Image Forgery Detection using Local Binary Patterns

Amol Shinde[1], Mohit Saxena[2], Kalpesh Vishwakarma[3], Dr. Ashwini Kunte[4]

[1,2,3]*Final-year B.E.-EXTC, Thadomal Shahani Engineering College, Mumbai, India*
[4]*Professor, Department of EXTC, Thadomal Shahani Engineering College, Mumbai, India*

---***---

Abstract - *The paper proposes Image forgery detection using Local Binary Patterns. Local Binary Pattern (LBP) is basically used for feature extraction. A simple LBP operator is calculated in a rectangular window. The main advantage of this original LBP implementation is that we can gain extremely small details in the image. However, the* biggest downside of this algorithm *is that we cannot capture fine details at varying scales, only the fixed* 3 x 3 scale. *To handle this drawback of variable neighbourhood sizes we use an extension to the original LBP implementation, it has varying* p *and* r *which are used to construct Local Binary Patterns where p is the number of points 'p' in a circularly symmetric neighbourhood and r is the radius of the circle 'r', which allows us to have values at different scales after the LBP features are obtained. We generate a normalized Histogram which is called as LBP histogram and for the classification purpose linear SVM classifier is used which will determine whether the input image is authentic or forged.*

***Key Words*: Forgery, LBP, Histogram, SVM.**

1. INTRODUCTION

Image forgery means manipulation or tampering of the digital image to extract some meaningful or valuable information from it. Images play a vital role in several areas, including forensic investigation, a criminal investigation, surveillance systems, intelligence services, medical imaging, and journalism. But, in today's digital age, it is possible to very easily change the information represented by an image without leaving any visible traces of tampering [1]. Tampering of digital images is done for hiding some meaningful or useful information to forged images. When image tampering is done the authenticity of the image is lost. The digitally forged images are sometimes so genuine that they are nearly impossible to distinguish from the original image.

Since the authenticity of the image is lost. Integrity and authenticity verification of digital images is one of the popular and serious research issues in the field of image processing. There are so many image forgery techniques introduced over the years which are categorized into two approaches an active and a passive approach.

The active approach consists of Digital Watermarking and Digital Signature, Watermarking involves injecting a watermark which is used for the authenticity of the digital image which is inseparable from the image, whereas in digital signature some bit patterns are embedded in the digital image to avoid image tampering. The passive approach is the blind approach in which it never needs any prior information to include in the digital image. There are many methods in passive approach such as copy-move, splicing, image retouching and image re-sampling, but the most frequently used technique for image tampering is copy-move which aims to copy any content of the image and moving the same content in that image.

2. LITERATURE REVIEW

Digital imaging has experienced immense growth in late decades, and digital camera images have been used in a number of applications [2]. Detecting these types of tampering has become a serious problem at present. To determine whether a digital image is original or tampered is a significant challenge.

Luo et al. proposed a strong identification of the duplicated region in digital images in 2006 [3]. In this paper, the authors divide an image into overlapping blocks and then apply the similarity matching method on these blocks. The similarity matching method detects the copy-move forgery in the applied image.

Zhang et al. proposed a new method for copy-move forgery detection in a digital image in 2008 [3]. Authors used DWT and divided the given image into several non-overlapping sub-images and phase correlation is used to compute the spatial offset between the copy-move forgery regions. During this point, they applied similarity matching principle between the pixels for detecting tampered regions.

Kang et al. proposed a method to detect copy-move forgery in a digital image in 2010 [3]. The image is divided into subblocks and then applied to an improved SVD on each block. During this point, similarity matching method is used on each block based on the lexicographically sorted SV vectors. At last, the forged image region is detected.

M. Sridevi attempted to check the authenticity of the image using the image's necessary features like Markov and moment-based features [4]. This method was having the best results for splicing technique.

There is a technique based on the Radon transform and phase correlation which improves the robustness of the forgery detection [5]. The proposed method can detect forgeries even if the forged images were undergone some image processing operations such as scaling, Gaussian noise addition, rotation etc.

Rota et al. [6] proposed a passive deep learning approach based on convolutional neural networks (CNN) for forged image classification. They used the CAISA v2.0 dataset for experiments and achieved 97.44% detection rate.

