



Image forgery Detection

Prof. Swati Dronkar¹, Mauli Bhalotiya², Yashshree Bidkar³, Mayank Futane⁴, Umesh Amru⁵

Assistant Professor, Dept. of Computer Science and Engineering, Priyadarshini College of Engineering,
Nagpur, India¹

Dept. of Computer Science and Engineering, Priyadarshini College of Engineering Nagpur, India²⁻⁵

Abstract: Advancements in technology and globalization have made digital cameras widely accessible and affordable. As a result, people capture and collect numerous images using various camera sensors, often in soft copy for online documents and sharing on social media. Following the explosive use of social networking services, there has been an exponential increase in the volume of data of image. Moreover, the development of image processing software such as Adobe Photoshop has given rise to doctored images. These manipulated images can be used for malicious purposes, such as spreading false information or inciting violence. This image spam detection program allows users to detect even the smallest signs of fraud in images. With the rise in crime, image fraud has become a major problem that needs attention.

Moreover, the main goal of forgery detection in the digital age is to ensure immaculacy and validity. As research progresses, many deep learning methods are being implemented to identify fraud in images. Deep learning approaches have shown much better results for image manipulation compared to traditional methods. In this study, we have also aimed to determine the detection of image forgery using a deep learning approach. We propose a novel image forgery detection system based on Convolutional Neural Networks (CNNs) that can detect various types of image modifications, such as copy-move, splicing, and resampling. Our proposed system integrates Error Level Analysis (ELA) with deep learning techniques to provide an accuracy of 93% for detected images. Our proposed system even integrates Visual Geometry Group; it is a standard deep Convolutional Neural Network (CNN) architecture with multiple layers. After evaluating the proposed system on a database of real-world images and achieving a high detection VGG16's training accuracy of 93.21% and a training accuracy of 95.12% for VGG19. VGG16 is the first VGG network and VGG19 is the last hence we decided to use both of them as they give better accuracy than any other networks.

Keywords: digital cameras, doctored images, malicious purposes, authenticity, integrity, image forgery, deep learning, CNN, ELA, VGG16, VGG19

I. INTRODUCTION

As we know the real world is 3D but when we want or have to capture moments or instances for storing the as memories, we need to capture them in photos, audios and videos. Photos and videos capture everything in 2D. The photos are called Digital images which are represented numerically in binary form. The Government is moving towards the digital processing of administrative procedures to save time and money and optimize the processing. Hence digital images are also the softcopy of our personal documents such as Aadhar card, PAN card, etc.

These documents are confidential and cannot be shared with anyone except Government authorities and some authorized third parties. In this era there is an exponential increase in the rate of digital crimes or cybercrimes using the false documents. This means that the original documents of an individual are being tampered or forged by the criminals to use their identity for malicious purposes. We are building this project to detect this type of forgeries. Unfortunately, there are various tools available across the web to make modifications in images such as Adobe photoshop, Snapseed, Remini, etc. and this basically means tampering the images for various purposes such as frauds.

Digital modification of a images to alter meaningful and useful information is defined as image forgery; and detection of image forgery means to identifying the forgeries in the image. There are various types forgeries that can be done on any image such as Copy-Move, Splicing, Resampling, etc. In this project we trying to detect Copy-Move and Splicing type of forgeries. In this paper we have discussed about the previous work of the researchers for the forgery detection. We will see the comparison between the existing methods and apply the best alternative in our project. We are using a passive approach of detection in this project. As mentioned above we will be using three training models i.e., the Convolution neural networks VGG16 and VGG19 alongside Error level analysis.



II. APPROACHES TO IMAGE FORGERY DETECTION

There has been an exponential growth in the communication technology over the time. Digital Image forgery has become the area of interest for many scientists or researchers to secure the administrative, business and personal communications. Hence the researchers are now aware of various methods to detect doctored images. Image forgery can be done in many ways, to detect these forgeries there two approaches as indicated in the figure below.

A. Active Approach

In Active approach, all the knowledge about the image is known and is crucial for authentication and integrity. Active methods are used for hiding the data, as some key information which is added during generating the image. By checking the key information, the image is authenticated or validated. This can be done in two possible ways namely, Digital Signature and Digital watermarking.

The Digital watermarks are embedded inside the image in the processing's stage itself, where additional information as Digital signature is added to an image during the capturing end.

