Journal of Innovative Image Processing (ISSN: 2582-4252)
www.irojournals.com/iroiip/

Identification of Image Forgeries using Machine Learning - A Review

Nagarathna C R[1], Jayasri A[2], Chandana S[3], Amrutha A[4]

[1]Assistant Professor, AI and ML department, BNM Institute of Technology, VTU, Bangalore, India

[2,3,4]Students, AI and ML, BNM Institute of Technology, VTU, Bangalore, India

E-mail: [1]nagarathna.binu@gmail.com, [2]jayasria45@gmail.com, [3]shivakumarchandana2002@gmail.com, [4]amrutha.amar21@gmail.com

Abstract

Forgery in images is the manipulation of digital images using techniques like copy-move, splicing, removal of parts of image. Image forgery detection is a crucial task in digital image processing field. The growth and use of digital images in various industries such as forensics, journalism and scientific research has increased the number of manipulated and forged images. New and advanced editing tools and techniques are capable of easily manipulating images without leaving traces, which can lead to negative impact for individuals and society. Therefore, the need for reliable and efficient forgery detection techniques has become more important than ever. They are required to protect the authenticity of images and avoid the spread of fabricated and fake news. In this study the overview of the existing methods for identifying forgeries in images, and the summary of the issues found in these methods are discussed.

Keywords: Forgery Detection, Digital Image Forensics, CNN, Splicing Detection

1. Introduction

In the digital age, image fabrication has increased as more people and companies produce fake images for various uses. These forgeries could be used for propaganda, deceit, or other nefarious motives. Therefore, it is becoming more and more important to have the right tools and techniques to detect and stop image fraud. One of the most promising tactics is to

DOI: https://doi.org/10.36548/jiip.2023.3.007
Received: 14.09.2023, received in revised form: 02.10.2023, accepted: 16.10.2023, published: 01.11.2023

apply machine learning [1]. Image forgery, which is simple to perform with software or editing tools, has increased due to the growing usage of digital images in a variety of industries, including forensic investigation, surveillance systems, intelligence systems, criminal investigation, medical imaging, and legal services among many others.

Image Forging can have negative consequences, including the deterioration of public confidence in visual representations and the use of altered images as supporting documentation in court cases. [2]. Several image forgery techniques like splicing, copy-move and removal are used. Advanced image manipulation techniques are developing quickly, making it possible to change images without leaving any visible traces. Exceptional forgeries are so excellent that they escape detection from the unaided eye and do not show any signs of manipulation to conventional image tamper detection tools. [3]. As a result, several detection techniques have been established in image forensics due to the significance and applicability of digital image forensics. This survey aims to include a comprehensive analysis of existing methods, from conventional to current progress including the Deep learning (DL), and a review of recent developments in the field [4-5]. There are many technical challenges associated with detecting image forgery, such as the need for robust and accurate image features, the ability to distinguish between different types of image manipulations, and efficient algorithms which can process large volumes of digital images. These challenges require the development of innovative and complex techniques in the computer vision and the image processing fields. This study presents the review the few existing techniques to detect the image forgery. The outline of the study is presented with literature survey in section 2, the gaps identified in existing methods are listed in section 3, the section 4 presents the summarization of existing research methodology and section 5 concludes with the contribution of this study for future researchers.

2. Related Work

The existing research often focuses on the detection of a single type of image forgery detection and hence fails to detect other types of forgeries. The research aims to find an algorithm that can work well with all forms of forgeries. Splicing, removal, and copy move fraud which are the three most used methods for detecting image forgeries. There are several detection techniques which are applied for “copy move forgery detection” (CMFD). The author “C. Wang, Z. Zhang, Q. Li and X. Zhou”[6] have used SURF in combination with PCET

