International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2017 IJSRCSEIT | Volume 2 | Issue 6 | ISSN : 2456-3307

Forged Image Detection by Analyzing Edge, Visual Saliency and Textural Features using SVM Classifier

R Gomathi
Department of ECE, University College of Engineering Dindigul, Dindigul, Tamilnadu, India

ABSTRACT

Nowadays, the use of Digital images is everywhere, in the magazines, in newspapers, in hospitals, in shopping malls and all over the Internet. As the development in technology is increasing day by day, at the same time the trust in images is decreasing day by day. Most common type of Image forgery is Image composition, which is also termed by the name Image Splicing. Combination of two or more images to generate a fake image is known as Image composition. It becomes very hard to differentiate between real image and fake images because of the presence of various powerful editing software. As a result, in most of the cases, there is a need to prove whether the images are real or not. This paper describes a technique for detecting forgery of composite images using Support Vector Machine (SVM) classifier. In the state of art of work, the forged image is detected by extracting Edge and Visual Saliency features. The proposed work detects the forged image by extracting Textural features in addition with Edge and Visual Saliency features. By using True Negative (TN) rate, True Positive (TP) rate and Accuracy parameters, it is found that the proposed method gives improved efficiency when compared with the existing methods.
Keywords: Digital Image, Support Vector Machine, Textural Features

I. INTRODUCTION

The trustworthiness of digital images has been questioned, because of the ease with which these images can be manipulated in both its origin & content because of tremendous growth of digital image manipulation tools. Digital image forensics is the latest research field, which intends to authorize the genuineness of images. Forgeries are not new to humankind but are a very old problem. In the past, it was limited to art and literature but did not affect the public [1], [2], [3].

Nowadays, due to the advancement of digital image processing software and editing tools, an image can be easily manipulated and modified. It is very difficult for humans to identify visually whether the image is original or manipulated. There is rapid increase in digitally manipulated forgeries in mainstream media and on the Internet. This trend indicates serious vulnerabilities and decreases the credibility of digital images. Therefore, developing techniques to verify the integrity and authenticity of the digital images is very important, especially this challenges the reliability of digital images offered as medical diagnosis, as evidence in courts, as newspaper items or as legal documents because of difficulty in differentiating original and modified contents.

Digital forensics field has developed significantly to combat the problem of image forgeries in many domains like legal services, medical images, forensics, intelligence and sports [4], [5], [6]. The objectives of this research are,

- To detect the forged images using SVM algorithm which uses adaptive selection of features
- To compare the performance of the proposed system with the existing systems

II. SUPPORT VECTOR MACHINE CLASSIFIER

The Support Vector Machine (SVM) is a state-of-the-art classification method introduced in 1992 by Boser, Guyon, and Vapnik. The SVM classifier is widely used in bioinformatics (and other disciplines) due to its high accuracy, ability to deal with high-dimensional data such as gene expression, and flexibility in modeling diverse sources of data. Yet, obtaining the best results

with SVMs requires an understanding of their workings and the various ways a user can influence their accuracy [7], [8], [9].

SVMs belong to the general category of kernel methods. A kernel method is an algorithm that depends on the data only through dot products. When this is the case, the dot product can be replaced by a kernel function, which computes a dot product in some possibly high dimensional feature space [10], [11], [12].

This has two advantages:

- First, the ability to generate non-linear decision boundaries using methods designed for linear classifiers.
- Second, the use of kernel functions allows the user to apply a classifier to data that have no obvious fixed-dimensional vector space representation. The prime example of such data in bioinformatics are sequence, either DNA or protein, and protein structure.

When training an SVM, the practitioner needs to make a number of decisions: how to pre-process the data, what kernel to use, and finally, setting the parameters of the SVM and the kernel. Uninformed choices may result in severely reduced performance.

III. PROPOSED METHODOLOGY

In the forensic field, it is necessary to verify and preserve the integrity of the evidence images during the course of investigation. In the state of art of work, many methods are proposed. This proposed work has better method to detect authentication of the image. The steps in the classification and detection of spliced images are discussed under this heading.