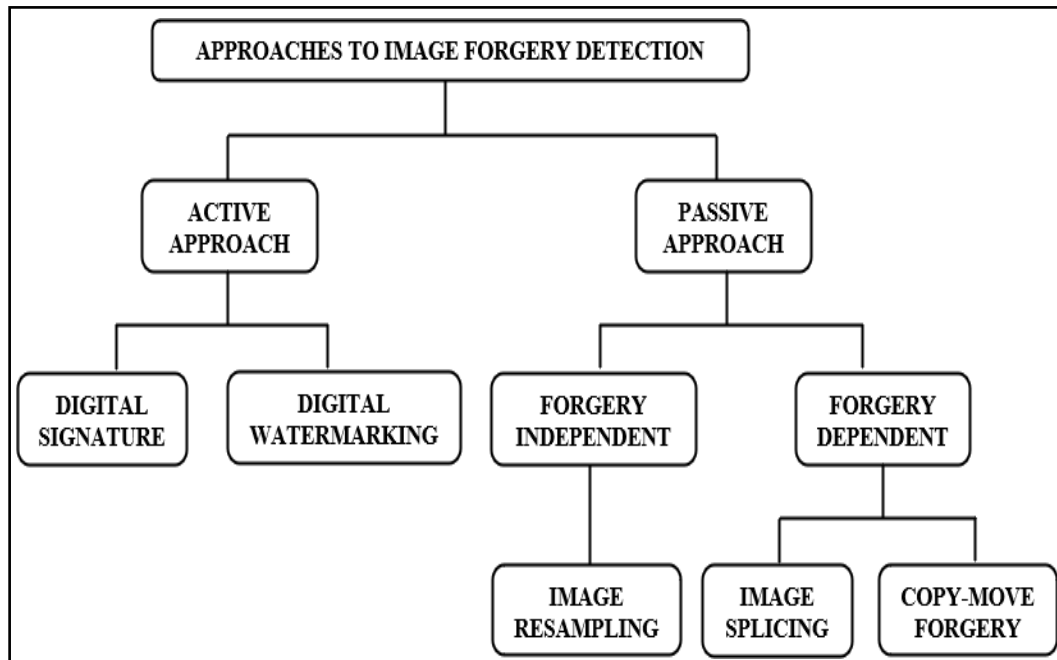


Fig. 1. Approaches to Image Forgery Detection

B. Passive Approach

Passive approach is also called as Blind approach. When authenticating using passive approach we only need the image not any additional information. The concept behind these methods is that even if there is no obvious evidence of manipulation, the underlying data will be altered.

Hence passive methods are more suitable to detect doctored digital image. In passive method we do not need any additionally embedded watermark or original image to check for forgery. The two types of Passive approach are as follows:

- Forgery Independent:

These approaches look for forgeries that aren't based on the sort of forgery, but rather on artefact traces left behind during the resampling process or owing to lighting discrepancies.

1. Image Resampling: It is also known as Image Retouching. It is done to enhance the image, improve the image quality to grab attention.

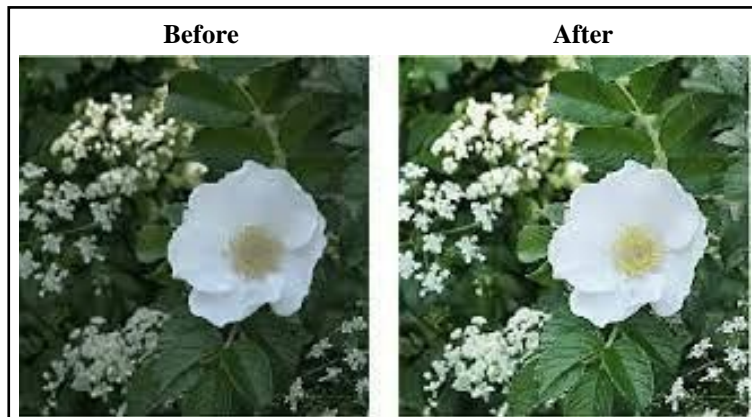


Fig. 2. Image Resampling Forgery Example

• Forgery Dependent:

These approaches look for forgeries that are reliant to forgery types that are performed on images.

1. Copy-Move Forgery: It is also called cloning. In this type any part of image is copied and pasted somewhere inside the image only. e. Doing so can be used to conceal critical information.
2. Image Splicing: In this type of forgery part one image is cut and pasted into another image. There can be two or more sources in the newly created image. If the splicing is done well, the boundaries between the spliced portions might be unnoticeable visually.



