Gurukul International Multidisciplinary Research Journal (GIMRJ)*with*
International Impact Factor 8.249
Peer Reviewed Journal
https://doi.org/10.69758/NCII4879

e-ISSN No. 2394-8426
Special Issue On Advanced Computational Techniques: Emerging Trends from Postgraduate Studies
Issue–I(VI), Volume–XII

Advance Image Forgery Detection

Nikita Nitin Tinkhede
PG Student
Department of computer science,
G H Raisoni Amravati University Nagpur, India

Received on: *11 April ,2024* ***Revised on:*** *26 May ,2024,* ***Published on:*** *01 June ,2024*

Abstract :- In the digital era, the proliferation of sophisticated image editing tools has made image forgery a prevalent issue, posing significant challenges in various fields, including legal proceedings, journalism, and digital security. Advanced image forgery detection techniques are crucial to ensuring the authenticity and integrity of digital images. This paper explores the state-of-the-art methods and technologies employed in the detection of image forgeries. We focus on both traditional and deep learning-based approaches, highlighting their respective strengths and weaknesses.

Traditional methods, such as pixel-based, format-based, and geometry-based techniques, rely on the intrinsic properties of images to identify inconsistencies and manipulations. These methods are effective in detecting simple forgeries but often fall short against more sophisticated alterations. On the other hand, deep learning-based approaches, leveraging convolutional neural networks (CNNs) and generative adversarial networks (GANs), have shown significant promise in identifying subtle and complex forgeries by learning high-level features from large datasets.

In conclusion, while significant advancements have been made in image forgery detection, continuous research and development are necessary to stay ahead of emerging forgery techniques. The integration of multidisciplinary approaches and the collaboration between academia and industry will be pivotal in advancing the efficacy of forgery detection systems.

Keywords : Image Forgery Detection, Deep Learning, Digital Image Authentication, Pixel-based Methods, Image Manipulation. Image Manipulation.

1. INTRODUCTION

In today's digital age, the manipulation and alteration of images have become increasingly prevalent, leading to significant challenges in ensuring the authenticity and integrity of visual content. With the proliferation of sophisticated image editing tools and techniques, detecting image forgeries has become a critical area of research and development. While traditional forgery detection methods have primarily focused on detecting manipulations in realistic or natural images, the detection of forgeries in abstract images presents unique challenges due to the absence of concrete objects or recognizable features.

Abstract art, characterized by its non-representational and often ambiguous nature, poses a distinct set of challenges for forgery detection. Unlike realistic images, which contain recognizable objects and scenes, abstract images rely on elements such as color, texture, shape, and composition to convey meaning and evoke emotions. As a result, detecting forgeries in abstract art requires specialized techniques that can analyze the underlying visual patterns and structures inherent to abstract imagery.

However, the field of image forgery detection faces several challenges. The adversarial nature of forgery means that as detection methods improve, so do the techniques used by forgers. This ongoing battle necessitates continuous research and innovation. Additionally, deep learning models require vast amounts of annotated data for training, which can be difficult to obtain. There are also significant computational requirements for training and deploying these models, which can be a barrier to widespread adoption.

This paper explores the current landscape of image forgery detection, examining both traditional and deep learning-based approaches. We analyze their respective strengths and limitations and propose a hybrid model that leverages the advantages of both methods. Furthermore, we discuss the potential for emerging technologies, such as blockchain, to enhance the traceability and verification of digital images. The goal is to provide a comprehensive overview of the field and highlight future directions for research and development.

2. RELATED WORK:

Related work in a research paper. The field of image forgery detection has seen significant advancements over the years, with various methodologies being developed to tackle the problem. The existing literature can broadly be categorized into traditional approaches and deep learning-based methods, each with its own set of techniques and applications.

Convolutional Neural Networks (CNNs):

- **Bayar and Stamm (2016)**: Developed a novel CNN architecture specifically designed to learn features indicative of image manipulation, demonstrating significant improvements over traditional methods.
- **Rao and Ni (2016)**: Introduced multi-scale CNNs, which analyze images at various resolutions to capture both local and global anomalies indicative of forgery.

