International Journal of Engineering & Technology, 7 (3.27) (2018) 215-219

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

A Novel Approach for Efficient Forgery Image Detection Using Hybrid Feature Extraction and Classification

G. Clara Shanthi[1*], V. Cyril Raj[2]

[1]*Research Scholar, Department of CSE, Dr.M.G.R. Educational and Research Institute University, Tamil Nadu, India.*
[2]*Professor, Department of CSE & IT, Dr.M.G.R. Educational and Research Institute University, Tamil Nadu, India.*
E-mail:cyrilraj@drmgrdu.ac.in
**Corresponding author E-mail:clarashanthi2@gmail.com*

Abstract

Image forgery detection is developing as one of the major research topic among researchers in the area of image forensics. These image forgery detection is addressed by two different types: (i) Active, (ii) Passive. Further consist of some different methods, such as Copy-Move, Image Splicing, and Retouching. Development of the image forgery is very necessary to detect as the image is true or it is forgery. In this paper, an efficient forgery detection and classification technique is proposed by three different stages. At first stage, preprocessing is carried out using bilateral filtering to remove noise. At second stage, extract unique features from forged image by using efficient feature extraction technique namely Gray Level Co-occurance Matrices (GLCM). Here, the GLCM improves the feature extraction accuracy. Finally, forged image is detected by classifying the type of image forgery using Multi Class- Support Vector Machine (SVM). Also, the performance of the proposed method is analyzed using the following metrics: accuracy, sensitivity and specificity.

Keywords: *Image forgery detection, image splicing, image retouching, copy-move, support vector machine, gray level co-occurrence matrices.*

1. Introduction

For the past decades, the image forgery detection has become the important research in the application of biomedical technology, digital image processing, forensics image and criminal investigation, etc,

It is more difficult when potent software tools for image processing are so prevalent and complex that we can't affirm whether an image is controlled by naked eyes[1-2]. For the most part, the Image Forgery Detection (IFD) methods are characterized into two categories [3].

Because of the differences of tampering controls, the analogous passive detection systems are proposed to distinguish forgery images [4]. This paper centers around algorithm and procedures for image Copy-Move, Retouching and image spliced finding methods.

Image Retouching

This method is common in photo edition. It is an attractive method which improves certain features of image. This type of forgery image is present in all magazine cover[6]. the following figure can shows the image retouching attacks.

Fig. 1: Image retouching attack on images

Copy–Move

It is the most prevalent image tampering technique. In this technique the main purpose is to hide the original image with some different other part of the same image [6]. The figure 1 shows the example of copy-move technique. The copy-moved picture on the right has four missiles whereas the original image contains three missiles.

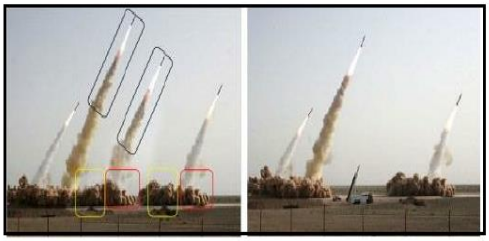


Fig. 2: Copy-move attacks on images

Image Splicing

This method is used to make a more aggressive forgery images. It

is an easy procedure which can be crops and paste on the desired areas from other sources. The following figure shows the image splicing based forgery image. It can be used to copy a spliced portion from original image to desired image [7].

Fig. 3: Image splicing attack on Images

This paper presents an efficient forgery detection and classification technique that can propose three major stages. At first stage, preprocessing is carried out using bilateral filtering to remove noise. At second stage, extract unique features from forged image by using efficient feature extraction technique namely Gray Level Co-occurance Matrices (GLCM) has been used to improve the accuracy of feature extraction.
Lastly, forgery image is identified by classifying the type of image forgery using Multi Class- Support Vector Machine (SVM). Also, the performance of the proposed method is analyzed using metrics include specificity, accuracy and sensitivity.

