International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Image Forgery Detection Using Copy-Move Technique

Md. Iftekhar Hossian[a] *Md Tasnim Alam*[b], *Jyotirmoy Ghose*[c]

[a]Lecturer, North Bengal International University, Chowddopai, Natore Road, Motihar, Rajshahi, Bangladesh.
[b] Lecturer, North Bengal International University, Chowddopai, Natore Road, Motihar, Rajshahi, Bangladesh
[c] Lecturer, North Bengal International University, Chowddopai, Natore Road, Motihar, Rajshahi, Bangladesh

DOI: https://doi.org/10.55248/gengpi.2023.32077

ABSTRACT

A digital image is a data representation of a two-dimension scene. Digital images are easy to manipulate and edit due to the availability of powerful image processing and editing software. So, it is possible to add or remove important features from an image without leaving any obvious traces of tampering. In today's world of advanced computer technology, tampering with a digital image can be easily performed by a novice with a number of available sophisticated image processing software like Adobe Photoshop, Corel Draw, etc. So, it is very difficult for a viewer to judge the authenticity of a given image. In fields such as forensics, medical imaging, e-commerce, and the film industry, the authenticity and integrity of digital image is essential. In the medical field physicians and researchers make diagnoses based on imaging. In this paper, we proposed a method for image tamper detection that is based on copy move technique. Copying parts of an image and pasting in the same image for creating a fake image is called copy move.

Keywords: Forgery, DCT, Robust Match, and Lexicographical Sorting.

1. INTRODUCTION

Nowadays, with the reputation of low-cost and high-resolution digital cameras, digital media is playing a more and more important role in our daily life, however, due to the sophisticated editing software (for example, Photoshop, 3D Max), digital images can be easily manipulated and distorted without leaving visible clues, thus, it poses a grave social problem as to how much of their content can be trusted, whether it is authentic or tampered especially as a witness in a courtroom, insurance claims and scientific deception. An example of Copy move is shown in fig 1. In the last decade, any photographs were generally acknowledged as a proof for any crime. Computers nowadays however are being used in almost all fields of business, banking, agriculture, health and many other domains to up- keep report in the form of digital images. But due to easy accessibility of efficient tools, no one nowadays is confident for the integrity of images. Presently there is no efficient method to test the authenticity and integrity of any digital image which can be put as proof in a court. Due to a vast range of image forgery methods and tools, it has become very hard to have a single tool which can handle detection of all facets of image forgery. Detecting image forgery is an rising field of research. In the last decades, many image manipulations were seen on the cover of magazines, news documents, and other media resources. In respect of cases being at present reported, three approaches are famous as commonly being used for image forgery. These are i) Image Enhancement (e.g. blurring the portions, brightness, Contrast or Color etc.), ii) Image Compositing (Mixing up diverse features from two or more different images) and iii) Copy move forgery (Copying some part of image from image itself and pasting it in the same image elsewhere.). The present paper focuses on the last type of forgery.

2. METHODOLOGY

In our proposed forgery scheme, we have introduced a combine DCT and Robust match method to detect the image forgerydetection. Since the forgery regions are unknown both in size and shape, if we compare every possible pairs pixel by pixel, the computational complexity will be very higher and none can continue that. Obviously, it is more practical to divide the image into blocks for detecting the forgery regions. In order to take an efficient detection, some proper and robust features must be extracted from the blocks, therefore, a high-quality features extraction can not only represent the whole blocks, but also reduce the length of feature vector, and what's more, make the detection algorithm has lower computational complexity.

2.1 Discrete Cosine Transformation (DCT):

DCT is one of the most accepted linear transforms on digital image processing. It has been broadly used because of its good ability of energy compression. The DCT transforms a image from a spatial representation into a frequency representation. The DCT is conceptually similar to the DFT, except the DCT does a improved job of intent energy into lower order coefficients than does the DFT for image data. The DCT is purely real, the DFT is complex (magnitude and phase). A DCT operation on a block of pixels produces coefficients that are similar to the frequency domain coefficients produced by a DFT function. Assuming a periodic input, the value of the DFT coefficients is spatially invariant (phase of the input does not matter). This is not true for the DCT. The Discrete Cosine Transform (DCT) has special property that most of the visually important information of the image is concentrated in just a few coefficients of the DCT. Moreover, DCT based watermarking techniques propose compression while DWT based watermarking techniques propose scalability

The DCT can be defined by the following equation:

Two Dimensional (2D) DCT:

$$F(x,y) = C(x)C(y)\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} f(i,j)\cos\left[\frac{\pi(2i+1)x}{2N}\right] * cos\left[\frac{\pi(2j+1)x}{2N}\right] (2.1)$$

Where$C(x), C(y) = \begin{cases} \sqrt{\frac{1}{N}}, \; x,y = 0 \\ \sqrt{\frac{2}{N}}, \; otherwise \end{cases}$