Generative Adversarial Networks (GANs):

- **Wang et al. (2019)**: Employed GANs to generate synthetic forgeries for adversarial training, enhancing the robustness of detection models against sophisticated forgeries.
- **Zhou et al. (2018)**: Proposed a two-stream network, where one stream focuses on image content and the other on artifacts introduced by GANs, achieving high accuracy in detecting GAN-generated images.

3. LITERATURE SURVEY

Image Forgery Detection Using Recompressing Images, carried out by Syed Sadaf Ali [1] The techniques used are adapted to the individual needs, interests, and preferences of the user or society. Image compression involves reducing the pixels, size, or colour components of images in order to reduce the file size for forgery detection. Advanced image optimization techniques can detect the more important image components and discard the less vital ones. Image Forgery Detection by using Support Vector Machine developed by J.math [2] Forgery detection technique that uses illuminan color inconsistency and machine learning classifiers such as Support Vector Machine (SVM). SVM is a supervised classification algorithm that is used to differentiate between two separate categories by drawing a line between them. In this technique, the illuminant color of input images is estimated, and illuminant maps are created for each image.

2. Furthermore, all faces present in one image and corresponding faces of other individual images are extracted for investigation. However, it seems that this technique has some drawbacks, such as requiring clear textural and inclination highlighting and affecting the acknowledged substance of the image entirely.It is worth noting that there are several other forgery detection techniques available that use different approaches, such as image forensics, watermarking, and deep learning-based methods. Each technique has its advantages and limitations, and the selection of an appropriate technique depends on various factors such as the type of forgery, the available data, and the required level of accuracy.

A Full-Image Full-Resolution End-to-End-Trainable CNN Framework for Image Forgery Detection, [6] carried out by F. Marra It proposes a framework for detecting image forgery using a convolutional neural network (CNN).The framework includes a feature extraction module and a classification module, both using CNNs, and it operates on full-resolution images.The dataset used is authentic and forged images, including various types of forgeries, to train and test the framework. It also proposes a data augmentation method to improve the framework's robustness.

(Table 1 Summary of related work on Image Forgery Detection)

Literature	IMAGE FORGERY DETECTION			
	DNN	**SVM**	**GBI**	**Copy Move**
J.Malathi, et al. 2019 [2]	Yes	Yes		
F. Matern, et al. 2020 [3]			Yes	
Anushka Singh and Jyotsna Singh, 2022 [5]	Yes			
S. B. G. T. Babu and C. S. Rao, 2020 [8]				Yes
H. Chen, et al. 2020 [12]	Yes			Yes

(Table 2 Literature survey summary with techniques)

Paper	**Technique**
Syed Sadaf Ali, et al. 2022 [1]	Recompression of Images

J.Malathi, et al. 2019 [2]	SVM
F. Matern, et al. 2020 [3]	Gradient-Based Illumination
Anushka Singh and Jyotsna Singh, 2022 [5]	ResNet
F. Marra, et al. 2020 [6]	ResNet
S. B. G. T. Babu and C. S. Rao, 2020 [8]	Copy-move
M. H. Alkawaz, et al. 2020 [9]	Expectation Maximization
S. alZahir and R. Hammad, et al. 2020 [10]	Clustering Algorithm
H. Chen, et al. 2020 [12]	Copy-move

4.SYSTEM METHODOLOGY

Methodology ELA (Error Level Analysis) is a technique used for detecting image forgery. It involves compressing an image to a low quality, then re-saving it at a higher quality, and then calculating the difference between the two versions of the image. The resulting image is known as an ELA image, and it highlights the parts of the image that have been manipulated or edited.