Flow Diagram of Proposed Method

Training Phase

The training dataset of 20 images, which contains 10 original images, and 10 forged images is collected and pre-processed. From the pre-processed image, edge, visual saliency and textural features are extracted. Hash value is calculated for the extracted features and given as input to the SVM classifier. Authority key is set after training so that only authorized person can run the application. The diagrams are shown in Figures 1(a) and 1(b).

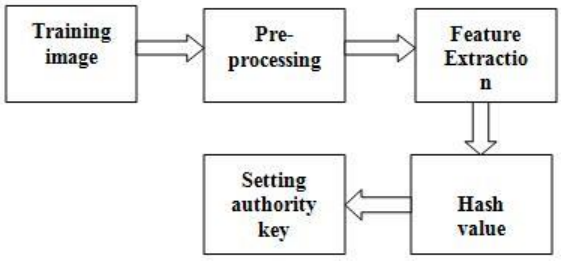

Figure 1 (a) Training Phase

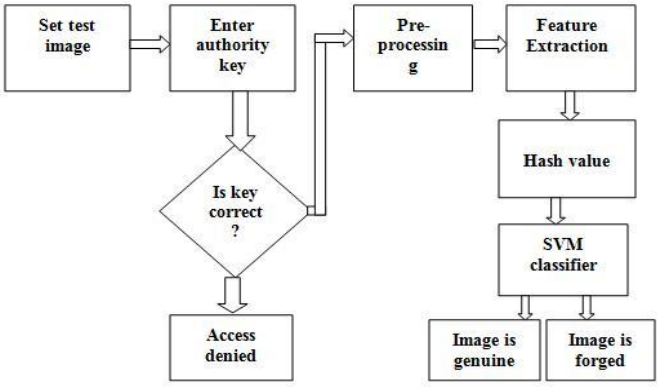

Figure 1 (b) Testing Phase

Training image set

A database is created and trained in the training phase with a number of images. The jpg and jpeg images either can be downloaded from the Internet or captured using digital cameras. The images can be of different sizes. The authority key is set after training images in the database. When any user tries to run the application for checking the authenticity of query image, he has to enter the same key, which was set during the training phase. Thus making sure that only authorized person can run the application.

A. Image Pre processing

The images are converted into gray scale from RGB. Median filter is applied to remove noises present in the images. Image enhancement techniques are applied. It includes gray level & contrast manipulation, noise reduction, edge crispening and sharpening, filtering, interpolation and magnification, pseudo colouring, and so on.

B. Extraction of Features

Edge Analysis

Edge analysis includes a variety of mathematical methods that aim at identifying points in a digital image at which the image brightness changes sharply or more formally, has discontinuities. The point at which image brightness changes sharply are typically organized into a set of curved line segments termed edges. In the image analysis, the edges are analyzed using the sobel operator in the MATLAB.

Visual Saliency Analysis

Visual saliency is the distinct subjective perceptual quality, which makes some items stand out from neighbours and immediately grab the attention. Centre surround saliency deals with the set of motivated features such as intensity, orientation, colour and motion. Dozens of feature channels are extracted and finally combined to form master saliency map.

Texture Analysis

Texture analysis can be used to find the texture boundaries. Texture analysis refers to the characterization of regions in an image by their texture content. The GLCM functions characterize the texture of an image by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image, creating a GLCM, and then extracting statistical measures from this matrix. For texture analysis, haralick function and the gray co-matrix function are used. The gray-level co-occurrence matrix (GLCM) includes calculating contrast, correlation, homogeneity, energy.

GLCM={contrast, correlation, homogeneity, energy}

C. Classification

SVM determine the decision boundaries in the training step and the method can provide good generalization in high dimensional input spaces. SVM classification is based on the concept of decision planes that define decision boundaries. A decision plane is one that separates between a set of objects having different class memberships. SVM finds the vectors ("support vectors") that define the separators giving the widest separation of classes. SVM classification supports both binary and multiclass targets. SVM models have similar functional form to neural networks and radial basis functions, both popular data mining techniques. However, neither of these algorithms has the well founded theoretical approach to regularization that forms the basis of SVM [14]. The quality of generalization and ease of training of SVM is far beyond the capacities of these more traditional methods. The SVMs map the original data points from the input space to a high dimensional, or even infinite-dimensional, feature space making classification problem simpler in feature space. The mapping is done by a suitable choice of a kernel function.