algorithms whereas “Jixiang Yang, Zhiyao Liang, Yanfen Gan, Junliu Zhong”[7] have proposed a novel method using two-stage filtering which uses SURF along with SIFT. The author “Goel, N, Kaur, S, Bala, R”[8] have used Convolutional Neural Networks(CNN) Architecture and achieved an accuracy of 96-97% . The authors “Nitish Kumar & Toshanlal Meenpal”[9] have used SIFT and KAZE algorithms to extract the features while Koul, S., Kumar, M., Khurana, S.S. et al.[12] and. “Paul, S., Pal, A.K”[10] have used overlapping block based Discrete Cosine Transforms (DCT). The methodologies used for image splicing detection often do not overlap with the methodologies used for CMFD. Authors “Muhammad Hameed Siddiqi, Khurshed Asghar, Umar Draz, Amjad Ali, Madallah Alruwaili, Yousef Alhwaiti, Saad Alanazi, M. M. Kamruzzaman, and Usman Habib”[11] used “Discrete Wavelet Transform’ (DWT) and “Edge Weighted Local Binary Patterns (EW-LBP)”; “Bo Liu, Chi-Man Pun”[4] used “Deep fusion network”; Patrick Niyishaka and Chakravarthy Bhagvati[12] proposed a methodology using Local Binary Pattern (LBP) and “Bin Xiao, Yang Wei, Xiuli Bi, Weisheng Li, Jianfeng Ma”[3] used Cascaded convolutional neural network (C2RNet) for splicing detection. Several researchers do not explicitly mention the image forgery technique detected by their methodology. The research on identification of image forgery applying the methods of detection is limited.

The state art of literature survey shows that existing methods of image forgery detection uses pre-processing techniques such as image normalization, image compression [5], image resizing [14], super pixel segmentation [6] and conversion to grayscale images [2][13]. The features are extracted using machine learning methods such as CNN [5][12][10][4], LCA [17], SURF[6][11], etc. The images are then classified using classification techniques like binary classification [8], SVM [5][16][13][1] and ELM [14]. The performance of the methods are evaluated using various datasets like Dresden [8][17], FAU [5][6], CASIA[7][3][18][13][1] and MICC-F2000[12][10]. An average these conventional methods have achieved accuracy of 98.95%[13]. The detailed survey of the existing methods are tabulated in table 1.

Table 1. Summarization of Existing Methods for Detection of Forgery in Images.

Authors	**Dataset**	**Pre-processing techniques**	**Feature extraction technique**	**Classification techniques**	**Accuracy**
Ahmed Ghoneim, Ghulam Muhammad et al	CASIA 1, CASIA 2	Images noise map	"Multi-resolution regression filter"	Classifier integrated with extreme learning and SVM	97.4% ,98.2%
Saif alZahir et al	CoMoFoD	Conversion to grayscale image	Steerable pyramid decomposition technique	Copulas ensemble	95.90%
Bin Xiaoa et al	COLUMB, CASIA, FORENSICS	Image normalization	Cascaded convolutional neural network (C2RNet)	Adaptive clustering that groups the extracted features into clusters	8% higher than R-CNN
Bo Liu et al	Splicing forged pictures	Image normalization	Deep fusion network that combines the outputs of multiple CNN models	Neural network	97%
Boubacar Diallo et al	CMI(Dresden dataset)	Compression using JPEG format	CNN	Support vector machine (SVM)	90%
Chengyou Wang et al	GRIP, FAU and SBU-CM16	Superpixel segmentation	SURF and PCET	"Random sample consensus (RANSAC) algorithm and filtering scheme"	96%
Haipeng Chen et al	NIST16, COVER, AGE, CASIA	Image normalization	Rotating residual units	Semantic reinforcement network that combines the outputs of multiple CNN models	98.90%

Francesco Marra et al	MFC2019, MFC2018, NC2017, FAU / DSO-1,Korus, /Dresden	Image normalization	Patch-wise processing	Binary classification	85.10%
Falko Matern et al	ALOI, COCO, SDO-1, IEEE IFS-TC Challenge, OpenImages Splices (OIS)	Pre-segmentation	2-D lighting environment	Salient objects comparison	97%
Nidhi Goel et al	MICC F-2000	Image transforms, color space transformation, and dimensionality reduction	CNN architecture	Dual branch CNN that combines the outputs of two CNN models	96-97%
Jixiang Yang et al	IMD, CoMoFoD, CMHD	Block based algorithms	Enhanced SURF and SIFT were used.	2-stage filters grid and cluster filters along with the Delaunay triangulation algorithm	82%
Saboor Koul et al	MICC-F2000	Image normalization	CNN architecture	Fully connected layer and a softmax activation function.	97.52%
Muhammad Hameed Siddiqi et al	DVMM, CASIA v1.0 and CASIA v2.0, Columbia	Transformation into YCbCr color space	DWT and EWLBP (“Discrete Wavelet Transform combined with Edge Weighted Local Binary Patterns”)	Support Vector Machine (SVM)	98.95%