Convolutional neural networks (CNNs) are a popular choice for image forgery detection because they can learn to recognize patterns and features in images. The CNN is trained on a dataset of real and manipulated images to learn the characteristics of forged images. Once the CNN is trained, it can be used to classify new images as either real or manipulated. To implement an image forgery detection system using ELA and CNN, the following steps can be taken: Convert the input image into an ELA image

Preprocess the ELA image and prepare it for input into the CNN. Use the CNN to classify the ELA image a either real or manipulated. If the image is classified as manipulated, further analysis can be performed to determine the type of forgery that was used. It is important to note that while ELA can be a useful technique for detecting image forgery, it is not foolproof and can produce false positives or false negatives. Therefore, it is important to combine ELA with other techniques and methods for a more accurate and robust detection system.

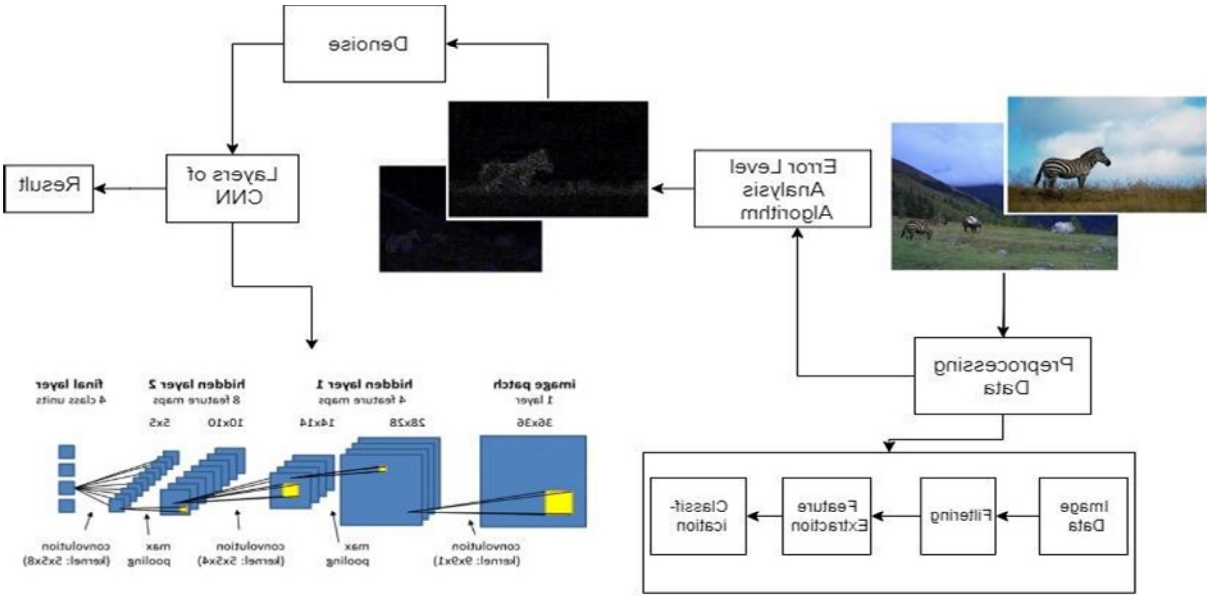


Fig 1

The input layer in a convolutional neural network (CNN) is where the images from the dataset are fed. The images are usually in the form of 3-dimensional arrays, with the first two dimensions representing the height and width of the image (the number of pixels), and the third dimension representing the red, green, and blue (RGB) colors present in each pixel.

In the feature-extraction part of the CNN architecture, the input image is passed through a series of convolutional layers, which apply a set of learnable filters to extract features from the image. Each filter produces a feature map that highlights a particular pattern or feature in the input image. These feature maps are then passed through activation functions like ReLU to introduce non-linearity and avoid the vanishing gradient problem.

After the feature extraction process, the output of the last convolutional layer is flattened into a 1-dimensional vector and fed into a series of fully connected layers for classification. The fully connected layers use the extracted features to make predictions about the class of the input image. The final output layer usually employs the softmax function to generate a probability distribution over the classes, indicating the most likely class for the input image.