2. Literature Review

Chen-Ming Hsu et al (2015) has proposed an proficient algorithm for detecting Copy-Move Forgery by using the Histogram Oriented Gabor Magnitude (HOGM) algorithm to identify the replicated region based on the gabor magnitude of the histogram. Thus the implementation results established the overlapping blocks of HOGM and it can identify the multiple copy-move forgery instances such as blurring, JPEG compression, brightness adjustment, slight image rotation; therefore the computational complexity also very less, low complexity, and it easily locates the duplicate image but cannot use large block size.
Kaur et al (2015) recommended a Copy-Move Forgery Detection using DCT and SIFT approach to identify the mixture of dissimilar post-processing operations by single method. Also the key point-based SIFT method is used for rotating and scaling the image. Thus the mixture of SIFT and DCT approach is capable to identify forgery under post-processing of rotation, Gaussian noise, scaling, JPEG compression and thus the forgery detection has effectively improved by this method but detection method have many false matches.
Although the researches XuanjingShen et al (2016) described Splicing Image Forgery Detection by using Textural Features based Gray Level Co-occurrence Matrices (TF-GLCM). In TF-GLCM, the GLCM is measured by using Diffrence Block Discrete Cosine Transform (DBDCT) arrays that is used to get the textural information and the spatial relationship sufficiently among the image pixels. The experimental results showed that the TF-GLCM 98% &97% of finding rate on CASIA v1.0, & CASIA v2.0 respectively.
Kaur & Manoj Kumar (2015) proposed An Effective finding algorithm for Copy-Move Forgery. The researchers concluded that the CMF detection is a major security factor in image processing. The survey showed that there are several works completed by CMF detection technique for image cloning. Thus the system achieved efficient performance such as better enhancement, overlap, and better compression. The researcher has described various copy move forgery detection techniques but the detection accuracy is less.

3. Forgery Detection Framework

The proposed functional arrangement of image forgery detection process is shown in following figure. Hence the process requires major stages includes image Preprocessing, feature extraction, and classifier. Image pre-processing is also an improvement stage of the image data, which can suppress the unwanted distortions and to enhance the important features of the remaining processing. Additionally, the preprocessing is to improve the accuracy rate.

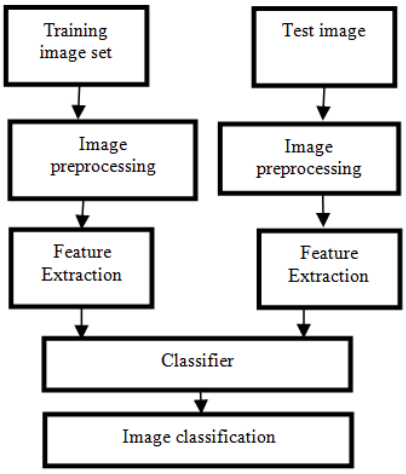


Fig. 4: Block diagram for proposed Forgery detection

Enhancement and image resize conversion from RGB to gray scale, filtering, edge detection. In this research, the process of image preprocessing includes two major steps: fast bilateral filtering and canny edge detection. Edge detection output is an input for feature extraction; the input is given through image preprocessing.

Image Preprocessing

1) Bilateral filter

The bilateral filter is used to remove the noise and reduces the blurriness of color images. It is very simple and non- interactive based implementation method. Basically the bilateral filter is used to make a nonlinear combination of similar pixel values. It is employed to filter the noise range of image which acts as a domain filter [12]. The values selected to view the required amount of arrangement of pixels, while the range filtering select values based on the desired amount of low pass filtering. A low pass domain filter image is derived as I(x):

$$k_d(c) = \iint_{\infty}^{\infty} c(x,y)dy \quad (1)$$

Where
$kd(x)$ - normalized constant;
C(x, y) is used calculate the geometric closeness among neighborhood center x &y,
h(x) -output image;

$$h(x) = k_d^{-1} \iint_{-\infty}^{\infty} i(y)c(x,y)dy \quad (2)$$

$$k_r(x) = \iint_{-\infty}^{\infty} s(i(x)i(y)dy \quad (3)$$

s(f(x),f(y)) calculated the photometric similarity of neighborhood pixel center x & y.

$$h(x) = k_r^{-1} \iint_{-\infty}^{\infty} i(y) * s\big(i(x), i(y)\big) * c(x,y)dy \quad (4)$$

$$k_r(x) = \iint_{-\infty}^{\infty} s(i(x), i(y) * c(x,y)dy \quad (5)$$

The output image is redefined by,

$$s(i(x), i(y) = e^{\frac{-(i(y)-i(x))^2}{2\sigma_r^2}} \quad (6)$$

$$c(x,y) = e^{\frac{-(i(y)-i(x))^2}{2\sigma_r^2}} \quad (7)$$

Pixel location (x,y) and output I(x,y) of bilateral filter measured by the following equation:

$$I'(X,Y) = \in y_{\in N(X)} e^{\frac{-(i(y)-i(x))^2}{2\sigma_r^2}} e^{\frac{-(i(y)-i(x))^2}{2\sigma_D^2}} i(x,y) \quad (8)$$

Where & σr - spatial & intensity of fall-off of weights control parameter,
N(x) - pixel of I(x),
σd - geometric spread parameter, used to select the required low pass filter. The bilateral filter helps to understand its entire variants.

