https://doi.org/10.48047/AFJBS.6.13.2024.472-481

African Journal of Biological Sciences

Journal homepage: http://www.afjbs.com

Research Paper Open Access

IMAGE COPY-MOVE FORGERY DETECTION BASED ONMACHINE LEARNING

Shimal Das[1], Jhunu Debbarma[2] ,Bibhash Roy[3], Bharat Sarkar[4]

[1,2,3]Deptt. of Computer Sci. & Engineering,
[1,2,3,4]Tripura Institute of Technology, Narsingarh, India,
[4]Deptt. of Civil Engineering,
Corresponding author: DR. SHIMAL DAS
Email: shimalcs.tit@gmail.com

Article History
Volume 6, Issue 13, 2024
Received: 18 June 2024
Accepted: 02 July 2024
doi:
10.48047/AFJBS.6.13.2024.472-481

Abstract
Digital images are easy to manipulate and edit due to availability of powerful image processing and editing software. Nowadays, it is possible to add or remove important features from an image without leaving any obvious traces of tampering. As digital cameras and video cameras replace their analog counterparts, the need for authenticating digital images, validating their content and detecting forgeries will only increase. Detection of malicious manipulation with digital images (digital forgeries) is the topic of this paper. In particular, we focus on detection of a special type of digital forgery – the copy-move attack in which a part of the image is copied and pasted somewhere else in the image with the intent to cover an important image feature. If we can detect any forged region in an image then we can consider the image as fake. In this project work, we investigate the problem of detecting the copy-move forgery and describe an efficient and reliable detection method. The method may successfully detect the forged part even when the copied area is enhanced/retouched to merge it with the background and when the forged image is saved in a lossy format, such as JPEG. The performance of the proposed method is demonstrated on several forged images.

Keywords: Digital forgeries, Copy-move attack, Video surveillance, Deep learning, Machine Learning

INTRODUCTION

A copy-move forgery is created by copying and pasting content within the same image, and potentially post-processing it. In recent years, the detection of copy-move forgeries has become one of the most actively researched topics in blind image forensics. A considerable number of different algorithms have been proposed focusing on different types of post-processed copies.

With the advancements in imaging technologies, the digital images are becoming a concrete information source. Meanwhile, a large variety of image editing tools have placed the authenticity of images at risk. The ambition behind the image content forgery is to perform the manipulations in a way, making them hard to reveal through the naked eye, and use these creations for malicious purposes. For instance, in 2001, after the 9/11 incident, several videos of Osama bin Laden over the social media were found counterfeited through the forensic analysis [1]. In the same way, in 2007, an image of tiger in forest forced the people to believe in the existence of tigers in the Shanxi province of China. The forensic analysis, however, proved the tiger to be a “paper tiger” [2]. Similarly, in 2008, an official image of four Iranian ballistic missiles was found to be doctored, as one missile was revealed to be duplicated [3]. Hence, the famous saying “seeing is believing” [4], [5] is no longer effective. Therefore, ways that can ensure the integrity of the images especially in the evidence centered applications are required.

Digital forensics is a key part of proving digital media authenticity. Media content such as an image or a video may be presented as evidence of a crime in the courtroom, In that condition, Digital forensics analysis is needed, during this analysis forgery detection algorithms play a vital role. Video surveillance is one of the technologies which can be used for security reason and monitoring. In such situation, there is a tendency of criminal's to do suspicious activities and always tries to alter the CCTV footage to hide their presence from the scene. So the proposed research work is a promising research field. Several computer vision applications, including Fake media content identification from social media and, prove the authenticity of an image or video in the courtroom.