2.2 Robust Match:

The idea for the robust match is similar to the exact match except we do not arrange and match the pixel representation of the blocks but their robust representation that consists of quantized DCT coefficients. The quantization steps are calculated from a user-specified parameter Q. This parameter is equivalent to the quality factor in JPEG compression, i.e., the Q factor determines the quantization steps for DCT transform coefficients. Because higher values of the Q-factor lead to finer quantization, the blocks must match more closely in order to be identified as similar. Lower values of the Q-factor produce more matching blocks, possibly some false matches. The image is scanned from the upper left corner to the lower right corner while sliding a B×B block. For each block, the DCT transform is calculated; the DCT coefficients are quantized and stored as one row in the matrix A. The rows of A are lexicographically sorted as before. Because quantized values of DCT coefficients for each block are now being compared instead of the pixel representation, the algorithm might find too many matching blocks (false matches). Towards this goal, if two consecutive rows of the sorted matrix A are found, the algorithm stores the positions of the matching blocks in a separate list (for example, the coordinates of the upper left pixel of a block can be taken as its position) and increments a shift-vector counter C. Formally, let (i1, i2) and (j1, j2) be the positions of the two matching blocks. The shift vector s between the two matching blocks is calculated as s = (s1, s2) = (i1 – j1, i2 – j2). We increment the normalized shift vector counter C by one: C(s1, s2) = C(s1 , s2) + 1 . Then the algorithm finds all normalized shift vectors s(1), s(2), …, s(K), whose occurrence exceeds a user-specified threshold T: C(s(r)) > T for all r = 1, …, K. For all normalized shift vectors, the matching blocks that contributed to that specific shift vector are colored with the same color and thus identified as segments that might have been copied and moved. For color images analyzed, it is first converted to a gray scale image using the standard formula I = 0.299 R + 0.587 G + 0.114 B, before proceeding with further analysis.

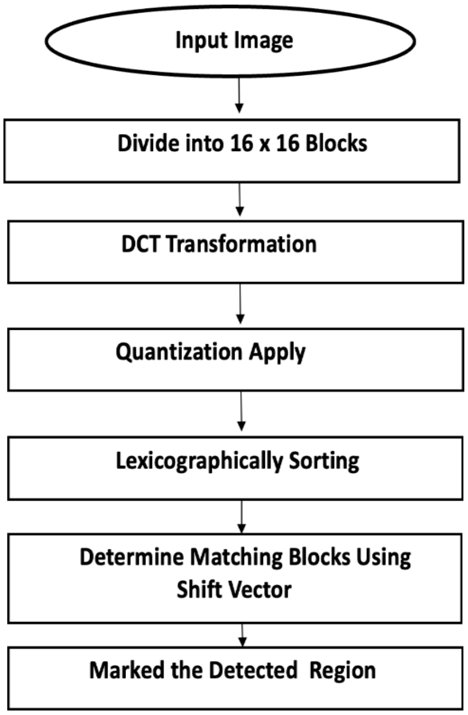
Input Image
Divide into 16 x 16 Blocks
DCT Transformation
Quantization Apply
Lexicographically Sorting
Determine Matching Blocks Using Shift Vector
Marked the Detected Region

Fig 2.2: Block diagram of the image forgery detection

According to the above discussion, the whole detection framework is given as follows:

Take a input image as grey scale image, if it is RGB image first converted to a grey scale image using the standard formula I = 0.299 R + 0.587 G + 0.114 B, before proceeding with further analysis and we assume a quantized 8 x 8 matrix and extend it 16 x 16 matrix.

Dividing the image into fixed-size blocks

DCT is applied to each block to generate the quantized coefficients

Divide the DCT coefficients matrix by the extended 16 x 16 matrix and get the DCT quantized matrix.

Apply the lexicographical sorting the DCT quantized matrix and searching similar blocks.

Finding correct block and increment the shift vector counter and finally show the output blocks above the threshold.

3. RESULT AND DISCUSSION

The entire process is proposed in the preceding section was implemented in C^{++} and executed on a computer of CPU 2.2GHz with. The test images used are free from any type of other tampering which we proposed in this approach. Tests were conducted on such hundred images. In the present study, all these hundred images are copy moved forged images without any further modifications. We have considered a square block of size 16*16 pixels and checked for different threshold values of discrimination say 5, 10 and 15.