Fig. 3. Copy-Move Image Forgery Example

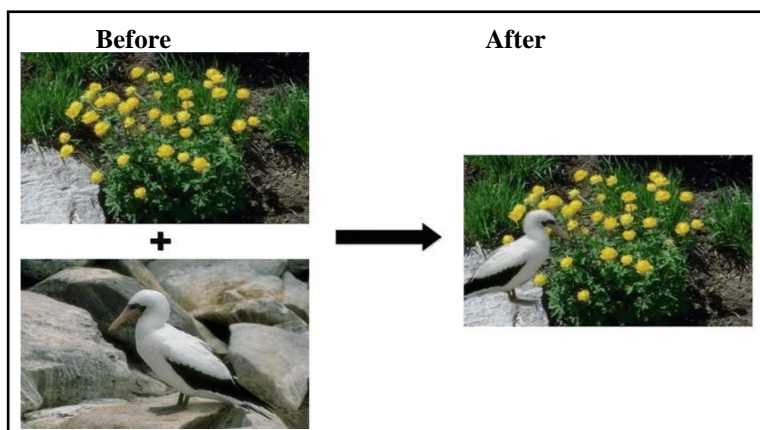


Fig. 4. Image Splicing Forgery Example



III. METHODOLOGY

As mentioned above we will be using deep convolutional neural network, Visual Geometry Group (VGG16 and VGG19) along with Error level analysis (ELA) in the methodology or working of the project. Let us understand each of them.

A. Error level Analysis

Error Level Analysis (ELA) identifies areas within an image which have different compression levels. In JPEG images, the entire picture should roughly be at the same level of compression. If any section in any picture is at a significantly different level of error, then it likely indicates a modification or tampering in the image. ELA is used to highlight differences in the compression rate of JPEG image. Regions with uniform colouring, like a solid blue sky or a white wall, will likely have a lower ELA result than high-contrast edges (darker colour). Look around the image and identify high contrast edges, low contrast edges, surfaces and textures. Compare those areas with the results of ELA. If there are significant differences, identify the digitally modified caution areas. JPEG restoration removes high frequencies and reduces differences between edges, textures and high-contrast surfaces. Very small JPEGs will appear too dark. Changing the image to a smaller size can emphasize sharp edges that will be clearer in ELA. Also, when you save a JPEG with an Adobe product, strange edges and textures are automatically sharpened and appear brighter than areas of low texture.

With ELA, grids that are not optimized for your quality level will display grid squares that change during rendering. For example, digital cameras do not optimize images for the camera's quality level (high, medium, low, etc.). Original digital camera images appear heavily modified when restored (high ELA values). Each retrieval results in a darker ELA result and a lower error rate. If you save the times again, the grid squares will reach a small error level and will not change.

B. VGG16

The Convolutional neural network (CNN) also known as a ConvNet is a type of artificial neural network. A CNN consists of a convolutional layer, a pooling layer, a fully connected (thick) layer, an input layer, and an output layer, in these layers the number of first three layers can be greater than one in any model. VGG16 is a type of CNN (Convolutional Neural Network) and is considered one of the best computer vision models to date. VGG stands for Visual Geometry Group, which created the VGG network. The authors of this model evaluated the network and increased its depth by using an architecture with very small (3×3) convolutional filters, which showed a significant improvement over previous technical configuration. They pushed the depth to 16–19 weight layers making it approximately- 138 trainable parameters.

What is VGG16 used for:

VGG16 is object detection and classification algorithm which is able to classify 1000 images of 1000 different categories with 92.7% accuracy. It is one of the most popular algorithms used for image classification and is easy to operate with transfer learning.

Architecture of VGG16:

- The 16 in VGG16 refers to 16 layers that have weights. VGG16 consist 13 convolutional layers, 5 Max Pooling layers, and 3 Dense layers i.e., total of 21 layers, but there are only 16 weight layers i.e., learnable parameters layer.
- VGG16 takes input size as 224*224*3 where 224*224 are dimension (height-width) and 3 is the RGB channel.
- Most unique thing about VGG16 is that instead of having a large number of hyper-parameters they focused on having convolutional layers of 3x3 filter with stride 1 and having a maxpool layer of 2x2 filters of stride 2.
- Conv-1 Layer has 64 number of selected filters, Conv-2 has 128 selected filters, Conv-3 has 256 selected filters, Conv 4 and Conv 5 has 512 selected filters.
- For three Fully-Connected (FC) layers: the first two have 4096 channels each, the third performs 1000- way ILSVRC classification and thus contains 1000 channels. The last layer is the soft-max layer which assigns probability to each class.

