International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-12, October 2019

Image Forgery Detection and Localization using Modified JPEG Ghost

Pooja Patil, Namita Pulgam, Vanita Mane

Abstract*—Authenticity of an image taken digitally suffers severe threats as a result of increase in various powerful digital image editing tools. These tools modifies the image contents without leaving footprint of such modifications. We come up with a technique that analyzes digital image forgery detection in JPEG images which goes through multiple compression. Nearly all digital devices uses JPEG as a standard storage format to maintain the storage space. JPEG is a lossy compression standard. By using any image processing tools, when assailant changes any part of a JPEG image and save it, the alter part of the image has different compression artifacts. JPEG ghost algorithm is used to detect disparity in JPEG blocks that rise from improper alignments of JPEG blocks respect to original structure and detect local footprint of JPEG compression. In our work, our proposed technique will modify JPEG ghost detection to detect and localize digital image forgery.

***Keywords*—Authenticity, JPEG Ghost, Image forgery**

I.INTRODUCTION

Digital Forensics science has one of its part known as digital image forgery. It emphasizes on discovering and re-pairing forgery in an image[1]. The main reason why images are forged so easily are due to low budget tools, software and hardware which are easily available and easy to use hence any alteration and manipulation are not easily traceable to human eyes. It may not possible to distinguish whether a given digital image is original or a modified. Use of this kind of forgery is widely used in fields like Journalism, media, publications, social networks, Investigations where the data is manipulated accordingly. Image splicing and copy-move are two most familiar tech- niques for producing image forgeries. In image splicing forgery some parts from base image is copied and pasted it to another source of image. Other type is copy move that paste copied part in the same image to add or hide content.

Revised Manuscript Received on October 30, 2019.
***** Correspondence Author
Ms. Pooja Patil*, Department of Computer Engineering, Ramrao Adik Institute of Technology Mumbai, India Email:erpatilpooja@gmail.com
Prof. Namita Pulgam, Department of Computer Engineering Department Ramrao Adik Institute of Technology Mumbai, India Email:ashwinigudewar@gmail.com
Prof. Vanita Mane, Department of Computer Engineering Department Ramrao Adik Institute of Technology Mumbai, Email:vanitamane1@gmail.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license http://creativecommons.org/licenses/by-nc-nd/4.0/

Normally, some processing done before like scaling and rotation or some after like blurring and adding noise on copied part and then paste it in image using editing tools. This discard inconsistencies from image.So no one could recognize it as forged one. There are two techniques to classify Image Forgery detection: active (intrusive) and passive (non-intrusive). Forgery in intru- sive techniques can be detected by verifying the integrity of pre-embedded digital watermark or signature. Non-intrusive techniques depends on the original features of the image for forgery detection.
So passive techniques are used widely. But some post pro-cessing and compression weakens the system performance because these ways are intuitive to a definite forgery detection type. Therefore, new blind techniques are necessary to separate authentic image from forged one[2].
Our method gives better results of forgery detection in low-quality images. Also it has some drawbacks i.e lack of automa- tion and issues with non-aligned Discrete Cosine Transform (DCT) blocks. Our aim is to remove limitations, by proposing an modified version of JPEG ghost which adds automation to algorithm.
The paper is further subdivided in sections mentioned bellow: Literature survey is part of Section II. Problem definition is presented in section III. Actual method is proposed in Section IV. The final conclusion is in Section V.

II.LITERATURE REVIEW

H. Farid[3] came up with a way to detect double JPEG compression by further compressing the image at different quality factors and found one of those with the same quality factor as the tampered image produced a JPEG ghost indicating the forged region. The paper[4] author derive a model for the probability distributions of DCT coefficients in case of single and double compression and then a probability map is used to distinguish between original and forged areas. In paper[5] JPEG forgery detection and localization technique based on finding an optimal error matrix image that clearly depicts the forged regions is proposed. The paper[6] propose JPEG ghost detection by adding some post-processing and iterations to it. In this, SE-MinCut segmentation algorithm extract JPEG ghost borders and then the classifier classify forged and tampered region.
As studied above we have gone through various image splicing techniques for JPEG images. Out of which we have consider JPEG Ghost which can further enhance to overcome some limitations.