N. Krishnaraj et al	Benchmark datasets (MNIST, CIFAR-10)	Image resizing and image normalization	DenseNet	Extreme learning machine (ELM) classifier	95.42% and 96.94%
Nitish Kumar et al	CoMoFoD and MICC-F220	Region proposal approach	"Scale-Invariant Feature Transform" (SIFT) with "KAZE"algorithms	Feature descriptor matching	97.90%
Njood Mohammed AlShariah et al	Images extracted from Instagram application	Image resizing	Color histogram, edge detection, and texture analysis	Support Vector Machine (SVM) classifier	97%
Owen Mayer et al m	Dresden Image Database	Camera response removal	Lateral Chromatic Aberration	Hypotheses testing	84%
Patrick Niyishaka et al	CASIA v2.0, Digital Image Forensics	Luminance and Chrominance are extracted using Illumination-Reflectance model	Local Binary Pattern (LBP)	Support Vector Machine, Linear Discriminant Analysis, Logistic Regression, K-Nearest Neighbours, Decision Tree, Naive Bayes	93.79%
Srilekha Paul et al	Self generated dataset and BMP images from public dataset	Gaussian image pyramid	Discrete cosine transform (DCT)	Reduced space search	96.14%

Through the research conducted above some of the research questions regarding image forgery detection are:

- Can image forgery detection be improved by combining multiple detection algorithms or techniques?

As mentioned in introduction section the image forgery can be identified by mixture of techniques like splicing, resampling, cloning, region removal, and other techniques are used to create realistic image forgeries. While copy-move detection algorithms are excellent at spotting cloning and region removal, method to detect resamples are efficient in spotting splicing and resampling. These detection techniques can be combined to the detection of image modification is improved overall by the use of complimentary techniques [21].

- What are the most effective algorithms and techniques for detection of image forgeries?

Forgeries in images can be identified using techniques like active approach and passive approach techniques. According to this viewpoint, digital pictures must undergo preprocessing such as adding a watermark or creating a signature on the image, which limits their practical application[3]. Without requiring any explicit additional activities for the purpose of authentication, the client's identity is verified and checked in this. It can also act blindly when detecting something. No prior knowledge of the image is required for this procedure.We did not use any active techniques, such as watermarking or digital signatures, while evaluating the originality and validity of photos. Instead, we employed passive detection. These are predicated on the presumptions that there are no indications of forged regions on digital images, and this may alter the underlying image regularity of our field of view.image that kickstarts the production of new artifacts in many different sorts the anomalies[22]

- What are the key characteristics and features of manipulated images that can be used for forgery detection?

The image forgery can be found using a variety of key factors in passive forgery detection. Figure 1 illustrates some of the key factors that are important in forgery detection. Although it is challenging to identify forgeries due to JPEG compression, this technique analyses every pixel of the provided images in order to do so, in addition to camera-based

criteria..- Several types of lighting can be employed to take pictures that look natural. These physical parameters are used to detect image fraud because the illumination of a counterfeit zone during splicing procedures may differ from the original lighting [23].