C. VGG19

The concept of the VGG19 model is the same as the VGG16 except that it supports 19 layers. The 16 and 19 in the name stand for the number of weight layers in the model(convolutional layers). This means that VGG19 has three more convolutional layers compared to VGG16. The VGG19 network model gives more accuracy than the VGG16 model. But these models are great individually. Architecture of VGG19:



- The size of input image for VGG19 is same as the size of the image for VGG16 model.
- The only preprocessing difference is the creators subtracted the average RGB value from each pixel, calculated over the entire training set.
- Used kernels of (3 * 3) size with a stride size of 1 pixel, this enabled them to cover the entire concept of the picture.
- spatial padding was used to maintain the spatial resolution of the image.
- max pooling was performed over a 2 * 2 pixel windows with stride 2.
- Since previous models used tanh or sigmoid, So here they used Rectified Linear Unit (ReLU) to include non-linearities to improve model classification and improve computation time compared to previous models.
- Now it includes three fully connected layers: out of which the first two were of size 4096 and after that the third layer with 1000 channels for 1000- way ILSVRC classification and the last layer is a layer of softmax function.

D. *The main existing methods*

The techniques that are used to address two primary forms of image manipulation are copy-move tampering (CMF) and image splicing tampering. Among these, copy-move forgery is the most commonly employed by counterfeiters. Researchers have shown a significant interest in copy-move forgery and have introduced numerous methods in this domain. CMF involves the act of copying a portion of an image and pasting it into another area within the same image. Conversely, splicing entails the fusion of various sections from different images to create a manipulated image. Detection of splicing can be accomplished by examining its impact on image statistics, assessing variations in lighting across the image, or scrutinizing the boundaries of foreign regions [1]. In 2018, Hesham A. Alberry and colleagues introduced a copy-move forgery detection method that utilizes a fast and efficient approach by optimizing the Scale-Invariant Feature Transform (SIFT) and fuzzy C-means (FCM) clustering. The technique is founded on the SIFT algorithm for feature extraction, with the fuzzy C-means clustering method employed to reduce the time complexity associated with SIFT. The evaluation of their method involved measuring True Positive Rate (TPR), False Positive Rate (FPR), and time complexity. To achieve the best results, the researchers manipulated three key parameters within the FCM algorithm, namely the number of clusters, the maximum number of clusters to create, and the minimum threshold for improvement. It's important to note that the results obtained are influenced by the specific datasets used. The study revealed that the TPR for the MICCF220 dataset outperformed that of their dataset [2].

In 2018, Shruti Ranjan and colleagues introduced a novel approach to detect digital image forgery. Their method involved a series of steps to enhance image quality and identify forged elements within images. First, they improved image quality through histogram equalization, a technique that adjusts the distribution of pixel values to enhance overall clarity. Next, they applied a median filter to remove noise from the images, ensuring that the final output was as clear as possible. To further analyze and segment the images effectively, they employed K-means clustering, a method that groups similar pixels together. This segmentation process aids in identifying distinct regions within the image, which can be crucial for forgery detection. The team then harnessed the Gray Level Co-occurrence Matrix (GLCM) to extract image features and a linear kernel Support Vector Machine (SVM) as a classifier to distinguish between genuine and manipulated images. Subsequently, they applied an Artificial Neural Network (ANN) classifier, which demonstrated superior performance compared to the linear SVM. The ANN classifier achieved an impressive accuracy rate of 96.4% [3]. As an illustration, Muhammad Jaleed Khan and his team introduced a method in their research that relies on Fuzzy C-means Clustering (FCM). To assess the effectiveness of their approach, they utilized the publicly available UWA Writing Inks Database. Their method involved employing FCM to analyze and process image data, aiming to detect any potential forgeries within the images [5].