5. SYSTEM IMPLEMENTATION

A. Software and Hardware: The system requirements for running an image forgery detection system using a CNN model implemented in Python 3.7.X (IDLE) and the CASIA dataset.It is recommended that the computer system has at least:

- RAM: 8GB or more
- Hard Disk Drive (HDD): 80GB or more
- Processor: i5 or higher These system requirements are necessary to handle large amounts of data and the constant nature of the environment. It is important to note that the specific hardware and software requirements may vary depending on the size of the dataset and the complexity of the CNN model being used. Therefore, it is always a good idea to check the specific requirements of the software and datasets being used before implementing an image forgery detection system.

Additionally, it is recommended to have sufficient cooling and power supply to ensure that the system can run smoothly and avoid any unexpected shutdowns or errors.

B. Dataset

The CASIA v2.0 database contains a total of 10,000 images, divided into two subsets: a training set of 5,000 images and a testing set of 5,000 images. Each subset includes eight categories of images: animal, architecture, article, character, nature, plant, scene, and texture. The images are in JPEG format and have a size of either 256 x 384 or 384 x 256 pixels.CASIA V2.0 dataset is used for image forgery detection.

Two classes make up this dataset: actual photos and tampering detection. There are 7354 images, which are classified into real images and altered images in JPG format.

Dataset	Size	Categories	Format
CASIA V2.0	5 GB	8 categories of images	JPEG

(Table 3 Details of CASIA Dataset)

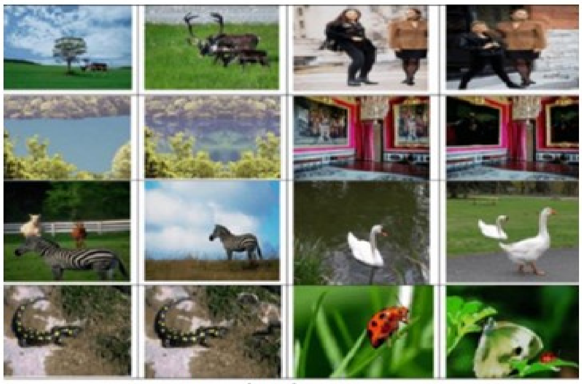
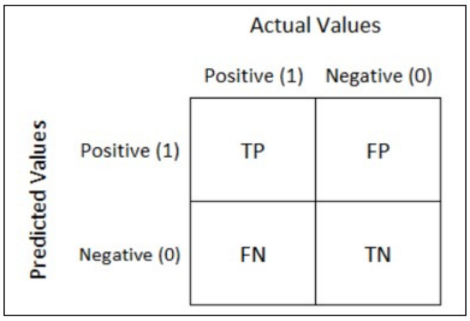


Fig.5. Images from the CASIA V2.0 Dataset

(Fig.4 CASIA Dataset)

C. Confusion Metrics

A confusion matrix is a table that is commonly used to evaluate the performance of a classification algorithm by comparing the predicted labels to the true labels of a set of test data. The matrix displays the number of true positive, false positive, true negative, and false negative predictions made by the algorithm



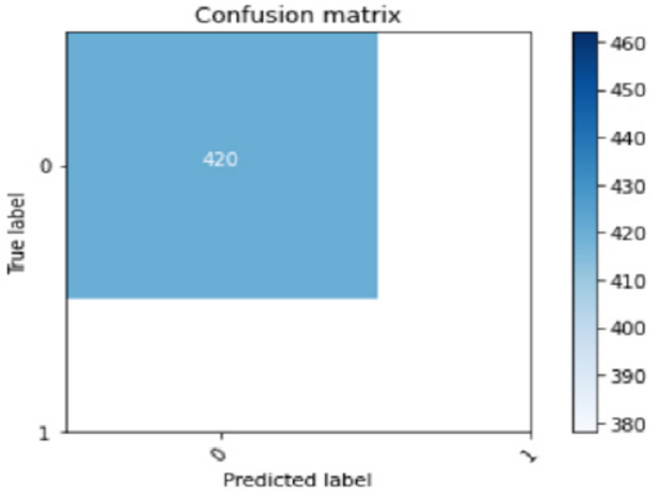
Actual Values

Predicted Values		Positive (1)	Negative (0)
	Positive (1)	TP	FP
	Negative (0)	FN	TN

(Fig.5 Confusion Matrix)

The above table is a confusion matrix that summarize the performance of a binary classification model, and it includes four possible outcomes: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). True positives occur when the model correctly predicts a positive outcome, and true negatives occur when the model correctly predicts a negative outcome. False positives occur when the model predicts a positive outcome, but the actual outcome is negative, and false negatives occur when the model predicts a negative outcome, but the actual outcome is positive.