Edge Detection

Edge detection is a most important step in image processing to locate an edge of desired image. This process is used to understand image features. Edge detection technique also used to decrease the image size and filter-out the data. Thus it could be preserving the structural properties of image. Hence some edge detection techniques are Canny, Sobel, Prewitt, Roberts, and so on. Among the various methods, 'Canny method' is mostly preferable for detect both strong and weak edges. Canny edge detection has following merits such as good detection, good localization, and minimal response. While the optimal function of canny detector is obtained by using 4 exponential terms, and the performance of these algorithm is depends on the adjustable parameters σ with the standard deviation of Gaussian filter. Hence the 'σ'is used to control size of the Gaussian filter. However, the larger value of σ has suggests a larger Gaussian filter and which can provide more blurring to detect larger edges. Small value of 'σ' suggests a smaller Gaussian filter and that can be limits the amount of blurring in image.

Feature Extraction

Hybrid feature extraction = PCA+GLCM

1) Principal component analysis

A standout amongst the most well-known types of dimensionality decline is Principal Components Analysis (PCA). Given an arrangement of information, PCA finds the direct lower-dimensional illustration of the information with the end goal that the difference of the recreated information is protected. Utilizing an arrangement of feature decrease is based on a joined rule part investigation on the element vectors that figured from the wavelets restricting the element vectors to the segment chosen by the PCA should prompt a powerful grouping algorithm creating controlled approach. In this way, the primary novel behind utilizing PCA in proposed strategy is to reduce the dimensionality of the wavelet coefficients. This prompts more proficient and exact classifier.

2) Gray level Co-occurrence Matrices (GLCM)

$$G = \begin{bmatrix} h(1,1) & \cdots & h(1,Bg) \\ \vdots & \ddots & \vdots \\ h(Bg,1) & \cdots & h(Bg,Bg) \end{bmatrix} \quad (9)$$

Where h (i, j) shows the relative frequencies;
I and j- pixel couple values of image. h(i, j) can be derived as follow:

$$h(i,j) = \#\{(p_1,q_1),(p_2,q_2) \in A * B \setminus f(p_1,q_1) = i, f(p_2,q_2) = J\} \quad (10)$$

Where, # - number of elements in a set.
The diverse arrangements of separation (indicated by d) and edges (meant by θ) between the two pixels can impact the method for computing the quantity of the pixel couples in GLCM. In TF-GLCM, the separation is set at 1 and the edges are 0°, 45°, 90°, and 135° [11]. What's more, these two parameters (d and θ) would bring about four types of grouping. Subsequent to being standardized, for example, the four GLCM removed from the horizontal contrast coefficient array can be given by

$$G_{d=1,\theta=0^\circ}(F_H(i,j)), \ldots . G_{d=1,\theta=135^\circ}(F_H(i,j)) \quad (11)$$

Although the researcher Haralick has extracted totally twenty four parameters from co-occurrence matrix, but selectively seven parameters are commonly used which given as follows

Energy

The textural uniformity that is pixel pair repetitions, which is measured by energy called as uniformity or angular second moment.

$$energy\,(E1) = \sum i \sum j\, p_{ij}^2 \quad (12)$$

Detects the disorders in textures. Energy reaches a maximum value equal to one. Energy (E1) is the sum of the square of the elements in the GLCM, which can measure textural uniformity.

Entropy

Entropy can be used to measure the disorder or complexity of desired images. The entropy is derived as larger when the image is not texturally uniform. The following entropy equation shows the strong formation and the complex textures tends to high entropy

$$entropy\ (ent) = -\sum i \sum j\, p_{ij} log_2 p_{ij} \quad (13)$$

This entropy formula is used to measures the disorder of an image. When the image not texturally uniform, the discrepancy of the elements in GLCM is large, which implies the entropy have consistently small values.

Contrast

Contrast is used to measure the spatial frequency of an image and it is difference moment of GLCM. Contrast is used to measures the change between the lowest and highest values of a adjoining set of pixels. It is a simple term to measure the quantity of local differences exist in the pictures.

$$contrast\,(con) = \sum i \sum j\, (i-j)^2\, pij \quad (14)$$

Contrast reflects the sharpness of the image. The greater contrast is obtained from above equation, and it shows the clearer visual effect, the fuzzier visual effect lower contrast.