In 2018, D.G. Savakar and their team introduced an innovative method that combines blind and non-blind watermarking techniques, as documented in their research. This hybrid approach seamlessly integrates both techniques to enhance watermarking effectiveness. For the internal watermarking process, they incorporated the blind watermarking technique. Within this phase, they embedded a secret watermark into the internal cover image. To accomplish this, they employed the Discrete Wavelet Transform (DWT) in conjunction with the blind watermarking technique, and for the external watermarking, they utilized the non-blind watermarking technique. They embedded the internal watermark into the external cover image using DWT. The output of this step was the creation of a Hybrid Watermarked Image. To assess the performance of their approach, the researchers employed similarity measurements such as correlation, structural similarity (SSIM), and Peak signal-to-noise ratio (PSNR) to compare the original watermark with the extracted watermark. The results of their evaluation indicated that this hybrid watermarking method exhibited robustness against noise, making it a promising approach for secure watermarking in various applications [4]. Active forgery detection techniques, which encompass image protection tools like watermarking and digital signatures, play a vital role in safeguarding digital content.



These methods involve the insertion of a recognized authentication code into the image before transmission, followed by a subsequent verification of its authenticity by the recipient. While active forgery detection techniques offer several advantages, there are limitations to consider. These methods are primarily effective when dealing with pre-processed images, and they might not be as versatile in identifying unaltered images. However, one of the key strengths of active forgery detection techniques lies in their reliability and certainty, making them a robust and trustworthy choice in comparison to passive forgery detection methods [6].

Introducing a deep neural network architecture called ManTra-Net for the detection and localization of real-life image forgery. This architecture addresses various types of image manipulations, including splicing, copy-move, removal, enhancement, and even unknown types, making it a comprehensive solution. In summary, ManTra-Net is presented as a unified deep neural network architecture for addressing the complex problem of real-life image forgery. It leverages self-supervised learning, anomaly detection, and LSTM techniques to achieve accurate detection and localization of a wide range of image manipulation types. The extensive experiments conducted in the study show its effectiveness and applicability in various scenarios [7].

Describes a method for detecting splicing forgery in images. The method comprises two main components: a Coarse-to-Refined Convolutional Neural Network (C2RNet) and diluted adaptive clustering. The paper appears to focus on the development of a specific deep learning-based method for detecting splicing forgery in images. The approach involves a combination of CNNs, feature extraction, and possibly clustering techniques to improve the accuracy and efficiency of forgery detection. The diluted adaptive clustering may play a role in grouping similar regions or patches for more effective analysis [8].

Introduction to a research paper that discusses the use of self-supervised learning methods for training deep neural networks in computer vision applications. The paper seems to provide an extensive review of various aspects related to self-supervised learning in the context of visual feature learning from images and videos. In essence, this paper seems to be a comprehensive review of self-supervised learning methods in the context of computer vision, with a focus on the learning of visual features from large-scale unlabeled data. It covers a wide range of topics related to this research area and aims to provide insights and guidance for future work in self-supervised learning for computer vision applications [9].

The paper discusses biometric feature-based human authentication systems and specifically focuses on fingerprint-based authentication. The paper likely goes on to discuss potential solutions or improvements in fingerprint-based authentication systems, addressing the limitations and security concerns associated with using minutiae points directly as user templates. Biometric authentication methods, including fingerprint recognition, are essential in various applications, including access control, mobile devices, and identity verification. Researchers continuously explore ways to enhance their accuracy and security [10].

The context and purpose of a review paper focused on visual media integrity verification, with an emphasis on the detection of manipulated images and videos, including the emerging threat of deepfakes. This review paper addresses the urgent need for automated tools to detect manipulated multimedia content, particularly deepfakes, and it aims to analyze the current state of forensic methods, limitations, challenges, and potential areas for future research in the context of visual media integrity verification and security [11].

A concerning technology that has advanced to the point where it is challenging to distinguish between real and fake content. It highlights the unethical and malicious applications of deepfake technology, such as the spread of misinformation about anything, impersonation of any person, and the defamation of innocent people. The article appears to serve as an informative and educational resource to help readers gain a deeper understanding of the deepfake phenomenon, its implications, and the ongoing efforts to develop effective countermeasures. Addressing the challenges posed by deepfake technology is essential for maintaining the integrity of digital content and safeguarding against misuse and malicious activities [12].

E. Proposed System

As mentioned earlier we are using VGG and ELA for forgery detection. We gave CASIAv2 dataset for model training the model. First the image is compressed to the optimal size required by the CNN. After training the image will be converted into binary format using Error level analysis and partitioned into different blocks. The blocks are analyzed and the error rate of each block is observed. If there is difference in error rate of a block then it is classified as forged. If there is no difference in error then the image is classified as original.