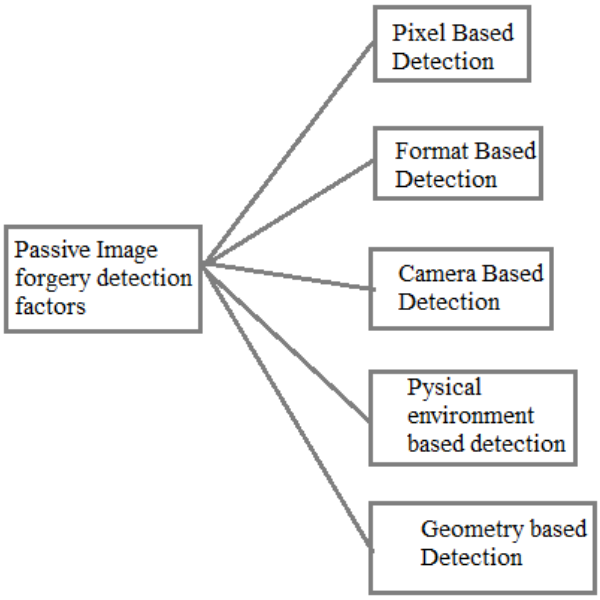

Figure 1. Key Factors used in Detection of Image Forgery

- What are the limitations and challenges of current image forensics tools in detecting image forgery?

The challenges of image forgery are

-Data Provenance: In applications like science, medical, financial transactions, government legal prosecutions, and many more everyday scenarios, where the information is valuable and reliable, the data provenance is essential for the protection of rights and may be a regulatory necessity.

-Digital information migration: with technology constantly evolving, it is more difficult to maintain the integrity of digital documents as they are transferred across organizations and over the internet while maintaining the capacity to retrieve and display integrated digital materials.

- Ethical, legal and institutional issues: Widespread ambiguities over the management and preservation of intellectual assets (such as text and other document-like objects, photos, film, software, and multimedia objects) present additional risks and difficulties[24].

3. Gaps Identified

The importance of detecting image forgeries has increased recently, and numerous studies and research projects are in progress. Many new techniques are discovered for image forgery detection. But there are few issues that are yet to be resolved. The gaps that were identified through the literature survey is listed below.

- Many image forgery techniques used small datasets that do not reflect the variety of real-world images. This limits the ability to generalize new images.
- Some image forgery detection models are not transparent in their decision-making process, making it difficult to understand how they are making their detections.
- It is challenging to determine whether the images are manipulated using the current approaches because of the sampling or interpolation of large-scale reduction or expansion in the image regions.
- There is no consensus on the metrics used in evaluation and benchmark datasets for image forgery exposure, which makes it challenging to compare the performance of different detection models.
- Most of the research in image forgery detection has focused on static images, with limited attention given to detecting forgeries in videos.
- Performance of deep learning models should be optimized for various types of image forgeries.

4. Existing Research Methodology

Many methods have been implemented for identifying the image forgery. Convolution Neural Networks (CNN) is the most widely used methodology for image forgery detection.

Few methods use deep learning along with CNN for more accurate results. These methods often focus on “Copy-move forgery detection and splicing detection”.

The detailed algorithm for image forgery detection is given in algorithm 1.

Algorithm 1: Image Forgery detection

Input : RGB images

Output: Forged or un forged images

Start

Step 1: Load the RGB images

Step 2: Preprocess the input dataset to improve the performance of a system.

Step 3: Extract the features from preprocessed dataset by using various machine learning techniques.

Step 4: Design and develop a classification model to classify the image as forged or unforged image.

End

The existing method for forgery detection is shown in figure 1.