Homogeneity

Homogeneity is likewise called as inverse difference moment. It is utilized to measures the picture of homogeneity as it accept substantial value for littler gray tone contrast in pair components. It is more complex to the nearness of close diagonal components in the GLCM. It has most extreme value when all components in the picture are same. Homogeneity reduces if contrast increases while vitality is kept steady.

$$homogeneity\ (hom) = \sum i \sum j \frac{1}{1+(i-j)^2} p_{ij} \quad (15)$$

Also the local homogeneity is measures the closeness of the distribution of elements in GLCM to diagonal of GLCM.

Correlation

$$\text{correlation} = \sum_{I,j=0}^{N-1} P_{ij} \frac{(i-\mu)(j-\mu)}{\sigma^2} \quad (16)$$

Correlation is used to measures the joint probability occurrence of pixel pairs and similarity of two images.
Where
P_{ij} - texture image pixel value in a position of (i,j),
N - Number of image gray levels,
$\mu = \sum_{i,j=0}^{N-1} i p_{ij}$; μ - texture image
$\sigma^2 = \sum_{i,j=0}^{N-1} p_{ij} (i-\mu)^2$; σ^2- variance of the texture image

Variance

This static is a measure of heterogeneity and is strongly correlated to heterogeneity and is strongly correlated to first order statistical variable such as standard deviation.
Variance increases when the gray level values differ from their mean.

$$variance\ (var) = \sum i \sum j\,(i-j)^2\ p_{ij} \quad (17)$$

Where, μ -mean of pixel value in position (p_{ij})
Hence, Difference variance = variance of P_{X-Y}
Difference of entropy = $\sum_{i=0}^{N-1} p_{x-y}(i) log\{p_{x-y}(i)\}$ (18)

Cluster shade

$$S_{CS} = \sum_{i,j=1}^{N} (i - M_x + j - M_y)^3\ p(i,j) \quad (19)$$

Skewness of the matrix is measured under cluster shade and the lack of symmetry is called as cluster shade feature.
Cluster prominence: $S_{cp} = \sum_{i,j=1}^{N} (i - M_x + j - M_y)^4\ p(i,j)$ (20)
Where,

$$M_x = \sum_{i,j=1}^{N} ip(i,j)\ ;$$

$$M_y = \sum_{i,j=1}^{N} jp(i,j);$$

$$mean(m) = \sum_{i=0}^{l-1} z_i\ p(z_i) \quad (21)$$

m- Average intensity.
If the cluster prominence is low, peak co-occurrence matrix around the mean values. For image, the means that there is little variation is occurred in gray-scales.

Classifier

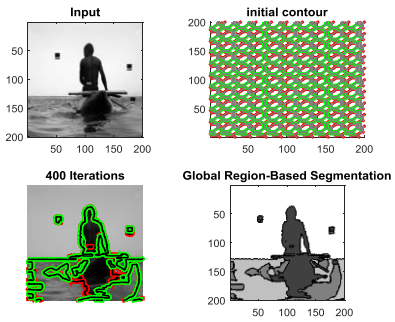
Multi Class-Support Vector Machine (MC-SVM) classifier is employed to classify the forgery detection, accuracy of detecting forgery is also enhanced by MC-SVM classifier. Also it can provide the results for additional complex complications and the output must be more one class that can be divided into (M) mutual exclusive classes. Also the MC-SVM is considered into four important methods such as: Directed Acyclic Graph (DAG), Binary Tree (BT), One Against-One (OAO), One-Against-All (OAA) classifiers.
Generally SVM class label requires two values such as ± 1, it can be called as bilinear classifier [11]. MC-SVM classifier is constructed by f 1, f 2, . . $f_{M,}$ each set is trained by each class. Multi-class maximum output is derived from $g^i(x)$:

$$g^j(x) = \sum_{i=1}^{m} y_i \alpha_i^j k(x, x_i) + b^j \quad (22)$$

The detection of forge picture was done by using SVM classifier on systematic basis by planning a simple process containing of two phases, namely training phase and testing phase.[16]

1. Segmentation-multi-Level set algorithm

The level set technique is a segmentation method that streamlines a target work and the zero level of that target work speaks to an understood form which fragments the picture. Multi-level set algorithm works in three stages. In the initial step, clusters are identified from the Gaussian filtered image utilizing multi-level set technique using both intensity and edge data. Filtering of the picture enhances the cluster segmentation enhancement. The middle regions are recognized by applying intensity based level set numerous circumstances until the center that no middle is cleared out. The last and most significant step uses a modified multi-level set technique for every nucleus to recognize their analogous cell limit.[15]