III. PROBLEM DEFINITION

Images found on the internet are probably result of image tampering because there is no control over the user distributed content. Several algorithms have been developed to evaluate such kind of problems but in reality it becomes quite a tedious job to detect the forged image accurately. Also image forgery detection i.e. to know if an image is tampered or not is one problem while localizing where the forgery actual is; becomes another major problem to be solved. Web

images that reach to us have generally go through at least one JPEG re-save. It may be by some mediator or automatically by the publisher. So we choose to work with JPEG images to detect forgery.The benefits of JPEG ghost detection method is that it gives better results of forgery detection in low-quality images. Also it has some drawbacks as:

1.This method is of lack of automation.

2.If ghost area in comparison to entire forged image is limited then it is bit difficult to distinguish actual ghost and other dark spot highlighted due to low intensity values in authenticate image.

We are going to overcome this limitation and propose the modified JPEG Ghost which will add automation to algorithm also it validate actual ghost size.

IV. PROPOSED SYSTEM

This section presents The proposed system as follows.

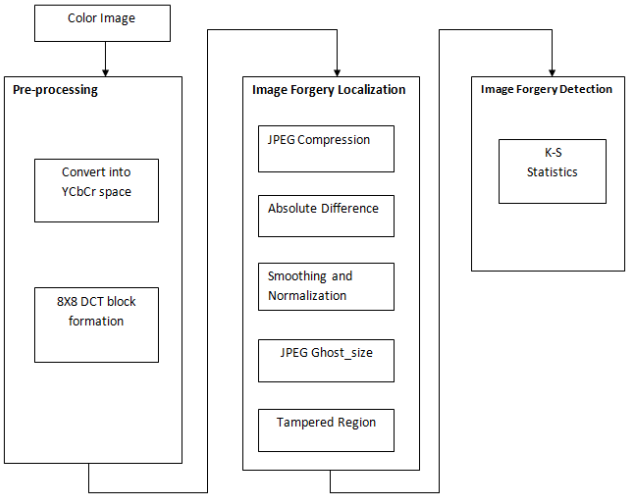

Fig. 1. System Architecture

Primarily a transformation of RGB color image into luminance/chrominance space (YCbCr) is done in standard JPEG compression format. Both the chrominance channels (CbCr) Compression, Smoothing and Normalization, Determine JPEG

Ghostsize and Localize tampered region and then apply K-S statistic for Image Forgery Detection. In these techniques we use the modified JPEG Ghost algorithm to detect JPEG Ghost. Finally we get whether the image is forged or not and then normalized into the range[0, 1] at each location

(x, y) so as to maintain averaged difference:

are com monly sub sampled by a factor's of 2 proportional to the luminance channel (Y). Particular channel is then further segregate into 8 × 8 pixel blocks. Each block using 2-D DCT is transform to frequency space.

B. Image Forgery Localization

Image from last preprocessing phase is given as a input in this phase. This phase include JPEG Compression, Smoothing and Normalization, Determine JPEG Ghost _size and Localize tampered region.

In a composite forged image, by the use of editing tools, two JPEG images with different quantization tables were spliced together. Then, resultant image is resaved in JPEG by another quantization table. Normally it uses a higher quality quantization table. This result in double-quantization of some portions of image which are known as the forged regions and other portions are quantized by a higher quality factor which are original one. Such a forged image is known as double- quantized image. For detecting double-quantized parts of an image, each DCT coefficient, c is then quantized by an amount q:

$$C = \text{round}(c/q) \quad (1)$$

Here the quantization q is based on the spatial frequency and channel. Image degradation is caused due to larger quanti- zation values of q which gives better compression. We consider the absolute difference calculated from the pixel values, other than calculating difference of quantized DCT coefficients as below:

$$d(x,y,q) = \frac{1}{3}\sum_{i=1}^{3}[f(x,y,i) - f_q(x,y,i)]^2 \quad (2)$$

The proposed system composed of following phases. As shows in fig, it takes the input as image which we plan to detect for forgery. The image undergo prepossessing steps which includes RGB image conversion into YCbCr space, 8×8 DCT block formation. Then the prepossessed image is given as input to Image Forgery Localization which include JPEG