3. IMPLEMENTATION

Image forgery detection is growing immensely in the field of image processing because the digital image forgeries are growing at an increasing rate in different fields of application and have made so much complications to accept the integrity and authenticity of the digital image. The proposed method aims to detect whether the input image is authentic or forged. To achieve this, the Local Binary Pattern (LBP) algorithm is used for feature extraction and SVM classifier is used for classification which presents maximum accuracy and has less complexity.

The proposed methodology starts with an input image which is applied to the pre-processing block which converts the input image into a grayscale image, then feature extraction process is executed which yields features of the image later it is presented as LBP histogram after this the SVM classifier does the prediction based on the LBP histogram of the input image and the trained images which concludes that the input image is authentic or forged. The proposed methodology which is adopted in the present work is shown in the Fig.1 in a very an abstracted manner with various blocks

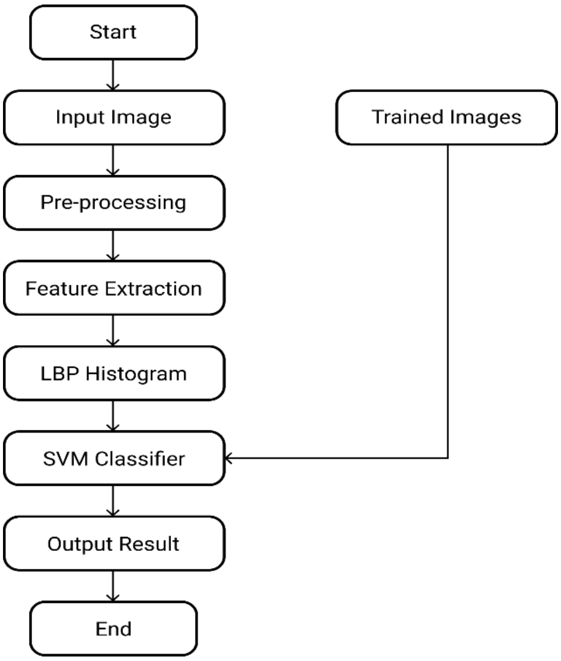


Fig – 1: Proposed methodology flowchart

I) Input Image

This methodology first starts with an input image. The input image is the test image which is to be tested whether it is an authentic image or a forged one. The input image can be of any type PNG, TIFF, JPEG etc

II) Pre-processing

The main purpose of pre-processing is to frame an image for feature extraction. The input image is converted into Grayscale image and this process comes under the pre-processing block. The grayscale images consist of only grey tones of colours, which are only 256 values where 0 stands for black and 255 stand for white, this process of conversion is done because the input image is usually an RGB image and when converted into a grayscale image, it's size is tremendously reduced and which helps to minimize the time occupied by the system.

III) Feature Extraction

Feature extraction is a necessary block because when you need to process a large amount of data, feature extraction reduces the details of the data without losing any relevant or essential information. When tampering is done, edges abnormalities are formed, which changes the texture of images, so there is a notable difference present in the texture of authentic and tampered images.

In the proposed method we are using a texture operator Local Binary Pattern (LBP) for feature extraction. So, the Grayscale image is divided into small blocks. LBP algorithm works with the given grayscale image and labels each pixel in the image by thresholding the neighbourhood pixels with the centre pixel and presenting the result as a binary number, but the real trouble is that the normal LBP cannot produce efficient output i.e. (it has a drawback while working in varying scales) so to overcome this obstacle we use an extension to the original LBP implementation, it has varying p and r which are used to construct Local Binary Patterns where p is the number of points 'p' in a circularly symmetric neighbourhood and r is the radius of the circle 'r', which produces binary values at varying scales, and these binary values are called as the LBP features of the image.

IV) LBP Histogram

LBP Histogram is nothing but graphical the representation of the LBP features of the input image. The input image, as well as the training images, go through the same above process to ensure that they can be used together in the upcoming block i.e. (SVM classifier).

V) SVM Classifier

In the proposed method, the detection of the forged image is done using a Support Vector Machine (SVM). An SVM is a machine learning algorithm that can be used for both classification and regression purposes. SVMs comprises of a hyperplane which separates or divides a dataset into classes, a hyperplane is a line that equally separates and classifies a set of data. It also has support vectors which are nothing but the points nearest to the hyperplane. As Fig.2. Shows an example of SVM classifier which has separated the dataset into two classes (Class A and B).