4. Result and Discussion

This section analyzes the experimental results of presented proposed Hybrid Feature Extraction and Classification method. The proposed set of experiments has established the better performance and effectiveness of image forgery detection using GLCM. Additionally the proposed methodology is to analyze the following merits such as
Accuracy: It is a ratio of number of correct assessment to the total number of assessments. It is measured as the term of (%)

$$\text{Accuracy}\ (\%) = \frac{((TN+TP))}{((TN+TP+FN+FP} \quad (23)$$

Sensitivity: It is defined as the ratio of the number of true positive assessments to the total number of true positive and false negative assessments. It is a degree of positive values that are correctly recognized.

$$\text{Sensitivity}\ (\%) = \frac{TP}{((TP+FN))} \quad (24)$$

Specificity: It is defined as the ratio of the amount of true negative valuations to the total amount of true negative and false positive assessments.

$$\text{Specificity}\ (\%) = \frac{TN}{((TN+FP))} \quad (25)$$

It is used to calculate the impact of changes in the output due to the change in the input dataset. It is a degree of negative values that are correctly recognized.

5. Conclusion

This paper presents an efficient forgery detection and classification technique that can propose three major stages. At first stage, preprocessing is carried out using bilateral filtering to remove noise. At second stage, extract unique features from forged image by using efficient feature extraction technique namely Gray Level Co-occurance Matrices (GLCM) and PCA. Here, the GLCM improved the feature extraction accuracy. Lastly, forged image is detected by classifying the type of image forgery using Multi Class-Support Vector Machine (SVM). The

performance of the proposed method is analyzed using metrics such as accuracy, sensitivity and specificity.

References

[1] Pun CM, Yuan XC & Bi XL, "Image forgery detection using adaptive oversegmentation and feature point matching", IEEE Transactions on Information Forensics and Security, Vol.10, No.8,(2015), pp.1705-1716.

[2] Hsu CM, Lee JC & Chen WK, "An efficient detection algorithm for copy-move forgery", 10th Asia Joint Conference on Information Security (AsiaJCIS), (2015), pp.33-36.

[3] Ardizzone E, Bruno A & Mazzola G, "Copy–move forgery detection by matching triangles of keypoints", IEEE Transactions on Information Forensics and Security, Vol.10, No.10,(2015), pp.2084-2094.

[4] Kaur A & Sharma R, "Copy-move forgery detection using DCT and SIFT", International Journal of Computer Applications, Vol.70, No.7,(2013).

[5] Shanthi MGC & Raj VC, "Image Forgery Detection Based On Local Texture Descriptors", Image, (2017), pp.268-273.

[6] Kaur G & Kumar M, "Study of Various Copy Move Forgery Attack Detection Techniques in Digital Images", International Journal of research in computer application and robotics, (2015).

[7] Cao Y, Gao T, Fan L & Yang Q, "A robust detection algorithm for copy-move forgery in digital images", Forensic science international, Vol.214, No.1-3,(2012), pp.33-43.

[8] Cattaneo G & Roscigno G, "A possible pitfall in the experimental analysis of tampering detection algorithms", 17th International Conference on Network-Based Information Systems, (2014), pp.279-286.

[9] Nguyen HC & Katzenbeisser S, "Detection of copy-move forgery in digital images using radon transformation and phase correlation", Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), (2012), pp.134-137.

[10] Panchumarthi, G. P., & Surendar, S. M. A. (2017). A review article on Fin-FET based self-checking full adders. Journal of Advanced Research in Dynamical and Control Systems, 9(4).

[11] Qureshi MA & Deriche M, "A bibliography of pixel-based blind image forgery detection techniques", Signal Processing: Image Communication, Vol.39, (2015), pp.46-74.

[12] Kim JINHO, Kim BS & Savarese S, "Comparing image classification methods: K-nearest-neighbor and support-vector-machines", Ann Arbor, (2012).

[13] Chaudhury KN & Dabhade SD, "Fast and provably accurate bilateral filtering", IEEE Transactions on Image Processing, Vol.25, No.6, (2016), pp.2519-2528.

[14] Shen X, Shi Z & Chen H, "Splicing image forgery detection using textural features based on the grey level co-occurrence matrices", IET Image Processing, Vol.11, No.1,(2016), pp.44-53.

[15] Villalobos Antúnez, JV (2017). Karl R. Popper, Heráclito y la invención del logos. Un contexto para la Filosofía de las Ciencias Sociales. Opción Vol. 33, Núm. 84. 5-11

[16] M Pallarès Piquer and O Chiva Bartoll (2017). La teoría de la educación desde la filosofía de Xavier Zubiri. Opción, Año 33, No. 82 (2017): 91-113