Digital Image Forensics Using Image-Similarity

ISSN (e) 2520-7393
ISSN (p) 2521-5027
www.estiri.com

Juveria Aslam[1], Moazzam Jawaid[2], Shahnawaz Talpur[3]

[1, 2, 3] *Department of Computer Systems Engineering, Mehran University of Engineering and Technology, Jamshoro*

Abstract: Digital image integrity and authenticity are becoming increasingly critical issues. Image forgeries are becoming quite simple to make. In order to demonstrate the integrity and validity of digital photographs, digital image forensics is required**.** Since there are so many effective methods for manipulating images, their veracity is being called into doubt, particularly when they are used to support claims for insurance, medical industry, or legal proceedings. By using numerous high-tech mechanisms discovered in the existing research, picture forensic approaches ascertain the integrity of photographs. In this study, the photographs are scrutinized for a specific kind of forgery in which a portion of an image is duplicated or hidden by being pasted onto another image. This type of forgery is called Copy-Move Forgery. In this paper, straightforward methodologies have been applied to determine the traces of forgery in the images with less code complexity and time consumption. The results of this research show that the suggested technique successfully detects several copy-move forgeries and accurately determines the copy-move forgery even when the images are tainted with noise, compression, and blurring. Therefore, the suggested method offers a computationally effective and trustworthy method of copy-move forgery detection that raises the credibility of photos in applications that focus on providing proof and contain sensitive information.

Keywords: *Digital Image Forensics, Image similarity, Image processing, Copy-Move forgery*

1. Introduction

Things have started to work more in the digital mode where images play a key part, especially after the spread of COVID'19. These photos include sensitive information and can be used in a variety of settings, including courtrooms, hospitals, banks, and educational organizations. For the purposes of quality assurance, verification, authorization, and identity, most sectors rely on digital data. Given that we're discussing digital image processing, it's impossible to ignore the reality that advances in this discipline have also been exploited to make highly plausible fabricated photos. Several types of research have been undertaken in this area in order to improve the system's dependability and efficiency with regard to the use of images. Nonetheless, the usage of image similarity in this area is the most useful.

As our process of transferring as well as information exchange globally has become considerably more sophisticated and practical thanks to the internet. However, this virtual world has also given criminals a forum to carry out illegal acts. As more people have access to computers globally, cybercrime is growing in importance and posing a significant challenge to law enforcement. There is a critical need to comprehend and improve the current investigation methods and mechanisms for preventing cybercrime because they are still not foolproof and have had limited effectiveness in prosecuting lawbreakers. One of them is digital image forensics. Images are a great source of information that are dispersed over the internet.

Although modification of images is quite easy. Attackers alter the photographs applying modern photo manipulation tools in order to distort or hide their significance. These software programs are widely accessible today, not just on desktop and laptop computers but also on handheld mobile devices. The society, the administration, and companies are all at risk because of this, which is a serious worry. Therefore, it is necessary to validate these photos. The field of *Digital Image Forensics* (*DIF*) entails verifying the integrity and provenance of the digital image in addition to its content [1]. Over the past ten years, *DIF* has become incredibly important to the research community. Finding the source and determining the authenticity of captured images are the basic issues that digital image forensics approaches try to fix.

The goal of *Digital Image Forensics* is to identify forgery traces in order to improve the validity of digital images. It focuses on how to make images authentic enough to deliver trustworthy responses or conclusions. Image forgery is a process that adds, changes, or even removes portions or characteristics of an original image covertly. The majority of the time, copy-move forgery is employed to tamper with photos, hence this study concentrates on that. The image is faked by copying and pasting pieces of the same image (of any size and shape).

The two main techniques used to modify images are region duplication by copy-move forgeries and image splicing [4]. Image splicing is the process of combining portions of different images to produce a fabricated image. To hide or amplify certain crucial content in the portrayed image, copy-move forgery, on the other hand, involves copying and pasting image portions onto the same image. It becomes difficult to distinguish the tempered sections from legitimate parts because replicated regions appear to be identical with suitable components (such as color, brightness, contrast, and noise).

