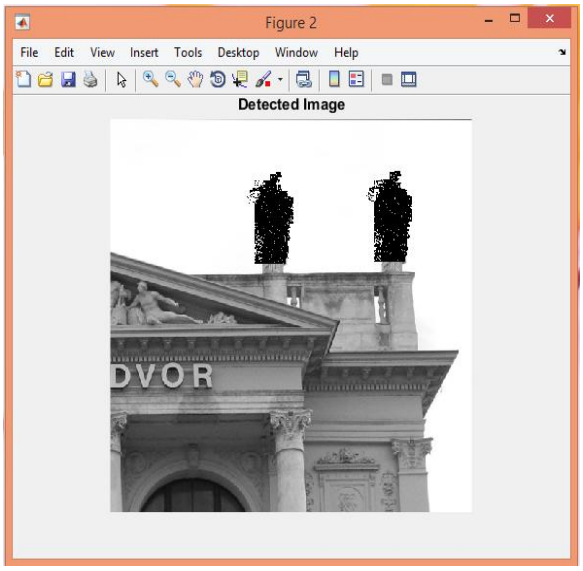


Fig 5: Output image with detected forged areas

The figure 5 reveals that Principal Component Analysis (PCA) is implemented to compute the principle components of image. The mismatched components of picture are marked which the black color.

Fig 6: Input Image

The input image after being processed is shown in figure 6 The above image is a grayscale image. Now GLCM is applied on this image to perform the extraction of features.

Fig 7: Output image with detected areas

The figure 7 illustrates that the Principal Component Analysis (PCA) algorithm is utilized to principle components of the image. The mismatched components of image are marked which the black color. This result also exhibits the multiple forged areas.

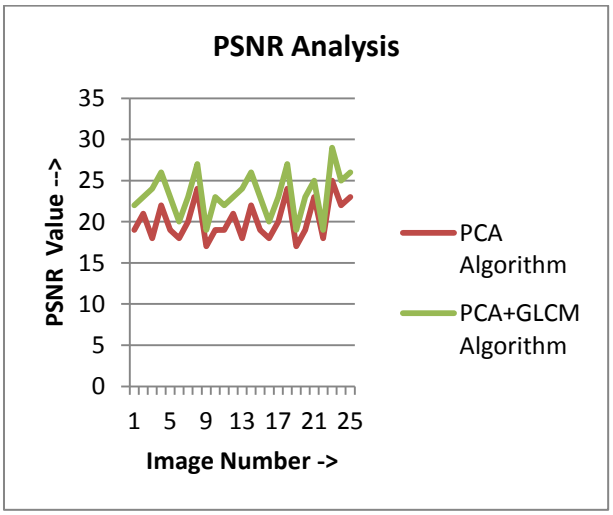


Fig 8: PSNR Comparison

The figure 8 represents that the forgery in copy-move is identified using PCA algorithm and PCA algorithm in combination with Grey Level Co-occurrence Matrix (GLCM). A superior PSNR value obtained from PCA along with GLCM algorithm as compare to Principal Component Analysis (PCA) algorithm.

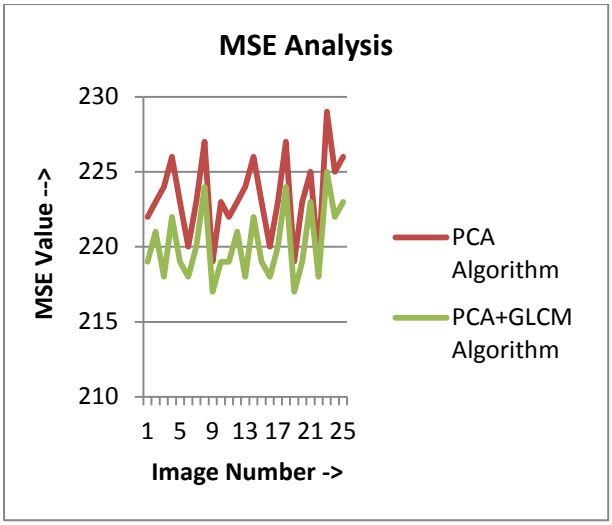


Fig 9: MSE Comparison

The figure 9 depicts that the comparison of mean squared error (MSE) value of PCA is done with the GLCM along with PCA algorithm. This demonstrates that the MSE value obtained from GLCM with PCA algorithm is found lower incomparison with the Principal Component Analysis (PCA) algorithm.

5 CONCLUSION

Image processing approach is implemented for processing the data that is accumulated as picture elements. The CMFD technique is executed in order to detect the tampered regions of input images. The prior approach makes the deployment of PCA algorithm for the feature extraction. The detection of tampered region of an image is done with the help of this algorithm. This investigative study focuses on detecting the forgery using the Grey Level Co-occurrence Matrix (GLCM) along with the PCA algorithm. The GLCM algorithm is suggested. This algorithm evaluates the co-occurrence matrix for recognizing the textural properties of image.

6 REFERENCES

[1] G. H. Li, Q. Wu, D. Tu, and S. J. Sun., "A sorted neighborhood approach for detecting duplicated regions in image forgeries based on DWTand SVD," in Proceedings of IEEE International Conference on Multimedia and Expo, Beijing, Jul. 2007, pp. 17503.

[2] Geetika Gupta, Akshay Girdhar., "A ROBUST PASSIVE METHOD FOR DETECTION OF COPY-MOVE FORGERY IN IMAGES", 2017, IEEE.

[3] H. Huang, W. Guo, and Y. Zhang., "Detection of copy-move forgery in digital images using SIFT algorithm," in Pacific-Asia Workshop on Computational Intelligence and Industrial Application, Vol. 2, Dec. 2008, pp. 2726.

[4] Hanieh Shabanian, Farshad Mashhadi., "A New Approach for Detecting Copy-Move Forgery in Digital Images", 2017, IEEE.

[5] I. Amerini., "A SIFT-based forensic method for copymove attack detection and transformation recovery", IEEE Trans. Inf. Foren. Sec., 2011.

[6] M. Ghorbani, M. Firouzmand, and A. Faraahi., "DWT-DCT (QCD) based copy-move image forgery detection", in 18th IEEE International Conference on Systems, Signals and Image Processing (IWSSIP), 2011, pp. 14.

[7] R.Dhanya, R.Kalai Selvi., "A State of the Art Review on Copy Move Forgery Detection Techniques", Proceedings of 2017 IEEE International Conference on Circuits and Systems (ICCS 2017).

[8] Rahul Dixit, RuchiraNaskar and Aditi Sahoo., "Copy–Move Forgery Detection Exploiting Statistical Image Features", IEEE WiSPNET 2017.

[9] X. Kang, and S. Wei., "Identifying tampered regions using singular value decomposition in digital image forensics", in International Conference on Computer Science and Software Engineering, 2008.

[10] Yong Yew Yeap, U. U. Sheikh, Ab Al-Hadi Ab Rahman., "Image Forensic for Digital Image Copy Move Forgery Detection", 2018 IEEE 14th International Colloquium on Signal Processing & its Applications (CSPA 2018).

[11] Yue Wu, Wael Abd-Almageed, and Prem Natarajan., "Image Copy-Move Forgery Detection via an End-to-End Deep Neural Network", 2018 IEEE Winter Conference on Applications of Computer Vision.