Copy-Move Forgery

Because of the extraordinary difficulty of the problem and its largely unexplored character, the authors believe that the research should start with categorizing forgeries by their mechanism, starting with the simple ones, and analyzing each forgery type separately. In doing so, one will build a diverse Forensic Tool Set (FTS). Even though each tool considered separately may not be reliable enough to provide sufficient evidence for a digital forgery, when the complete set of tools is used, a human expert can fuse the collective evidence and hopefully provide a decisive answer. In this paper, the first step towards building the FTS is taken by identifying one very common class of forgeries, the Copy-Move forgery, and developing efficient algorithms for its detection. In a Copy-Move forgery, a part of the image itself is copied and pasted into another part of the same image. This is usually performed with the intention to make an object “disappear” from the image by covering it with a segment copied from another part of the image. Textured areas, such as grass, foliage, gravel, or fabric with irregular patterns, are ideal for this purpose because the copied areas will likely blend with the background and the human eye cannot easily discern any suspicious artifacts. Because the copied parts come from the same image, its noise component, color palette, dynamic range, and most other important properties will be compatible with the rest of the image and thus will not be detectable using methods that look for incompatibilities in statistical measures in different parts of the image. To make the forgery even harder to detect, one can use the feathered crop or the retouch tool to further mask any traces of the copied-and-moved segments. Examples of the Copy-Move forgery are given in Figures 1.1–1.4. Figure 1.1 is an obvious forgery that was created solely for testing purposes. In Figure 1.2 and 1.3, you can see a less obvious forgery in which a truck was covered with a portion of the foliage left of the truck (compare the forged image with its original). It is still not too difficult to identify the forged area visually because the original and copied parts of the foliage bear a suspicious similarity. Figure 1.4 shows another Copy-Move forgery that is much harder to identify visually. This image has been sent to the authors by a third party who did not disclose the nature or extent of the forgery. We used this image as a real-life test for evaluating our detection tools. A visual inspection of the image did not reveal the presence of anything suspicious.

Figure 1.1: Test image “Hats”

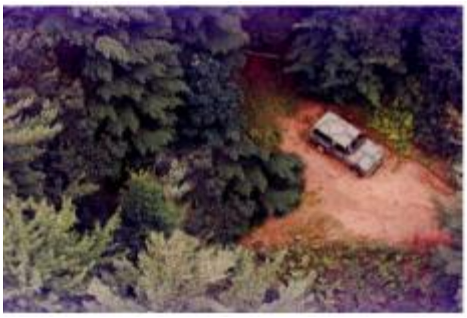


Figure 1.2: Forged image

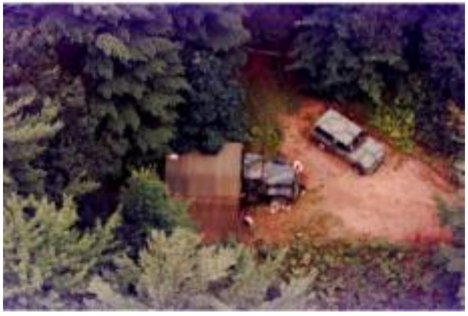


Figure 1.3: Original image

Figure 1.4: Test image “Golf” with an unknown original.