D. Hash Values

The hash values are calculated for the above extracted features. It is a numeric value of a fixed length that uniquely identifies data. Hash values represent large amount of data as much smaller numerical values. Hash values are useful tool for the examination, discovery and authentication of image. A unique numerical identifier that can be assigned based on the standard mathematical algorithm applied to the extracted feature. The most commonly used algorithms known asMD5 and SHA will generate numerical values so distinctive that the chance that any two data sets will have the same hash value, no matter how similar they appear, is less than one in one billion.

'Hashing' is used to guarantee the authenticity of an original data set and can be used as a digital equivalent. The basic idea behind hashing is to take a field in a record, known as key, and convert it through some fixed process to a numeric value is known as hash key. The numeric value will be in the range of 0 to n-1, where n is the maximum number of slots in the table. The fixed process to convert a key to a hash key is known as hash function. One common method of determining hash key is the division method of hashing.

Hash key = key %(number of slots in the table)

IV. RESULTS AND DISCUSSIONS

A. Training image set

A database is created and trained in the training phase with a number of images. The jpg and jpeg images either can be downloaded from the Internet or captured using digital cameras. The images can be of different sizes. The authority key is set after training images in the database. When any user tries to run the application

for checking the authenticity of query image, the user has to enter the same key, which was set during the training phase. Thus making sure that only authorized person can run the application. The data base is shown in Figure 2 and the results after processing are given in Figure 3.

Figure 2. Training image set

(a) (b)

(c) (d)

(e)

Figure 3. Results (a) Original Image (b) Noisy Image **(c) Pre processed Image (d) Edge Map (e) Saliency Map**

The output of textural analysis is shown in Figure 4.

Figure 4. Textural Features

The hash values are shown in Figure 5.

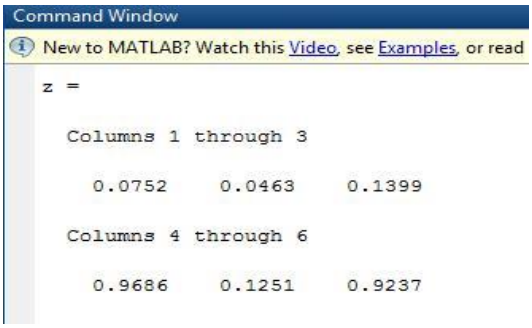

Figure 5. Hash values

The Hash value which are computed in previous step is given as input to the classifier. Then the SVM classifier classify whether the image is forged or real. The output is shown in Figure 6.

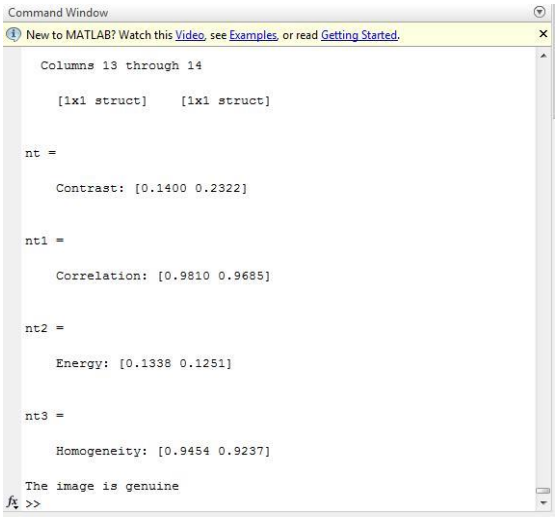

Figure 6. Classified output

Detection Rate

If the real image is considered as the positive sample and the spliced image is considered as the negative sample, the classifier will give four kinds of possible judgements:

(1) True Positive (TP): the number of positive samples which are predicted as positive sample;
(2) False Negative (FN): the number of positive samples which are predicted as negative sample;
(3) False Positive (FP): the number of negative samples which are predicted as positive sample;
(4) True Negative (TN): the number of negative samples which are predicted as negative sample;

TABLE I
DETECTION RATE FOR PROPOSED METHOD

Proposed Method	TP rate (%)	TN rate (%)	Accuracy (%)
	97.50	96.00	96.67

Table 1 illustrates the performance of the proposed method. This method can achieve a promising detection accuracy of 96.67% which is higher than the prior arts (in moment & HHT based technique detection date is reported as 85.86% and in chroma spaces based technique detection rate is reported as 94.70%).