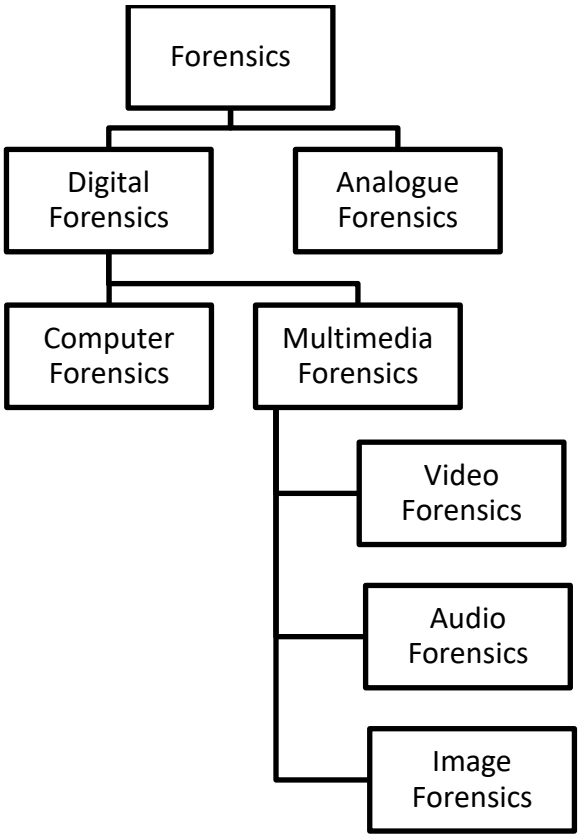


Figure 1. The ontology of forensics.

In some circumstances, tampering with images might be quite detrimental. In 2008, for example, an image of Iranian ballistic missiles was altered to depict four instead of three missiles. This created restlessness globally until it was revealed that the images have been forged to depict the wrong message. At the start of the Iraq War in 2003, the Los Angeles Times published a huge image in which the soldier was seen to be aggressively ordering the folks. Before it was discovered to be a forgery, the image was published in several additional newspapers and created global tension. It was, later on, discovered that $Brian\ Walski$ [3], the photographer, enhanced the image to make it more dramatic. Due to several scenarios like this, the scientific community has developed an interest in digital picture authentication and verification.

2. Related Work

Image Similarity detection has been a wide subject for a decade now due to a huge variety in its applications as well as techniques. However, there are some generalized approaches that have been followed in all the techniques to measure image similarity, like Comparing frames in a video, pattern recognition, and homographic transformation for image stabilization. For all the defined approaches, it is recommended in [7], to extract image features prior to comparing similarities as it saves time and cost.

Digital image forensics is a special field that has just evolved and looks for signs of forgery in images [6]. Digital image forensics' main objective is to examine the images for evidence of fraud using either active or passive (blind) techniques [2]. The information that is implanted in advance in the photos is what the active approaches, like watermarking [7] and digital signatures [6], rely on. The deployment of active approaches in practice may be constrained, nevertheless, due to the lack of information [8]. In addition to specifying images that don't need any prior knowledge about them, passive techniques are used [8–10]. The two main techniques used to modify images are region duplication by copy-move forgeries and image splicing. Image splicing is the process of combining portions of different photographs to produce a fabricated and forged image.

Forgery using copy-moves is another name for copy-paste forgery. The integrity of digital photographs has been significantly impacted in the last ten years by copy-paste forgeries. Studies focus a lot of work on identifying this form of counterfeit because of this. This kind of counterfeit fabrication involves copying and/or moving a portion of the original image to a different spot inside the same image. For this, there is a high correlation in between copied pasted sections, which might be used as proof of counterfeit. Locating effective algorithms to locate features and match these traits for finding associated segments, however, is the challenging task.

In this regard, $Hough\ Transform$ has been opted in [8]. It proposes sectioning the images into grids and then applying Hough transform which captures the features of different grids of the image. All the features are counted and the whole image is represented as a vector containing numeric values. Moreover, the vectors generated from two images are compared in the process. This technique shows an accuracy of 80% when the picture is sectioned into 100 grids, and it is expected to grow with the number of grid sections. However, the number of images used to test the accuracy is quite small.

In [19] the researcher introduced the copy-move forgery detection method employing DCT over small, overlapped blocks for perhaps the first time. The $DCT\ coefficients$ are used to create the feature vectors. Following lexicographic sorting of the feature vectors, blocks' similarity is examined. In [13], principal component analysis is used to represent visual blocks (PCA). The authors used nearly half as many features as [11] by utilizing one of the $PCA's$ properties. Although it makes this approach effective, copy-move forgery with rotation was not caught. [15] proposes a sorted neighborhood method based on the discrete wavelet transform (DWT). The picture is split into 4 subsets, as well as the feature vector is obtained by applying Singular Value Decomposition (SVD) to the low frequency components.