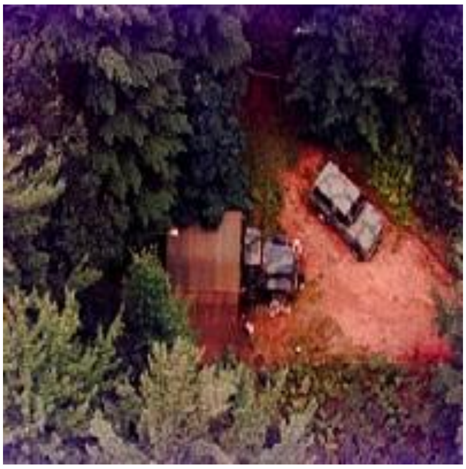

Fig 3.1: Original image

For example, the image taken in figure 3.1 is an original image without having any forgery. Image size is 1021 x 922. Block size selected by us is 16 x 16. In the image, the truck is hidden by pasting some portion of tree groves onto truck area. There was no other after manipulation except copy move. Fig 3.2 is showing the result of detection of copy moved region. The result obtained is very accurate and fast. The final image is found after performing erosion and dilation. Count for shift vector is taken 2. We performed various attempts on the same image with different threshold values and with different counts for shift vectors.

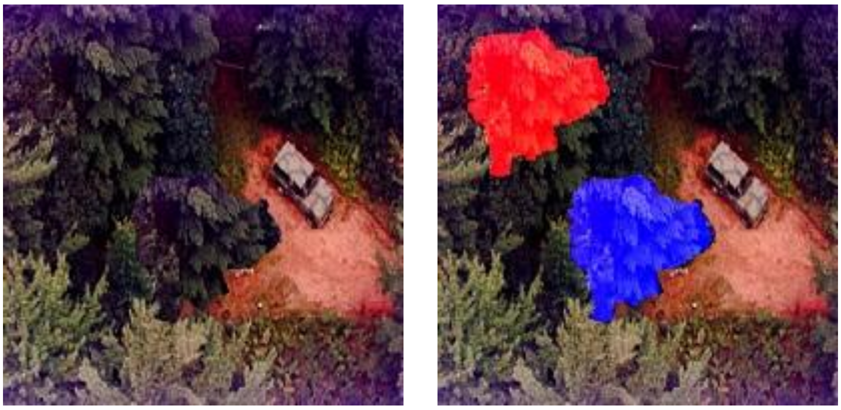

Fig 3.2: Result of Copy move without any Modifications

The result shown accordingly is very accurate. There were no modifications after performing copy move. Threshold set for above experiment is 5.

For another example, the image taken in figure 3.5 is an original image without having any forgery. Image size is 825 x 705. Block size selected by us is 16 x 16. In the image, there is no other after manipulation except copy move. (Please see Figure 3.6). Fig 3.7 is showing the result of detection of copy moved region. The result obtained is very accurate. Count for shift vector is taken 4. We performed various attempts on the same image with different threshold values and with different counts for shift vectors and find it threshold value 6.

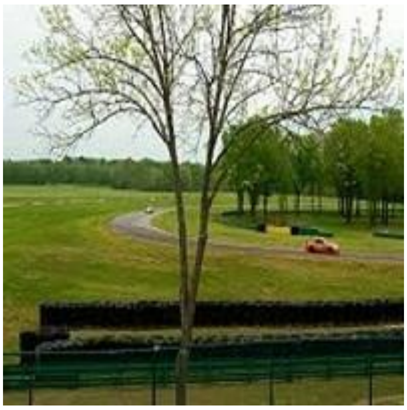

Fig 3.3: Original image

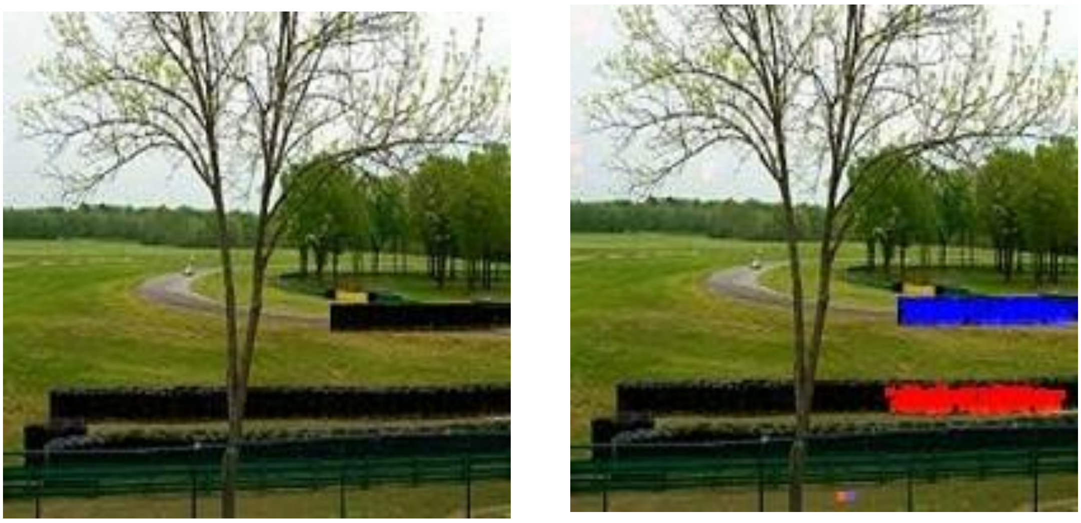

Fig 3.4: Result of Copy move without any Modifications

In our experiment computational cost reduced. The results are improved from two directions one is decomposition of image and another is by using less features of blocks. The result is faster and within few minutes it gives the result.