RELATED WORK

During the last few decades, various digital forensic techniques are proposed in the literature that has been made to forgery detection in videos by some manipulation operations. Ying Zhang et al. [6] proposed a two stage deep learning approach to learn features in order to detect tampered images in different image formats. For the first stage, they have utilized a Stacked Auto encoder model to learn the complex feature for each individual patch. For the second stage, they have integrated the contextual information of each patch so that the detection can be conducted more accurately. In Zhen Zhang et al. [7] authors have proposed an approach based on image quality metrics (IQMs) and moment features. They have analyzed the model creation and the extraction of features in digital image. In addition, they have compared these approaches and analyze the future works of digital image forensics. In [8] authors have proposed an improved scale invariant feature transform (SIFT)-based copy-move detection method, which combines broad first search neighbors (BFSN) clustering and color filter array (CFA) features. BFSN clustering algorithm is applied to detect multiple copied areas in tampered images. Angelo Ferreira etal. [9] have proposed different techniques that exploit the multi-directionality of the data to generate the final outcome detection map in a machine learning decision-making fashion and they have comparing the proposed techniques with a gamut of copy{move detection approaches and other fusion methodologies in the literature. In [10], a novel copy-move forgery detection method based on convolution neural network is proposed. The proposed method uses existing trained model from large database as Image Net, and then adjusts slightly the net structureusing small training samples. In [11] authors proposed a new method based on patch matching for detection and localization of video copy-move forgeries, they have used GRIP Dataset for their experiment purpose. The authors claim that the proposed method is out performed with 0.65 F-Measure. In [12] Jia, Shan, et al. have proposed Coarse-to-fine detection strategy based on optical flow for detect the copy move forgery in videos and they have evaluated their proposed methods in three publically available datasets namely, SULFA, DERF and VTL. In [13] authors have proposed Copy-Move forgery detection using Scale Invariant Features Transform (SIFT) features. In this research work authors adopt the SULFA dataset and their private dataset to evaluate the proposed methods. L. Zheng et.al. [14] proposed a new methodology based on Block-wise Brightness Variance Descriptor (BBVD) which is capable of fast detecting video inter-frame forgery. The proposed algorithm has been tested on a database consisting of 240 original and forged videos. In [15] author detect face tempering in videos based on mesoscopic properties of images they have proposed a deep neural network MesoNet. They have evaluated their methods on face2face dataset and deepfakes generated images, and got the accuracy 98% and 95% respectively, the rest of the papers are shown in Table 2.1.

Authors/Year	**Method Used**	**Used Dataset**	**Performance Measure**
L.D. Amiano et.al./2017 [11]	Patch based features, nearest- neighbourfield (NNF) and adhocvideo-oriented version of Patch Match	GRIP, REWIND	F-Measure: 0.65
S. Jia et.al./ 2018 [12]	Coarse-to-ne detection strategy basedon Optical Flow	SULFA, DERF, VTL	Precision: 0.96; Recall:0.92
R.C. Pandey et.al./ 2014 [13]	Scale Invariant Features Transform (SIFT), Correlation	SULFA and own Private	Accuracy: 98.98%

Harpreet Kaur et.al.[16]	Deep convolutional neural network tofind the spatial and temporal Correlation between authentic and forged	REWIND and GRIP	Accuracy: 98%
W. Wang et.al./ 2007[17]	De-interlacing algorithm	Own Private Dataset	Accuracy: 99.44%
C.C. Hsu et.al./ 2008[14]	Block level correlation, Gaussian Mixture Model, Bayesian classifier	Own Private Dataset	Recall: 63.29%; Precision:95.09%

G. Singh et.al./ 2018[18]	Mean based features and ThresholdingScheme	SULFA and Other collected
		videos from internet
			Precision: 1;
			Recall: 0.990;
			Accuracy: 0.995

H. Ravi et.al. 2014[19]	Modified Huber Markov Random Field(HMRF)	Open visual test media, Derf'scollection and YUV videos	Accuracy: 98.80%

III. PROPOSED DETECTION OF FAKE IMAGE BY BLOCK MATCHING

In this section we have discussed about identifying those segments in the image that match exactly. Even though the applicability of this tool is limited, it may still be useful for forensic analysis. In the beginning, the user specifies the minimal size of the segment that should be considered for match. Let us suppose that this segment is a square with BB pixels. The square is slid by one pixel along the image from the upper left corner right and down to the lower right corner. For each position of the B×B block, the pixel values from the block are extracted by columns into a row of a two-dimensional array A with B^2 columns and *(M-B + 1) (N-B + 1)* rows. Each row corresponds to one position of the sliding block. Two identical rows in the matrix A correspond to two identical *BB* blocks. To identify the identical rows, the rows of the matrix A are lexicographically ordered (as B×B integer tuples). This can be done in $M\ N\ log_2(M\ N)$ steps. The matching rows are easily searched by going through all MN rows of the ordered matrix A and looking for two consecutive rows that are identical. The matching blocks found in the BMP image of Jeep (Figure 1.2) for B=8 are shown in Figure 3.1. The blocks form an irregular pattern that closely matches the copied-and-moved foliage. The fact that the blocks form several disconnected pieces instead of one connected segment indicates that the person who did the forgery has probably used a retouch tool on the pasted segment to cover the traces of the forgery.