The approach is only resilient to JPEG compression up to quality level 70. A method for extracting the block features and $kd-tree$ matching is presented in [16] and therefore is focused primarily on blur moment invariants up to seventh order. According to [12], applying the scaling as well as rotating invariant $Fourier-Mellin\ Transform$ (FMT) in conjunction with bloom filters to the image blocks can help identify fake images. In order to reduce the dimension of the feature vector for forgery detection, a truncating method was introduced in [14] to introduce an enhanced DCT-based methodology. [17] suggests a method for detecting image forgeries using DCT as well as SVD.

A technique to identify Copy-move forgeries, a sort of tempering that would be frequently exploited to manipulate

digital photographs, was put forth by [15]. This technique involves copying and pasting a portion of an image into another area of the image. This study explains an effective non-intrusive method for copy-move forgery detection. This approach is based on dyadic wavelet transform background subtraction as well as similarity detection ($DyWT$). Applying $DyWT$ as well as statistical metrics, it is possible to identify the structural similarities between copied and pasted regions. The results demonstrate that this strategy works better than state-of-the-art approaches. The method used in this paper successfully detects tiling on images without the need for camera information or a sizable image database. The method can be applied to backgrounds and textures with complex geometry.

In digital photo modification, copy-move is frequently used. An effective copy-move detecting method with some post-processing resistances was put up by [27]. $Non-negative\ matrix\ factorization$ (NMF) coefficients typically recovered from a list of all the fixed-size overlapped blocks after the picture has been partitioned into the units of a certain size. To lower the likelihood of an invalid match, we employ the lexicographical sorting approach. Each block pair's hamming distance is measured throughout the matching process, but unless the distance is below a predetermined level, we identify the block pair as being in the tampering area.

The most typical kind of image fraud is copying and pasting, in which one portion of a picture is changed for a different portion of the same image. An effective method focused on translation invariant elements was put out by [7]. These mostly rely upon that Trace transform and are accomplished by making numerous changes to the $MPEG-7$ image identification set descriptions. This method of detecting fake images is therefore quite effective.

A top method for spotting copy-move fraud with rotation was put up by [12]. $Polar\ Harmonic\ Transform$ would be used to feature extraction of such spherical blocks, that are being used to carry out block matching. For rotated and noisy figures, this method works.

A technique to identify copy-move forgeries was put out by [25] and is focused on Multi-resolution $Weber\ law\ descriptors$ (WLD). The suggested multi-resolution WLD pulls patterns using chrominance parts, that can provide additional details that human vision cannot see. The suggested technique's prediction accuracy can be up to 91 percent when using a multi-resolution WLD descriptor on the image's chrominance space.

To identify picture copy-move forgeries, in [26] researchers suggested a technique relying on non-block matching. Utilizing phase correlation is done in this paper. The results of the studies indicate that the technique is effective in identifying duplicated image regions and is fairly resistant to impulse noise as well as blurring.

Advanced forgery techniques are day by day reaching near perfection, leaving a very small chance of ever getting caught. For such high-end forgery, the detective tools also need to be smart and advanced. There have been multiple approaches in this regard that result in partial success. Forgery can be induced in an image by using the techniques like copy-move, morphing, splicing, enhancing, retouching, and so on [22].

The system proposed in [10], detects forgery induced by cloning and splicing. It combines multiple previously used algorithms, which we're working separately for both types of forgeries on different sizes of images and carries out a system that works collectively for both. However, the time consumed by the system needs to be reduced.

Moreover, image similarity can be further used for multiple applications as in [11], cherry quality is being detected for different categories, and in [12], a rainfall forecasting model has been formulated. More methodologies and applications can be seen in [13] [14] [15]. This research depicts image similarity as a vast topic that must be explored further for better and to design efficient systems.

To sum up the above-provided literature review, the concept of image similarity for detecting the possibility of image tampering and fast retrieval of the image is gaining interest across the globe. The research will be directed towards making the current systems, as stated above, more efficient, and robust.