The SVM classifier uses two phases a) Training and b) Testing. In the training phase, a database is created and trained with a number of images which goes into several processes i.e. (Image -> Grayscale -> LBP features-> LBP Histogram) and in the testing phase, it goes with the same process as train phase except the test image is given as input image and later which is given to the SVM classifier, Here the SVM classifier does the classification based on the LBP histogram of the input image and the trained images.

Finally, the SVM does the prediction (classification) of the input image.

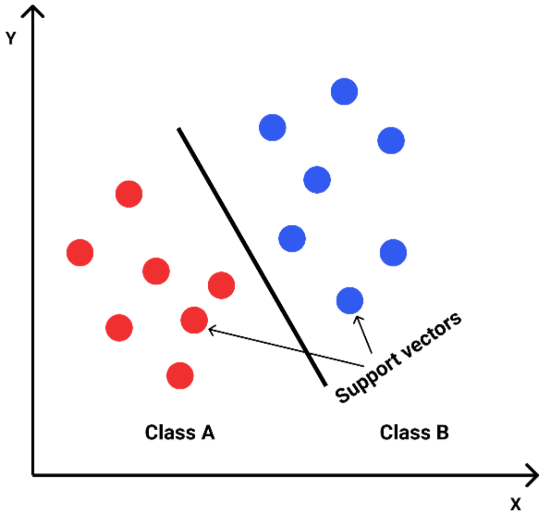


Fig - 2: Example of SVM classifier

VI) Output Result

Once the prediction of SVM classifier is done, it displays the image and reveals whether the test input image is authentic or forged.

4. EXPERIMENTAL RESULTS

4.1 Results

The proposed image forgery detection method is evaluated using different types of original and forged images. The images are taken from CASIA v2.0 dataset. The dataset we are using contains a total of 1755 images. Out of which 895 are authentic images and the remaining are forged images i.e. 860 images. The image sizes vary, and they are in PNG, TIFF, or JPEG formats. In more than half of the forged images, copy-move or spliced method is used as a forgery. Below the experimental results of each block are shown.

A) Input Image

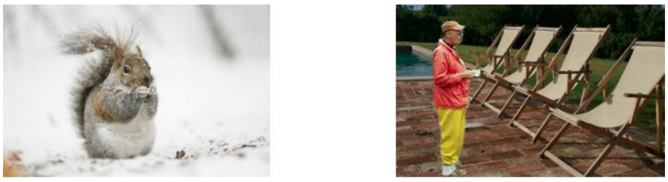


Fig - 3.1: Original Image

Two separate input images are shown as an example.

B) Pre-processing

Fig - 3.2 Gray scale Image

Original image gets converted into Grayscale Image.

C) Feature Extraction

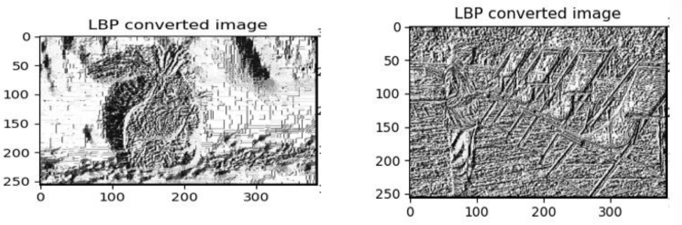


Fig - 3.3: LBP converted Image

Gray scale image gets converted into LBP converted image.

D) LBP Histogram

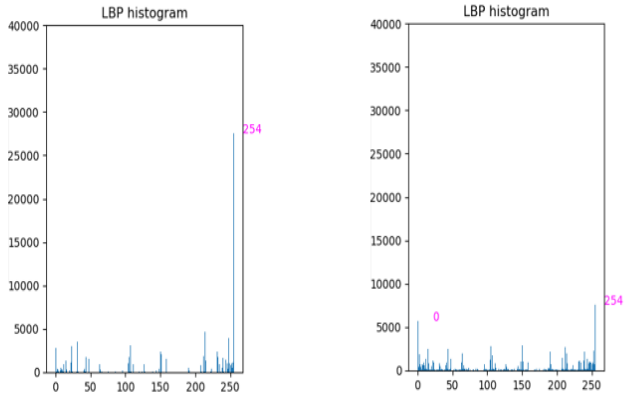


Fig - 3.4: LBP Histogram

LBP Histograms are generated by using LBP features of both the images.