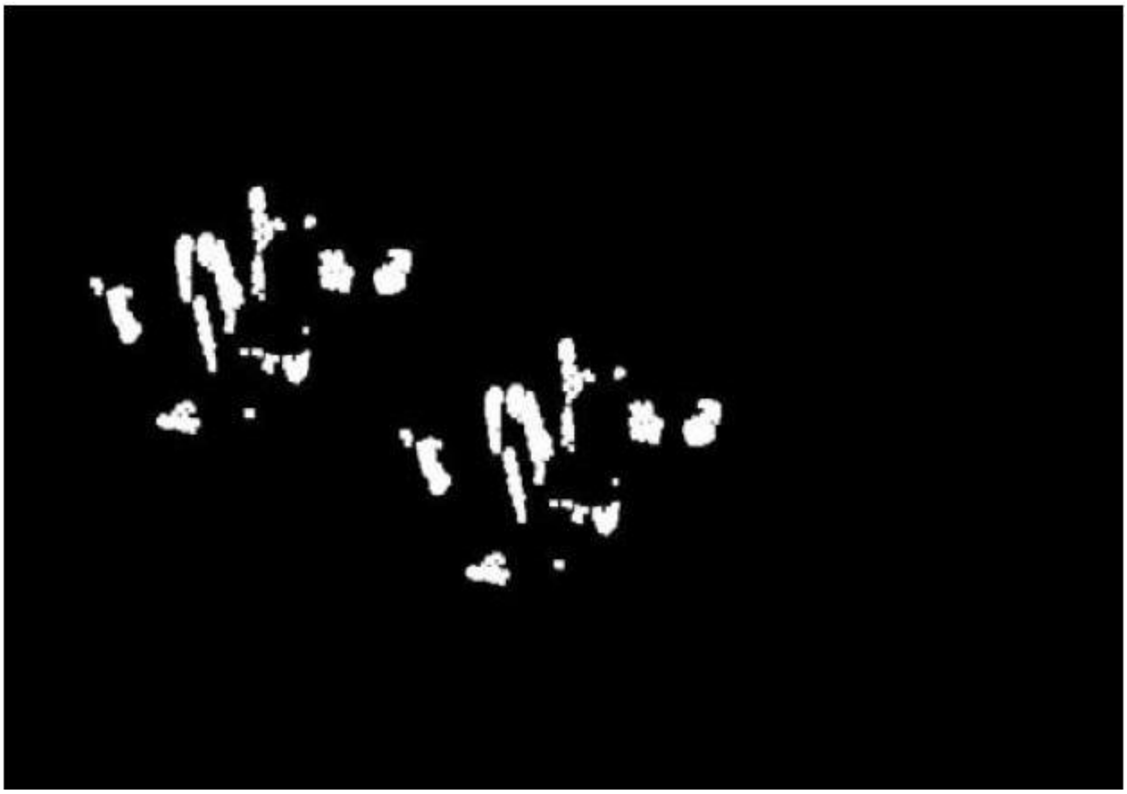


Figure 3.1: Results of the Block Match Copy-Detection forgery algorithm (the exact match modewith block size B=4).