Only translated, rotated, and scaled versions of the original photos were used to test the method, leaving the actual copy-move forged images out. The splicing algorithm has some speed issues, and the cloning algorithm does not function on areas that have been previously processed. The application is difficult and time-consuming overall. However, research has been done on copy-move fraud that is visually obvious and computationally dependable, as the paper claims. Prior to extracting the important information from an image, the image is segmented. Copy-move areas are then matched in two steps. The iterative process reduces the detection scheme while making the model more complex [3, 5].

The approaches now in use function satisfactorily, but they require a certain amount of dataset manipulation, complexity, and time, which makes them unsatisfactory for usage in delicate situations [8, 10, 12]. The COVID'19 left a significant mark on our lives, and it will continue to have an impact on how the world uses digital data moving forward. Therefore, cutting-edge, and trustworthy methods that can deliver quick digital image forensics are urgently needed. It is past time to implement some advanced strategies to check the validity and correctness of image data while making the best use of computer-based techniques quickly and accurately.

3. Methodology

To detect the forgery in different types of images, multiple image similarity techniques have been used in this research. Different types of forged images have been used. The Copy-Move-Forgery-detection ($CoMoFoD$) database was employed in this study. It comes from a public repository that is allowed to be used for research purposes. There are 260 forged image pairings in total, including 60 huge photos and 200 little ones. Each image has a distinct amount of complexity. Large photos are 3000 x 2000 pixels, whereas small images are 512 x 512 pixels. There are 3,361 photos in total that can be used in combinations. Scaling, Translation, Rotation, Contrast adjustment, image blurring, and other techniques were used to create the forgery. Images are

labeled 001 O for original and 001 F for forged in the $CoMoFoD$ Dataset. The Dataset hierarchy goes as mentioned below:

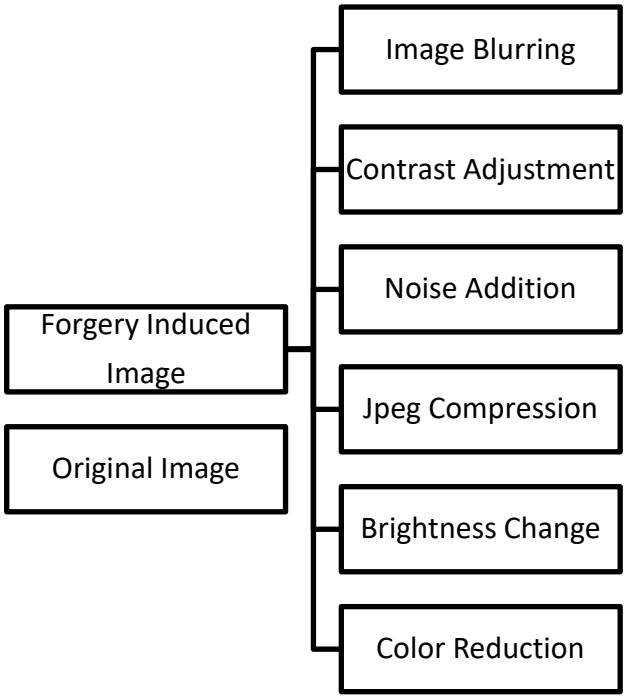


Figure 2. The hierarchy of CoMoFoD Dataset.