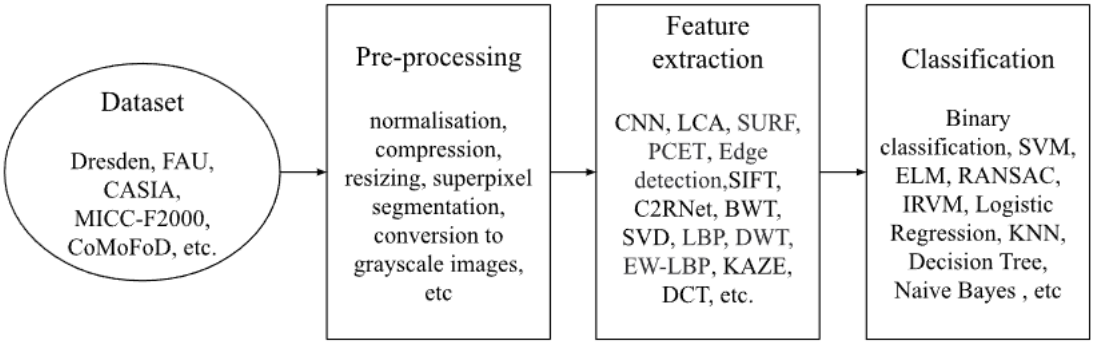

Figure 2. Existing Method

The figure 2 shows that , the preprocessing techniques like normalization, compression resizing are applied on dataset received from various organization in order to improve the performance of a system. Further the feature extraction techniques are applied on the preprocessed images in order to extract the relevant features from the images to do the classification. Finally, the extracted features are considered to do binary classification by applying the machine learning algorithm.

5. Discussion

This study presents the, review of the existing methods to identify the image forgery. The image forgery is essential for uses like forensic investigations and social media monitoring. There are many ways to fake an image, including splicing, retouching, and copy-move forgeries. Cutting, pasting, and reassembling images are all part of the copy-move forgery process. The act of splicing involves combining several images to create a fresh one. Retouching is the procedure used to alter the appearance of an image. Machine learning may be used to find the various types of evidence that each of these forgery types leaves behind. The identification of images using machine learning has many advantages over more traditional forgery detection systems. When compared to traditional methods, machine learning has many advantages, including speed, automation, accuracy, adaptability, scalability, and consistency. In this study the gaps in existing methods to identify the image forgery are figured out. As the literature shows existing machine learning technique shows an average performance to predict an image forgery by using limited dataset and static images. The future research has to concentrate on the performance-improving measurements by considering a large, dynamic dataset and transfer learning methods.

6. Conclusion

Image forgery is the manipulation of digital images. It is usually done with a malicious intent. Image forgery detection is a technique used to identify the various manipulation techniques that may be performed on an image and to check the authenticity of the image. Many existing methods of image forgery detection have achieved high accuracy in identifying forged images. But there are several disadvantages like un optimized performance and non-transparency of the model, limited training and testing as well as limited to only static images. Current image forgery detection techniques should be extended to different types of media, such as videos or live streams. The machine learning models should be trained to detect subtle or sophisticated image manipulation techniques. Existing image forgery detection algorithms should be adapted such that they can work in real-time applications. Focusing on developing algorithms that can effectively detect deep fake images and videos. These shortcomings can be overcome by developing new performance-oriented models and training the models with large datasets.

References

[1] Ghoneim, G. Muhammad, S. U. Amin and B. Gupta, "Medical Image Forgery Detection for Smart Healthcare," in *IEEE Communications Magazine*, vol. 56, no. 4, pp. 33-37, April 2018, doi: 10.1109/MCOM.2018.1700817.

[2] alZahir, Saif, and Radwa Hammad. "Image forgery detection using image similarity." Multimedia Tools and Applications 79, no. 39-40 (2020): 28643-28659.

[3] Niyishaka, Patrick, and Chakravarthy Bhagvati. "Image splicing detection technique based on Illumination-Reflectance model and LBP." Multimedia Tools and Applications 80 (2021): 2161-2175.

[4] Xiao, Bin, Yang Wei, Xiuli Bi, Weisheng Li, and Jianfeng Ma. "Image splicing forgery detection combining coarse to refined convolutional neural network and adaptive clustering." Information Sciences 511 (2020): 172-191.

[5] Liu, Bo, and Chi-Man Pun. "Exposing splicing forgery in realistic scenes using deep fusion network." Information Sciences 526 (2020): 133-150.

[6] Diallo, Boubacar, Thierry Urruty, Pascal Bourdon, and Christine Fernandez-Maloigne. "Robust forgery detection for compressed images using CNN supervision." Forensic Science International: Reports 2 (2020): 100112.

[7] Goel, Nidhi, Samarjeet Kaur, and Ruchika Bala. "Dual branch convolutional neural network for copy move forgery detection." IET Image Processing 15, no. 3 (2021): 656-665.