V. CONCLUSION

In this paper, forgery image detection is proposed using image processing techniques. At first, the data base is created and trained in the training phase with a number of images. Then the images are converted into gray scale images. After the gray scale conversion, the edges are detected by using sobel operator. After detecting edges, Visual saliency maps are analyzed and also Textural features are extracted. After that, the hash value is calculated for the above extracted features. Then the output of the hash value is given as input to the SVM Classifier. The SVM Classifier classifies whether the image is forged or real.

In this proposed work, a new modified splicing detection scheme is proposed. To detect the spliced images, the distinguishing image features are extracted. The methodology implemented in this work reduces the computation time and maintains good accuracy. This technique could be a more adaptive to perform competitively with other techniques with maintaining very low computational complexity in the field of forensic sciences.

VI. REFERENCES

[1]. Ajaz Hussain Mir and Saba Mushtaq, (2014), 'Digital Image Forgeries and Passive Image Authentication Techniques', Proc. International Journal of Advanced Science and Technology.

[2]. Anita Sahani, Srilatha, K.(2014), 'Image Forgery Detection using SVM Classifier', Proc. International Journal of Advanced Research in Electrical Electronics and Instrumentation Engineering.

[3]. Cao, X. Hou,Y.Qu, Y. Zhang, W. Zhao, H. and Zhang,C. (2010), 'Detecting and extracting the photo composites using planar homography and graph cut', Proc. IEEE Trans Inf. Forensics Security.

[4]. Chen, W. Shi, Y. and Su, W. (2007), 'Image splicing detection using 2-d phase congruency and statistical moments of characteristic function', Proc. Of SPIE electronic imaging: security, stenography, and watermarking of multimedia contents.

[5]. Fang,Z. Wang, S. and Zhang,X. (2010), 'Image splicing detection using color edge inconsistency', Int. Conf. on Multimedia Information Networking and Security (MINES).

[6]. Fu, D. Shi, Y. (2006), 'Detection of image splicing based on Hilbert-Huang transform and moments of characteristic functions with wavelet decomposition', Proc. of International workshop on digital watermarking.

[7]. Ng, T. Chang, S. and Sun,Q. (2004), 'Blind detection of photomontage using higher order statistics', Proc. IEEE International symposium on circuits and systems (ISCAS).

[8]. Mandeep Kaur, ER. Tamana Sharma (2016), 'Forgery detection of spliced images using machine learning classifiers and color illumination', Proc. International Journal of Innovative Research in Science, Engineering and Technology.

[9]. Pradyumna Deshpande, Prashasti Kanikar, (2012), 'Pixel based digital image forgery detection techniques', Proc. International Journal of Engineering Research and Applications.

[10]. Rajath, B. Suniha, K. (2016), 'Survey on passive image tampering detection', Proc. International Advanced Research Science, Engineering and Technology.

[11]. Shuguang Zhang, Qiang Zhang, Weidong Min and Yongzhen Ke (2014), 'Detecting image forgery based on noise estimation', Proc. International Journal of Multimedia and Ubiquitous Engineering, Vol.9, No.1, pp.325-336.

[12]. Siwei Lyu, Xunyu Pan (2010), 'Region duplication detection using image feature matching', Proc. IEEE Transactions on Information Forensics and Security.

[13]. Zhao, X. Li, J. Li, S. and Wang, S. (2010) 'Detecting digital image splicing in chroma spaces', Proc. International workshop on digital watermarking.

[14]. Zhenhua, Q. Guoping, Q. and Jiwu, H. (2009) 'Detect digital image splicing with visual cues', Proc.International workshop on information hiding.