Where $\mathbf{f}(x, y, i)$, $i = 1, 2, 3$ denotes each of three color channels and $\mathbf{f}_q(.)$is the result of compressing $\mathbf{f}(.)$at quality q. Spatial average and normalised difference measure are con- sidered for replacing these differences. The difference image is first averaged across a b × b pixel region:

$$\sigma(x,y,q) = \frac{1}{3}\sum_{i=1}^{3}\frac{1}{b^2}\sum_{b_x=0}^{b-1}\sum_{b_y=0}^{b-1}[f(x+b_x, y+b_y, i) - f_q(x+b_x, y+b_y, i)]^2$$

$$d(x,y,q) = \frac{\sigma(x,y,q) - \min_q[\sigma(x,y,q)]}{\quad} \quad (4)$$

A. Preprocessing

Pre-processing is important to improve of image data and enhance features important for later detection. Most pre-processing steps that are implemented are either to reduce the noise, to reconstruct an image, to perform morphological operations and to convert the image to binary/grayscale so that operations can be easily implemented on the image. Pre- processing steps are as follows.

$$\max_q[\sigma(x, y, q)] - \min_q[\sigma(x, y, q)]$$

After normalization it is important to validate actual ghost size because it is tedious to distinguish between actual ghost and dark spot arriving in authenticate image. So it uses algorithm to calculate JPEG ghost _ size. This algorithm is work as follows.

To segment image into different segments graph based segmentation approach is used. In this approach graph edges, and some undirected edges of neighbouring pixels are choose.

Dissimilarity between pixels can be measure using weights on each edge. Segmentation criterion of the neighbouring area depending on degree of variability is adjusted by technique.

By re-compressing the same image at different quality factors and then subtracted from original image. Effective average of three color channel is calculated by performing subtraction performed at each individual color channel. The final image shouldn't have ghost region too large as complete image or too small as noise all through the image. Hence we take into consideration only those images in which size of region is more than 1/8 compared to minimum size segment or less than twice of maximum size segment of original segmented image. Rest other are rejected. This ghost area obtained during segmentation will overlap with one of the segment. The ghost is considered as valid ghost only if size is within the given range identifying it as forged one.

Ghost _size defines that ghost is a part that is used to differentiate any dissimilarity in an image that can be a result of image forgery. There are various types of forgery like, Different compression level in same image; Different smoothing and sharpening edges due to copy paste forgery; Change in quantization level like one having 0-255 level of quantization another having 0-1024 levels. Our modified JPEG ghost is capable of detecting combination of all three image forgery effectively. There are various types of forgery like, Different compression level in same image; Different smoothing and sharpening edges due to copy paste forgery; Change in quantization level like one having 0-255 level of quantization another having 0-1024 levels. Our modified JPEG ghost is capable of detecting combination of all three image forgery effectively. Complete Automation Algorithm to validate JPEG Ghost _size is as given below.

1.I_g= Graph _Seg (I) // Segment of image I
2. Max _Seg _size= Maximum size of image segment I
3. Min _Seg _size= Minimum size of image segment I
4. Forged=0
5. For q=1:Q // Quality of JPEG image
a. Recompress Image I at JPEG quality q to get Image I_q
b. $I_d = I - I_q$ // Subtract recompressed image from original
compresses image I
c. Perform averaging on Image I_q . d. Normalize the I_q between 0 to1 e. Perform opening in Image I_d
f. g _size= Size of largest component present in Image I_g
g. (If g _size>=Min _Seg _size/8 && g _size <=Max Seg
_size×2)
i. seg _g _size=segment size of I_g overlapping with I_d
ii. if |seg _g _size - g _size| <=5000
forged=1 i. end-if
6. end-for