[8] Marra, Francesco, Diego Gragnaniello, Luisa Verdoliva, and Giovanni Poggi. "A full-image full-resolution end-to-end-trainable CNN framework for image forgery detection." IEEE Access 8 (2020): 133488-133502.

[9] Krishnaraj, N., B. Sivakumar, Ramya Kuppusamy, Yuvaraja Teekaraman, and Amruth Ramesh Thelkar. "Design of automated deep learning-based fusion model for copy-move image forgery detection." Computational Intelligence and Neuroscience 2022 (2022).

[10] Paul, Srilekha, and Arup Kumar Pal. "A fast copy-move image forgery detection approach on a reduced search space." Multimedia Tools and Applications (2023): 1-28.

[11] Matern, Falko, Christian Riess, and Marc Stamminger. "Gradient-based illumination description for image forgery detection." *IEEE Transactions on Information Forensics and Security* 15 (2019): 1303-1317.

[12] AlShariah, Njood Mohammed, A. Khader, and J. Saudagar. "Detecting fake images on social media using machine learning." International Journal of Advanced Computer Science and Applications 10, no. 12 (2019): 170-176.

[13] Rathore, Neeraj Kumar, Neelesh Kumar Jain, Prashant Kumar Shukla, UmaShankar Rawat, and Rachana Dubey. "Image forgery detection using singular value decomposition with some attacks." *National Academy Science Letters* 44 (2021): 331-338.

[14] Wang, Chengyou, Zhi Zhang, Qianwen Li, and Xiao Zhou. "An image copy-move forgery detection method based on SURF and PCET." IEEE Access 7 (2019): 170032-170047.

[15] Chen, Haipeng, Chaoqun Chang, Zenan Shi, and Yingda Lyu. "Hybrid features and semantic reinforcement network for image forgery detection." Multimedia Systems 28, no. 2 (2022): 363-374.

[16] Yang, Jixiang, Zhiyao Liang, Yanfen Gan, and Junliu Zhong. "A novel copy-move forgery detection algorithm via two-stage filtering." Digital Signal Processing 113 (2021): 103032.

[17] Koul, Saboor, Munish Kumar, Surinder Singh Khurana, Faisel Mushtaq, and Krishan Kumar. "An efficient approach for copy-move image forgery detection using convolution neural network." Multimedia Tools and Applications 81, no. 8 (2022): 11259-11277.

[18] Siddiqi, Muhammad Hameed, Khurshed Asghar, Umar Draz, Amjad Ali, Madallah Alruwaili, Yousef Alhwaiti, Saad Alanazi, and M. M. Kamruzzaman. "Image splicing-

based forgery detection using discrete wavelet transform and edge weighted local binary patterns." Security and Communication Networks 2021 (2021): 1-10.

[19] Kumar, Nitish, and Toshanlal Meenpal. "Salient keypoint-based copy–move image forgery detection." Australian Journal of Forensic Sciences 55, no. 3 (2023): 331-354.

[20] Mayer, Owen, and Matthew C. Stamm. "Accurate and efficient image forgery detection using lateral chromatic aberration." IEEE Transactions on information forensics and security 13, no. 7 (2018): 1762-1777.

[21] Mohammed, Tajuddin Manhar, Jason Bunk, Lakshmanan Nataraj, Jawadul H. Bappy, Arjuna Flenner, B. S. Manjunath, Shivkumar Chandrasekaran, Amit K. Roy-Chowdhury, and Lawrence Peterson. "Boosting image forgery detection using resampling features and copy-move analysis." arXiv preprint arXiv:1802.03154 (2018).

[22] Raja, A. "Active and Passive Detection of Image Forgery: A Review Analysis." IJERT-Proc 9, no. 5 (2021): 418-424.

[23] Sharma, Preeti, Manoj Kumar, and Hitesh Sharma. "Comprehensive analyses of image forgery detection methods from traditional to deep learning approaches: an evaluation." Multimedia Tools and Applications 82, no. 12 (2023): 18117-18150.

[24] Math, Shrishail, and R. C. Tripathi. "Digital forgeries: Problems and challenges." *International Journal of Computer Applications* 5.12 (2010): 9-12.