E) SVM Classifier

As shown in Fig.2. SVM separates the dataset into two classes here SVM does the prediction of the input image based on the LBP histogram of both the images and concludes that whether the input image is authentic or forged.

F) Output Result

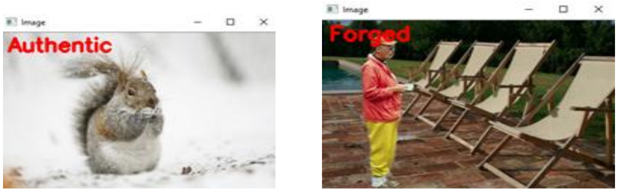


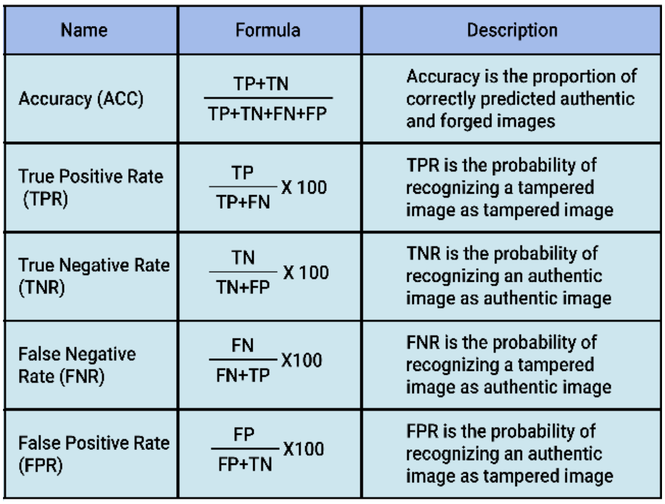
Fig - 3.5: Output of Original images

Fig - 3: Results of each block

After the SVM prediction, Fig.3.5. shows the output results indicating that the first image is an Authentic image and the second image is a Forged Image. Thus, from Fig.3. we can get a glance of results of each block.

4.2. Evaluation Criteria

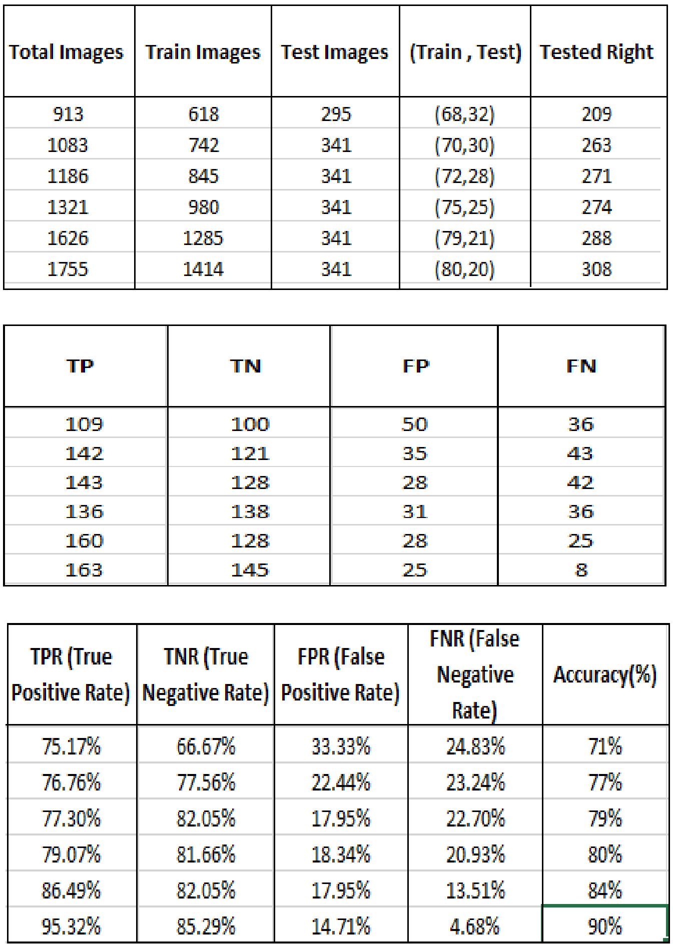
The performance of the proposed method is evaluated by Accuracy (ACC), True positive rate (TPR), true negative rate (TNR), False positive rate (FPR), False negative rate (FNR). These performance parameterers along with their formulas are defined below in Table-1 [7].