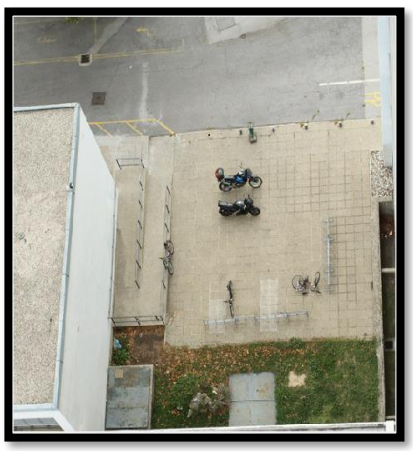
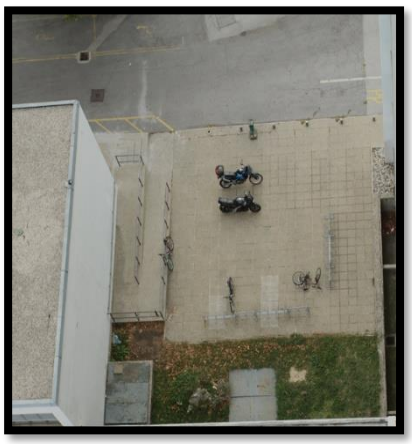
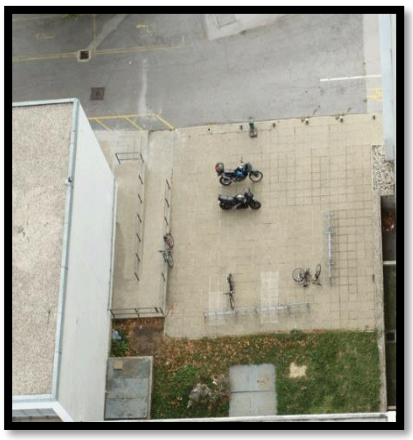
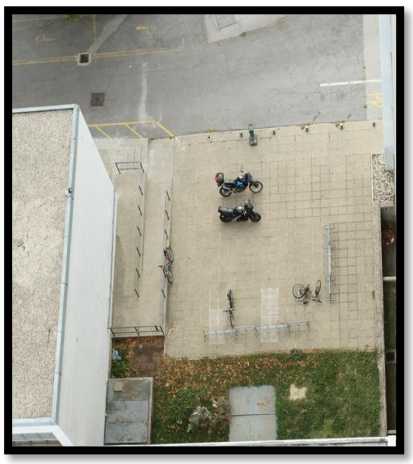
	Original Image
	Forged Image

	Image Blurring
	Contrast Adjustment
	Color Reduction
	Jpeg compression

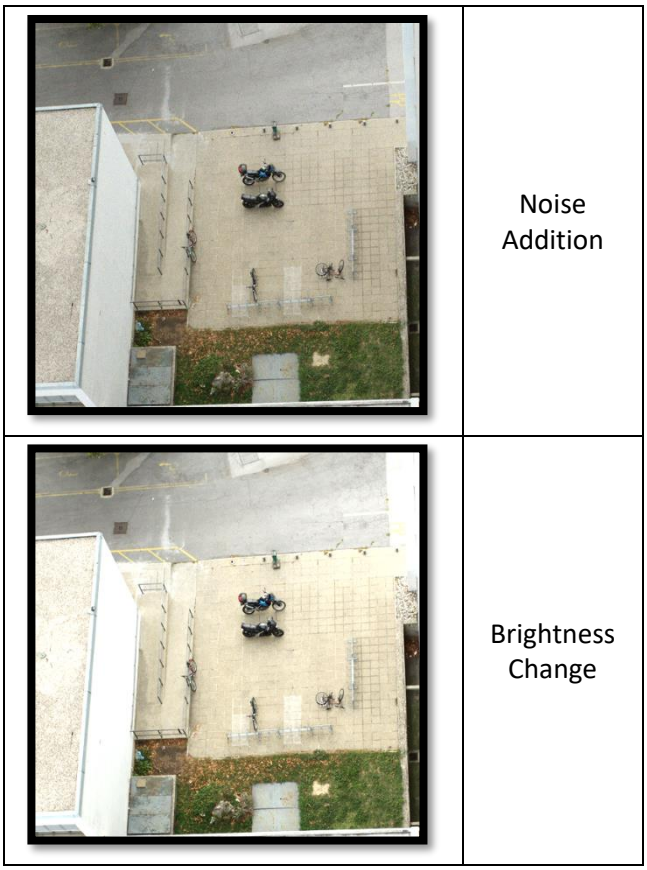


Figure 3. The sample images of CoMoFoD Dataset

3.1 Image Subtraction

Absolute image subtraction has been used on different image combinations to detect forgery with different preprocessing techniques. Techniques like image blurring, brightness, or contrast adjustment are commonly used to eliminate the chances of forgery being detected easily with the naked eye. The image forgery is detected easily in images where there is a huge visible difference however some color reduction and image blurring pre-processing techniques make this technique less efficient. It is not reliable as the results carry out false forged regions or do not detect the forgery at all.

$$Subtracted\_image(i,j) = |Image_1(i,j) - Image_2(i,j)|$$

There are only a few images, image subtraction works perfectly. While working with more intricate or processed photos, its accuracy declines.

4(a) original Image 4(b) Forged Image 4(c) Image Subtraction Output

4(d) original Image 4(e) Forged Image 4(f) Image Subtraction Output

4(g) original Image 4(h) Forged Image 4(i) Image Subtraction Output

Figure 4. The image subtraction from original minus forged image dataset.