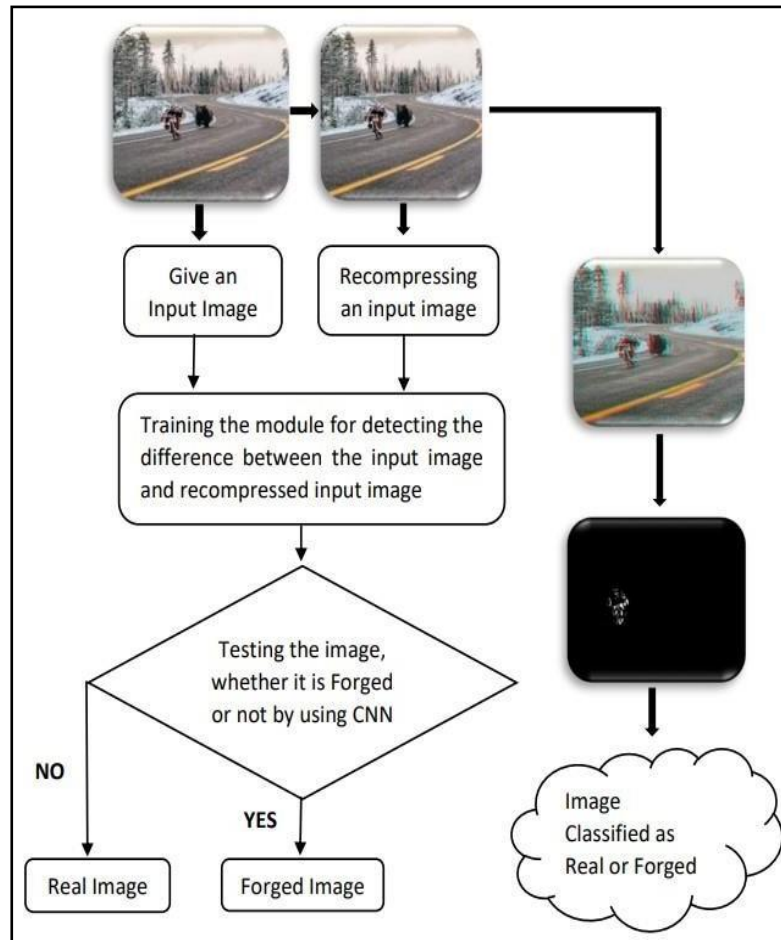


Fig. 5. Proposed System

IV. CONCLUSION

The proposed system tends to detect the copy-move and splicing type of forgeries. We used a dataset CASIAv2 that has 12.5K images, some of them are tampered and a majority is authentic. We used CNN networks with Error level Analysis to detect the forgeries. We came to a conclusion that CNN+ELA, ELA+VGG16 and ELA+VGG19 give an accuracy of 93%, 93.21% and 95.12% respectively. Hence it is preferable to use VGG19 CNN model to get the best result. Furthermore, our research found that even when done by specialists, picture manipulation can be identified with an accuracy of more than 84 percent. By performing this research, we can conclude that the model i.e., CNN+ELA has a scope or area of improvement.

REFERENCES

- [1]. N. Kanagavalli and L. Latha, "A survey of copy-move image forgery detection techniques", 2017 International Conference on Inventive Systems and Control (ICISC), 2017, pp. 1-6.
- [2]. H. A. Alberry, A. A. Hegazy and G. I. Salama, "A fast SIFT based method for copy move forgery detection", Future Computing and Informatics Journal, Elsevier, 2018, 3, pp. 159-165.
- [3]. S. Ranjan, P. Garhwal, A. Bhan, M. Arora, and A. Mehra, "Framework for image forgery detection and classification using machine learning," in 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI). IEEE, 2018, pp. 1-9.
- [4]. D. G. Savakar and A. Ghuli, "Robust invisible digital image watermarking using hybrid scheme," Arabian Journal for Science and Engineering, vol. 44, no. 4, pp. 3995-4008, 2019.
- [5]. M. J. Khan, A. Yousaf, K. Khurshid, A. Abbas, and F. Shafait, "Automated forgery detection in multispectral document images using fuzzy clustering," in 2018 13th IAPR International Workshop on Document Analysis Systems (DAS). IEEE, 2018, pp. 393-398.