Name	Formula	Description
Accuracy (ACC)	$\frac{TP+TN}{TP+TN+FN+FP}$	Accuracy is the proportion of correctly predicted authentic and forged images
True Positive Rate (TPR)	$\frac{TP}{TP+FN}$ X 100	TPR is the probability of recognizing a tampered image as tampered image
True Negative Rate (TNR)	$\frac{TN}{TN+FP}$ X 100	TNR is the probability of recognizing an authentic image as authentic image
False Negative Rate (FNR)	$\frac{FN}{FN+TP}$ X100	FNR is the probability of recognizing a tampered image as authentic image
False Positive Rate (FPR)	$\frac{FP}{FP+TN}$ X100	FPR is the probability of recognizing an authentic image as tampered image

Table -1: Evaluation criteria Table

As shown in Table-2 we can see that initially, the accuracy was less, later we analyzed that as the train images were increased the accuracy of the system increases and not only accuracy increases it also reduces the chances of false positives and negatives, we get the accuracy of the system near about 90% as earlier the accuracy was less. Thus, we can conclude that as train images were increased proposed methodology made more efforts to increase the system accuracy.



Total Images	Train Images	Test Images	(Train , Test)	Tested Right
913	618	295	(68,32)	209
1083	742	341	(70,30)	263
1186	845	341	(72,28)	271
1321	980	341	(75,25)	274
1626	1285	341	(79,21)	288
1755	1414	341	(80,20)	308

TP	TN	FP	FN
109	100	50	36
142	121	35	43
143	128	28	42
136	138	31	36
160	128	28	25
163	145	25	8

TPR (True Positive Rate)	TNR (True Negative Rate)	FPR (False Positive Rate)	FNR (False Negative Rate)	Accuracy(%)
75.17%	66.67%	33.33%	24.83%	71%
76.76%	77.56%	22.44%	23.24%	77%
77.30%	82.05%	17.95%	22.70%	79%
79.07%	81.66%	18.34%	20.93%	80%
86.49%	82.05%	17.95%	13.51%	84%
95.32%	85.29%	14.71%	4.68%	90%

Table -2: Calculations

5. CONCLUSION

This paper highlights image forgery detection using Local Binary Patterns (LBP). The detection was done by undergoing several processes like pre-processing which converted the input image into a grayscale image, then feature extraction process was executed which produced features of the image it was later then represented as LBP histogram, after that the SVM classifier was used for the prediction purpose, In the end, it displayed whether the image is authentic or forged. The performance of the proposed method was calculated which revealed that the accuracy was near about 90% and it had less false negatives.

ACKNOWLEDGEMENT

We are very much thankful to our project guide Dr. Ashwini kunte for helping us to understand the topic theoretically and practically and guiding us throughout the project.

REFERENCES

[1] M. Sridevi, C. Mala, Siddhant Sanyam, “Comparative Study of Image Forgery and Copy-Move Techniques”.

[2] Ashwin Swaminathan, Min Wu, K. J. Ray Liu “Digital Image Forensics via Intrinsic Fingerprints”, March 2008.

[3] Abhishek Kashyap, Rajesh Singh Parmar, Megha Agarwal, Hariom Gupta, “An Evaluation of Digital Image Forgery Detection Approaches”, 30 Mar 2017.

[4] M. Sridevi, C. Mala and S. Sandeep “Copy – move image forgery detection”, 2012.

[5] Hieu Cuong Nguyen and Stefan Katzenbeisser "Detection of copy-move forgery in digital images using Radon transformation and phase correlation”, 2012.

[6] Rota, P., Sangineto, E., Conotter, V., Pramerdorfer, C.: Bad teacher or unruly student: can deep learning say something in image forensics analysis? 2016.

[7] Monowar H. Bhuyan, Dhruba K. Bhattacharyya, Jugal K. Kalita, “Network Traffic Anomaly Detection and Prevention”.