Method	Detection time(sec)
[1]	900
[2]	850
[7]	500
Proposed method	300

Table 3.1: Detection time of various methods

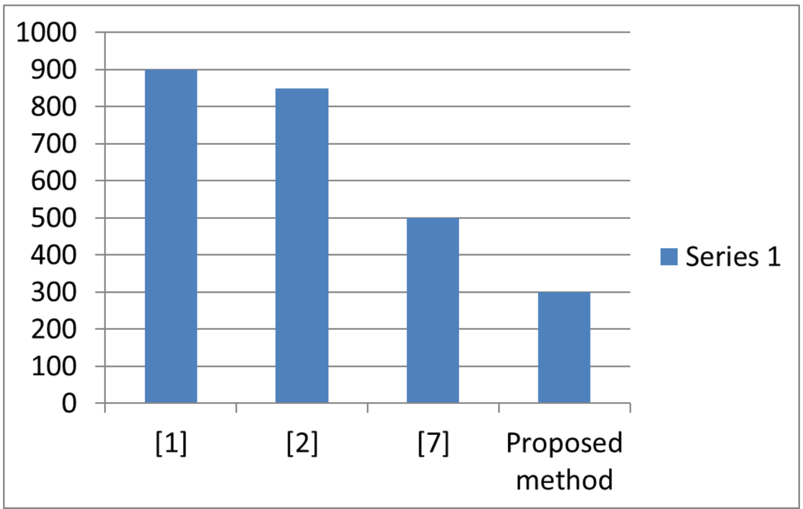

Table 3.2: Time taken (seconds) to detect the duplicated region

4. CONCLUSION:

When blood flow to a portion of the brain is disrupted due to a damaged or clogged blood artery, a stroke ensues. Hemorrhagic or ischemic strokes are both possible. When a blood vessel in the brain ruptures or breaks, blood leaks into the brain, causing a hemorrhagic stroke. When a blood channel transporting blood to the brain is stopped or restricted by severely narrowed arteries or a blood clot, an ischemic stroke develops. In this work, I analyze all the factors influencing this disease (STROKE) to extract their influence with exploratory data analysis and classifier models. In the dataset there are more female candidates than male. Basically, my work is based on the difference between male female candidates of how stroke can take effect on them. Private job holders are in big numbers but Self-employed candidates are usually affected by stroke than others. Older persons are more suffered from the risk of getting stroke. Hypertension patients are more suffered from the risk of getting stroke. Surprisingly, Unmarried persons are more suffered from the risk of getting stroke. Glucose level of 250-300 is more suffered from the risk of getting stroke. Persons with body mass 40-60 are more suffered from the risk of getting stroke. Persons who are formerly smoked and regularly smokes are more suffered from the risk of getting stroke. I take the accuracy tests also for successful project work. In Here I found that is the

References

[1] "*Detection of Copy-Move Forgery in digital Images*", J.Fridrich, D. Soukal, and J. Lukas, *Proc.* Of DigitalForensic Research Workshop, Aug. 2003.

[2] " *Exposing digital forgeries by detecting duplicated image regions*", A.C. Popescu and H. Farid, Technical Report TR2004-515, Dartmouth College, Aug. 2004.

[3] "*A robust detection algorithm for copy-move forgery in digital image*" ,Yanjun Cao , Tiegang Gao , Li Fan , Qunting Yang , College of Information Technical Science Nankai University, Tianjin 300071, China.

[4] **"***Pixel Based Digital Image Forgery Detection Techniques*" , Pradyumna Deshpande , PrashastiKanikar,Vol. 2, Issue 3, May-Jun 2012 .

[5] **"***A Novel Approach for Forgery Detection of Images*",Vimal Raj V 1, Lija Thomas 2,Volume 2, Issue 8, August 2013.

[6] "*Image Tamper Detection Scheme Using QR Code and DCT Transform Technology*", Ji-Hong Chen and Chen-Hsing Chen,International Journal of Computer, Consumer and Control (IJ3C), Vol. 1, No.2 (2012).

[7] "*Fast Copy-Move Forgery Detection*", Hwei-Jen Lin, Chun-Wei Wang, Yang-Ta Kao, WSEAS TRANSACTIONS on SIGNAL PROCESSING, Issue 5, Volume 5, May 2009, pp. 188-197.

[8] "*Digital Image Processing*", Gonzalez, Third Edition