- [6]. R. Dobre, R. Preda, and A. Marcu, "Improved active method for image forgery detection and localization on mobile devices," in 24th International Symposium for Design and Technology in Electronic Packaging. IEEE, 2018, pp. 255–260.
- [7]. Wu, Y.; Abd Almageed, W.; Natarajan, P. ManTra-Net: Manipulation Tracing Network for Detection and Localization of Image Forgeries With Anomalous Features. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 9535–9544.
- [8]. Xiao, B.; Wei, Y.; Bi, X.; Li, W.; Ma, J. Image splicing forgery detection combining coarse to refined convolutional neural network and adaptive clustering. *Inf. Sci.* 2020, 511, 172–191.
- [9]. Jing, L.; Tian, Y. Self-supervised Visual Feature Learning with Deep Neural Networks: A Survey. *IEEE Trans. Pattern Anal. Mach. Intell.* 2020, 43, 1.
- [10]. Ali, S.S.; Baghel, V.S.; Ganapathi, I.I.; Prakash, S. Robust biometric authentication system with a secure user template. *Image Vis. Comput.* 2020, 104, 104004.
- [11]. Verdoliva, L. Media Forensics and DeepFakes: An Overview. *IEEE J. Sel. Top. Signal Process.* 2020, 14, 910–932.
- [12]. Mirsky, Y.; Lee, W. The Creation and Detection of Deepfakes: A Survey. *ACM Comput. Surv.* 2021, 54, 1–41.
- [13]. Bappy JH, Simons C, Nataraj L, Manjunath BS, Roy-Chowdhury AK. Hybrid lstm and encoder–decoder architecture for detection of image forgeries. *IEEE Trans Image Process.* 2019;28(7):3286–3300. doi: 10.1109/TIP.2019.2895466.
- [14]. Camacho IC, Wang K (2021) Data-dependent scaling of CNN’s first layer for improved image manipulation detection. In: Digital Forensics and Watermarking: 19th International Workshop, IWDW 2020, Melbourne, VIC, Australia, November 25–27, 2020, Revised Selected Papers. Springer Nature, vol. 12617, p 208. 10.1007/978-3-030-69449-4_16.
- [15]. Cozzolino, D, Verdoliva, L (2020) Noiseprint: A CNN-based camera model fingerprint. *IEEE Trans Inf Forensics Secur* 15:144–159. 10.1109/TIFS.2019.2916364.
- [16]. Ding X, Raziei Z, Larson EC, Olinick EV, Krueger P, Hahsler M. Swapped face detection using deep learning and subjective assessment. *EURASIP J Inf Secur.* 2020;2020(1):1–12.
- [17]. Castillo Camacho I, Wang K. A comprehensive review of Deep- learning-based methods for image forensics. *J Imaging.* 2021;7(4):69. doi: 10.3390/jimaging7040069.
- [18]. L. Zheng and Y. Zhang, "A Survey on Image Tampering and Its Detection in Real world Photos A Survey on Image Tampering and Its Detection in Real-world Photos," no. December, 2018, doi: 10.1016/j.jvcir.2018.12.022.
- [19]. J. Ouyang, Y. Liu, and M. Liao, "Copy-move forgery detection based on deep learning," Proc. - 2017 10th Int. Congr. Image Signal Process. Biomed. Eng. Informatics, CISP-BMEI 2017, vol. 2018- Janua, pp. 1–5, 2018, doi: 10.1109/CISP BMEI.2017.8301940.
- [20]. W. Zhang and C. Zhao, "Exposing Face-Swap Images Based on Deep Learning and ELA Detection," vol. 5, no. November, p. 29, 2020, doi: 10.3390/ecea-5-06684.
- [21]. Adobe Photoshop. Accessed 16 Mar 2022. <https://www.adobe.com/it/products/photoshop.html>.
- [22]. Chen J, Liao X, Qin Z (2021) Identifying tampering operations in image operator chains based on decision fusion. *Sig Process Image Commun* 95:116287. <https://doi.org/10.1016/j.image.2021.116287>.
- [23]. Elaskily M, Elnemr H, Sedik A, Dessouky M, El Banby G, Elaskily O, Khalaf AAM, Aslan H, Fara gallah O, El-Samie FA (2020) A novel deep learning framework for copy-move forgery detection in images. *Multimed Tools Appl* 79. <https://doi.org/10.1007/s11042-020-08751-7>.
- [24]. Faceswap. <https://github.com/deepfakes/faceswap>. Accessed 16 Mar 2022.