IV. EXPERIMENTAL RESULTS AND ANALYSIS

The experimental results of proposed technique are presented in this chapter. Adobe Photoshop is used to forge the images and all the experiments are performed on a plate form with Intel 1.70GHz Core i3 processor and Python 3.7. The performance of the proposed technique is evaluated on two datasets. We used grayscale images with the size 128×128 pixels from the DVMM Columbia University dataset. The second dataset is collected from the Internet, containing the images of sizes 256×256 and 512×512 pixels. In the experimentation, we set the parameter values for B (current block) of size w = 16×16;Nn (number of rows to compare) = 20;Nt (block distance threshold) = 40,Nc (number of principal components) 10 and dt (similarity distance between vectors) = 0.0015, respectively. The experimentation details are presented in the following sections In Figure 4.1 we have shown the visual results of our proposed methodology, in Figure 4.1 first column indicates the fake images, second column indicate the Ground Truth of forged region and the last column shows the output of the proposed methodology

Performance Evaluation

Practically, the most significant property of a detection technique is its capability to discriminate forged and authentic images. In addition to this, the power of locating the forged area correctly is also very important which gives a strong evidence to expose digital forgeries. Thus, the performance of our algorithms evaluated at two levels: at image level, where we are concerned about the fact that the detected image is truly a forged image, and at pixel level, where we evaluate how accurately the forged areas can be located. To show the accuracy of the proposed technique at image level, the computation of precision \p" indicates the probability that an identified forgery is indeed a forgery; and recall \r" denotes the probability that actually a forged image is detected [20]:

$$P = \frac{T_P}{T_P+F_P}$$
$$T = \frac{T_P}{T_P+F_n}$$

Where T_P represents the total number of correctly detected forged images, F_P represents the total number of authentic images mistakenly detected as forged, and F_n represents the total number of forged images incorrectly missed.

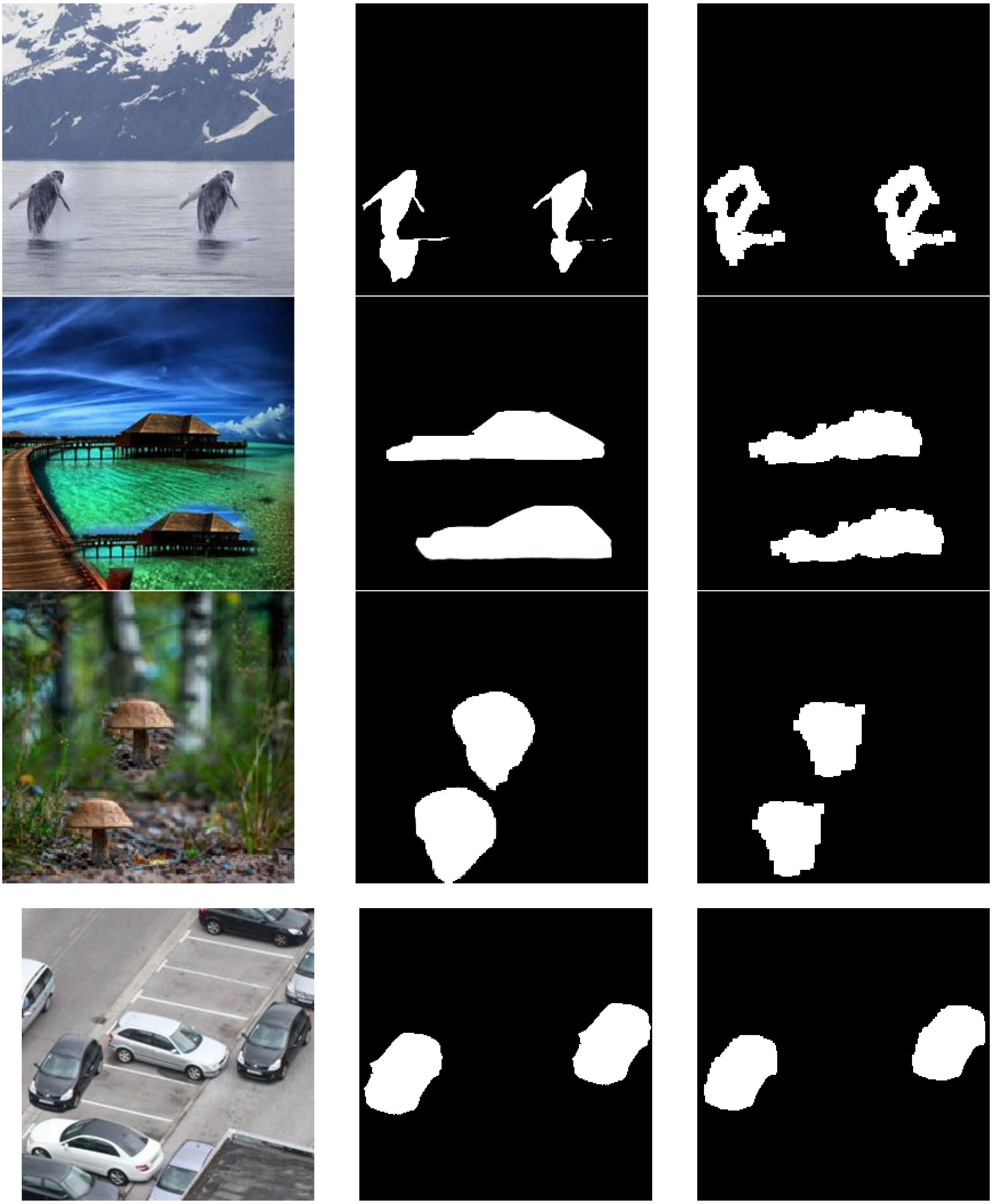
To show the accuracy at pixel level the true positive rate (TPR) and the false positive rate (FPR) are calculated as follows:

$$TPR = \frac{|\phi_s \cap \bar{\phi_s}| + |\phi_f \cap \bar{\phi_f}|}{|\phi_s| + |\phi_f|}, \quad TPR = \frac{|\bar{\phi_s} - \phi_s| + |\bar{\phi_f} - \phi_f|}{|\phi_s| + |\phi_f|}$$

Where Ø$_S$represents the pixels of original area, Ø $_f$ he pixels as the forged area, Ø$_S$the pixels as the detected original area, and Ø $_f$ the pixels as the detected forged area. Hence, the TPR shows the performance of technique by correctly identifying the pixels of the copy-moved areas in the forged image, while FPR reflects the pixels which are not contained in forged region butmistakenly included by the implemented technique. Therefore, both the above parameters point out how accurately the proposed technique can locate duplicated areas. The more the TPR is close to 1 and FPR is close to 0, the more precise the technique would be. We have shown the Similarity measurements of our proposed methods in Table 4.1.

Image Name	**Accuracy**	**FPR**
Figure 4.1 (Row 1)	0.9776	0.2372
Figure 4.1 (Row 2)	0.9556	0.2197
Figure 4.1 (Row 3)	0.9586	0.3814
Figure 4.1 (Row 4)	0.9876	0.3092
Figure 4.1 (Row 5)	0.9744	0.5112

Table4.1:Similarity Measurement Table of our proposed method



(a)InputImage (b) Ground Truth (c) Output

Figure4.1: Output ofOurProposedMethodology

CONCLUSION

In this paper, we focused on finding the ways through which we can assure the detection of copy-move forgery in digital images. The main consideration of this project work is to reduce the dimension of the feature length and find the forged objects in the suspected image. If once we can identify any forged region in an image then we can consider the image is fake image otherwise the image is Authentic. Therefore, we have applied block match methods which considers the identical objects found in the forged image. If we can detect any forged region in an image then we can consider the image as fake. In this work, we investigate the problem of detecting the copy-move forgery and describe an efficient and reliable detection method. The method may successfully detect the forged part even when the copied area is enhanced/retouched to merge it with the background and when the forged image is saved in a lossy format, such as JPEG. Furthermore, this technique does not require any prior information embedded into the image and works in the absence of digital signature or digital watermark. From the results, a conclusion can be drawn which is that the proposed technique is out performed.