3.2 Image Correlation

Image correlation is a more efficient way of detecting image forgery. The correlation coefficient determines the rate of similarity between the two given images. It detects the dissimilarity no matter how minute or invisible it is.

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

$$\rho(A,B) = \frac{1}{N-1} \sum_{i=1}^{N} \left(\frac{A_i - \mu_A}{\sigma_A} \right) \left(\frac{B_i - \mu_B}{\sigma_B} \right)$$

Where,

A=Original Image

B= Forged Image

Due to its potential to be an affordable, straightforward, yet precise method, digital imaging correlation (DIC) is a methodology that may prove to be perfectly suited for the investigation of fracture propagation and material deformation in practical applications. Even with the use of preprocessing methods, image correlation still provides good results for forgery detection. It recognizes even the slightest difference between the images. Image Correlation yields the following results:

Table I. Correlation of original and forged image.

Image Name	Original vs. Forgery induced image
001_	0.9946
002_	0.9896
003_	0.9963
004_	0.9995
007_	0.9992
011_	0.9956
012_	0.995
027_	0.9412
031_	0.9864
033_	0.9867
037_	0.9994
050_	0.9014
059_	0.981

Table II. Correlation of Image Blurring, Jpeg Compression and Noise Addition with Original Image.

Image Name	Image Blurring	Jpeg Compression	Noise Addition
001_	0.9391	0.9923	0.992
002_	0.9495	0.985	0.9863
003_	0.9643	0.9951	0.9952
004_	0.9485	0.9942	0.9932
007_	0.9664	0.9977	0.997
011_	0.9071	0.9896	0.993
012_	0.9835	0.9944	0.9941
027_	0.7966	0.9324	0.9387
031_	0.92	0.9747	0.9809
033_	0.9393	0.9851	0.9848
037_	0.9848	0.9987	0.9973
050_	0.6788	0.8602	0.8882
059_	0.9152	0.9765	0.9767

Table III. Correlation of Brightness Change, Color Reduction and Contrast Adjustment with Original Image.

Image Name	Brightness Change	Color Reduction	Contrast Adjustment
001_	0.9792	0.9941	0.9945
002_	0.9834	0.989	0.9895
003_	0.98	0.9961	0.9963
004_	0.9995	0.9984	0.9994
007_	0.9746	0.9988	0.9991
011_	0.9874	0.9952	0.9956
012_	0.9783	0.9949	0.995
027_	0.9323	0.9408	0.9412
031_	0.9831	0.9854	0.9863
033_	0.9768	0.9863	0.9867
037_	0.9949	0.999	0.9993
050_	0.9013	0.8993	0.9013
059_	0.9645	0.9803	0.9809

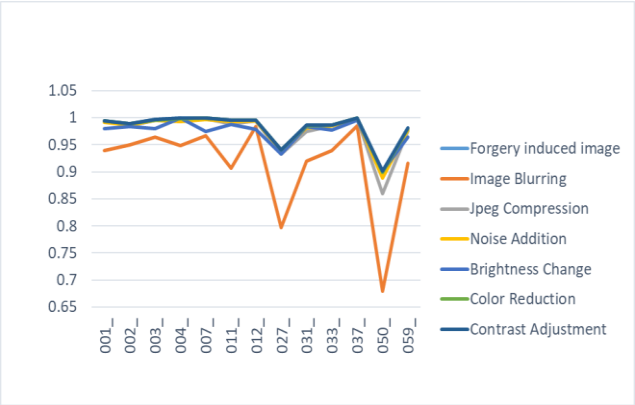


Figure 5. Comparison of correlation of CoMoFoD dataset.

4. Results and Discussion

In practical terms, a detection method's capacity to distinguish between genuine and fake photos is its most important feature. With a wide range of algorithms that are being introduced, a sophisticated one is needed to evaluate in simple terms the traces of forgery. The computation of precision "p" indicates the chance that a forgery identified is truly a forgery. The precision for the proposed system lies around 80.7% and has a low computational time.