C. Image Forgery Detection

This phase performs Kolmogorov Smirnov (K-S) statis- tic[8]. The similarity of distribution is conclude by (K-S) statistic. (K-S) statistic assist in finding the accuracy in finding out the forged region, and its accurateness to correctly categorize an authentic image. Our approach uses (K-S) statistic because it can work without a large database of images to train a support vector machine (SVM). Also it is computationally much simpler. Kolmogorov-Smirnov Statistic: This statistic can be used to obtain the combined distribution function of the two distributions that are to be compared. The Kolmogorov-Smirnov k is a particularly simple measure. It can be con- sidered as the maximum value of the absolute difference between either of functions used for cumulative distribution. For comparing two different cumulative distribution functions $C_1(u)$ and $C_2(u)$, the K-S statistic k is given by,

$$k = \max_u |C_1(u) - C_2(u)| \quad (5)$$

where $C_1(u)$ and $C_2(u)$, in cumulative difference d(x,y,q) these are the combined probability distributions , where indi- vidual values of q are to be considered . The image is classified as tampered if K-S statistics on exceeding a threshold which are specified. The selection criteria of this threshold is to give output that is less than 1% false positive rate.

V. CONCLUSION

We propose a improved JPEG splicing detection and local- ization technique, which covers two aims. First, it enables the slicing detection and localization processes to be completely automated. Second, the proposed technique is capable of detecting and localizing multiple forgeries in a JPEG image. We come up with a new technique for automatic image splicing detection and localization, based on JPEG ghost detection. It is difficult to detect JPEG Ghost by manual search. So system provides image forgery detection based on Modified JPEG ghost Algorithm. By using Modified JPEG Ghost, results were significantly higher than that of other recent JPEG image forgery detection methods. Also the approach gives significant true positive rates and at the same time a lower false positive value.

REFERENCES

1. A. Redi, W. Taktak, and J. Dugelay. “Digital image forensics: a booklet for beginners.”,Multimedia Tools and Applications, vol. 51, no. 1, pp.133162, May 2011.
2. S. Katzenbeisser and P. Fabien. “Information hiding techniques for steganography and digital watermarking.” Artech house, 2000.
3. H. Farid. “Exposing digital forgeries from JPEG ghosts.” ,IEEE Trans- actions on Information Forensics and Security, vol. 4, no. 1, pp. 154160, Feb. 2009.
4. Bianchi, T., Rosa, A.D., and Piva, A. “Improved Dct Coefficient Analysis For Forgery Localization in Jpeg Images.”,International conference on acoustics, speech and signal processing 2011. p. 24447.
5. Diangarti Bhalang Tariang and Ruchir Naskar. “ReCompressed based JPEG Forgery Detection and Localization through Automated Quality Factor Investigation.”,IEEE WiSPNET 2016 conference.
6. Sepideh Azarian-Pour, Massoud Babaie-Zadeh, Amir Reza Sadri. “An Automatic JPEG Ghost Detection Approach for Digital Image Foren- sics.”, 2016 24th Iranian Conference on Electrical Engineering (ICEE).

AUTHORS PROFILE

Ms. Pooja Shantaram Patil, B.E in Computer Engineering and Pursuing M.E. in computer engineering from RAIT, Mumbai University. Her research area includes Image Processing, digital communication. She has 4 years academic experience.

Mrs. Namita Damodar Pulgam, M.E. in Computer Engineering and Pursuing Ph.D in Computer Engineering from RAIT, Mumbai University. She is presently working as a Assistant professor in Department of Computer Engineering, RAIT. She has published 10 plus research papers in conferences and journals. Her research area includes Image Processing, Pattern Recognition and Cyber security. She is member of ISTE committee.

Mrs. Vanita Manik Mane, has Pursuing Ph.D in computer engineering from RAIT, Mumbai University. She is presently working as a Assistant professor with department of Computer Engineering in RAIT. She has 16 years academic experience and almost 50 research papers at international conference and journals. Her research area includes Image Processing, Pattern Recognition and Cyber security. She is member of professional bodies like CSI and ISTE.