REFERENCES

1. NealKrawetzandHackerFactorSolutions.“Apicture’sworth”.In:*HackerFactorSolutions*6.2 (2007), p. 2.
2. Shiguo Lian and Yan Zhang. “Multimedia forensics for detecting forgeries”. In: *Handbook of Information and Communication Security*. Springer, 2010, pp. 809–828.
3. Yuenan Li. “Image copy-move forgery detection based on polar cosine transform and ap- proximate nearest neighbor searching”. In: *Forensic science international*224.1-3 (2013),pp.59–67.
4. HanyFarid.“Digitaldoctoring:howtotelltherealfromthefake”.In:*Significance*3.4 (2006), pp. 162–166.
5. BinBZhu,MitchellDSwanson,andAhmedHTewfik.“Whenseeingisn’tbelieving[multimediaauthenticationtechnologies]”.In:*IEEESignalProcessingMagazine*21.2(2004),pp.40–49.
6. YingZhangetal.“ImageRegionForgeryDetection:ADeepLearningApproach.”In:*SG-CRC*2016(2016),pp.1–11.
7. ZhenZhang,JiquanKang,andYuanRen.“Aneffectivealgorithmofimagesplicingdetection”. In: *2008 international conference on computer science and software engineering*. Vol. 1. IEEE. 2008, pp. 1035–1039.
8. LuLiuetal.“ImprovedSIFT-basedcopy-movedetectionusingBFSNclusteringandCFA features”. In: *2014 Tenth International Conference on Intelligent Information Hiding and MultimediaSignalProcessing*.IEEE.2014,pp.626–629.
9. Anselmo Ferreira et al. “Behavior knowledge space-based fusion for copy–move forgery detection”. In: *IEEETransactions on Image Processing*25.10 (2016), pp. 4729–4742.
10. JunlinOuyang,YizhiLiu,andMiaoLiao.“Copy-moveforgerydetectionbasedondeep learning”. In: *2017 10th international congress on image and signal processing, biomedical engineering and informatics (CISP-BMEI)*. IEEE. 2017, pp. 1–5.
11. Luca D’Amiano et al. “A patchmatch-based dense-field algorithm for video copy–move detection and localization”. In:*IEEETransactions on Circuits and Systems for Video Technology*29.3 (2018), pp. 669–682.
12. Shan Jia et al. “Coarse-to-fine copy-move forgery detection for video forensics”. In: *IEEE Access*6 (2018), pp. 25323–25335.
13. Ramesh Chand Pandey, Sanjay Kumar Singh, and KK Shukla. “Passive copy-move forgery detection in videos”. In: *2014 International conference on computer and communication technology (ICCCT)*. IEEE. 2014, pp. 301–306.

14. Lu Zheng, Tanfeng Sun, and Yun-Qing Shi. “Inter-frame video forgery detection based on block-wise brightness variance descriptor”. In: *International Workshop on Digital Water- marking*. Springer. 2014, pp. 18–30.
15. Darius Afchar et al. “Mesonet: a compact facial video forgery detection network”. In: *2018 IEEE International Workshop on Information Forensics and Security (WIFS)*.IEEE.2018,pp.1–7.
16. Harpreet Kaur and Neeru Jindal. “Deep convolutional neural network for graphics forgery detection in video”. In: *Wireless Personal Communications*(2020), pp. 1–19.
17. Weihong Wang and Hany Farid. “Exposing digital forgeries in interlaced and deinterlaced video”. In: *IEEE Transactions on Information Forensics and Security*2.3 (2007), pp. 438– 449.
18. Gurvinder Singh and Kulbir Singh. “Video frame and region duplication forgery detection based on correlation coefficient and coefficient of variation”. In:*Multimedia Tools and Applications*78.9 (2019), pp. 11527–11562.
19. Hareesh Ravi et al. “Compression noise based video forgery detection”. In:*2014 IEEE International Conference on Image Processing(ICIP)*.IEEE.2014,pp.5352–5356.
20. Vincent Christlein et al. “An evaluation of popular copy-move forgery detection approaches”. In: *IEEETransactionsoninformationforensicsandsecurity*7.6 (2012), pp. 1841–1854.