5. Conclusion

A common and well-liked method for removing or hiding an object from a digital photograph is copy-paste forgery. In this research, we present an automated and efficient method to identify fake images. The experiments have been conducted using pictures of various types. In this paper, our main focus was on determining how to guarantee the detection of copy-move fraud in digital photographs. This paper's primary goal was to successfully identify forgery evidence in the images. Using image subtraction and image correlation, a straightforward methodology has been used to identify image forgery. Both produce results in various ways, but time consumption is minimal. The suggested strategy successfully locates the fabricated locations. It has been discovered that the suggested method can identify fabricated sections even after they have performed small-scale operations like image subtraction as well as image correlation. The approach is not entirely robust to it, though. The suggested algorithm will eventually be completely robust to the subtraction and correlation of images.

References

[1] alZahir, S., Hammad, R. Image forgery detection using image similarity. Multimedia Tools Applications 79, 28643–28659 (2020).

[2] Raneem Tassia, "Adaptive and Improved Approach for Image Forgery Detection," International Journal of Engineering Research & Technology (IJERT), vol. 09, no. 09, pp. 455-460, 2020.

[3] Toqeer Mahmood, Tabassum Nawaz, Rehan Ashraf, "Copy-Move Forgery Detection Technique for Forensic Analysis in Digital Images," Hindawi Publishing Corporation Mathematical Problems in Engineering, vol. 2016, 2016.

[4] S. Sadeghi, "State of the art in passive digital image forgery detection: copy-move image forgery," Springer-Pattern Analysis and Applications, 2018

[5] Jian Li, "Segmentation-Based Image Copy-MoveForgery Detection Scheme", IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, Vol. 10, NO. 3, March 2015.

[6] E. Ardizzone, "Copy–Move Forgery Detection by Matching Triangles of Keypoints", IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 10, NO. 10, OCTOBER 2015.

[7] A. Kashyap, "An Evaluation of Digital Image Forgery Detection Approaches," Jaypee Institute of Information Technology, 2017.

[8] S. TEERAKANOK, "Copy-Move Forgery Detection: A State-of-the-Art Technical Review and Analysis", IEEE Access, Vol. 7, 2019.

[9] Amira Baumy, "A Discriminative Statistical Model for Digital Image Forgery Detection"

[10] Xiang Lin, "Recent Advances in Passive Digital Image Security Forensics: A Brief Review ", Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company, 2018.

[11] Anil Dada Warbhe, "A Scaling Robust Copy-Paste Tampering Detection for Digital Image Forensics", 7th International Conference on Communication, Computing and Virtualization, 2016.

[12] Anil Dada Warbhe," Computationally Efficient Digital Image Forensic Method for Image Authentication", International Conference on Information Security & Privacy (ICISP2015), 11-12, Dec 2015.

[13] O. Mayer and M. C. Stamm, "Forensic Similarity for Digital Images," in IEEE Transactions on Information Forensics and Security, vol. 15, pp. 1331-1346, 2020

[14] Huang, HY., Ciou, AJ. Copy-move forgery detection for image forensics using the superpixel segmentation and the Helmert transformation. J Image Video Proc. 2019, 68 (2019).

[15] Kumar, Sunil & Nagori, Swati. (2017). Key-point based copy-move forgery detection in digital images. Journal of Statistics and Management Systems. 20. 611-621. 10.1080/09720510.2017.1395181.

[16] T. Chihaoui, S. Bourouis and K. Hamrouni, "Copy-move image forgery detection based on SIFT descriptors and SVD-matching," 1st International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), Sousse, 2014, pp. 125-129.

[17] S. M. Fadl, N. A. Semary and M. M. Hadhoud, "Copy-rotate-move forgery detection based on spatial domain," 9th International Conference on Computer Engineering and Systems (ICCES), Cairo, 2014, pp. 136-141.

[18] Kashyap, Abhishek; B. Suresh; Agrawal, Megha; Gupta, Hariom; Joshi, Shiv Dutt, "Detection of splicing forgery using wavelet decomposition," IEEE International Conference on Computing, Communication and Automation (ICCCA), Noida, 15-16 May 2015, pp. 843-848.

[19] J.-C. Lee, C.-P. Chang, and W.-K. Chen, "Detection of copymove image forgery using histogram of orientated gradients," Information Sciences, vol. 321, pp. 250–262, 2015.

[20] M.Hussain, S.Qasem, G. Bebis, G.Muhammad,H.Aboalsamh, and H.Mathkour, "Evaluation of image forgery detection using multi-scale weber local descriptors," International Journal on Artificial Intelligence Tools, vol. 24, no. 4, Article ID 1540016, 2015.