The above table has the following cases:



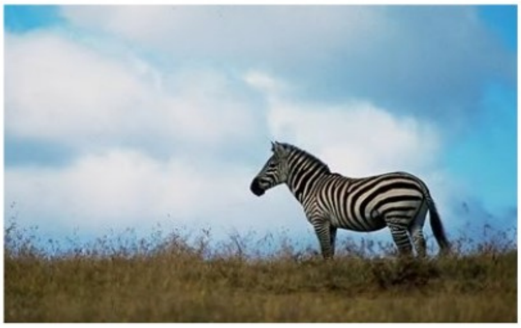
(Fig.6 Confusion matrix of image forgery detection)

A confusion matrix is a table used to evaluate the performance of a classification model on a set of test data whose true class labels are known

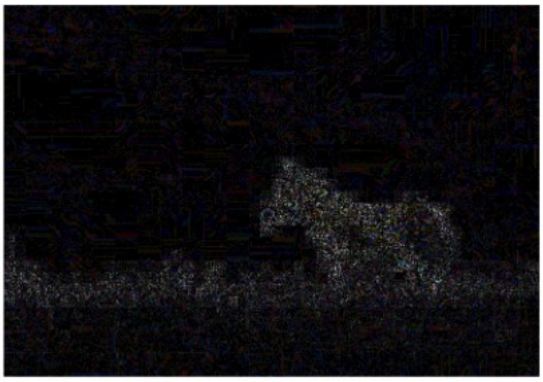
6. RESULT ANALYSIS

The original image and its ELA-converted counterpart are shown in Fig. 7 and Fig. 8 of the dataset, respectively. And the fake image and its corresponding ELA-converted image are shown in Figs. 9 and 10, respectively

In Fig. 11, the red line represents the model's training loss and training accuracy, while the blue line represents the model's validation loss and validation accuracy. The model is iteratively trained and has an accuracy of 78.08%.



(Fig.7 Original image from dataset)



(Fig.8 fake image from dataset)

7. CONCLUSION:

Once the forged images are recognized, they are displayed as output. A confusion matrix is used to evaluate performance, and the findings are displayed in a table that takes into account all of a classifier's anticipated and actual values. The confidence score is then computed as an evaluation standard.

1.Summary of Findings: Performance Evaluation: Evaluate the performance metrics including accuracy, precision, recall, F1 score, AUC-ROC, false positive rate, and false negative rate. Comparison with State-of-the-Art: Discuss how the proposed method compares with existing methods in terms of detection accuracy and computational efficiency.Real-World Testing: Summarize the results from testing the system under various conditions and scenarios, including mobile and edge devices.

2.Key Contributions: Innovative Approach: Highlight the novelty and contribution of the proposed approach, whether it's a hybrid model integrating traditional methods with deep learning techniques, the use of blockchain for provenance, or optimization for real-time detection. Performance Improvements: Discuss how the proposed method improves upon existing approaches, particularly in terms of accuracy, robustness, and efficiency.

3.Implications and Significance: Practical Applications: Discuss the implications of the research findings in real-world applications, such as forensic analysis, journalism, legal evidence, and digital security. Security and Trust: Emphasize the importance of reliable forgery detection in maintaining trust in digital media and preventing malicious manipulation.

4.Conclusion Statement: Overall Impact: Summarize the overall impact of the research on advancing the field of image forgery detection. Call to Action: Advocate for continued research and development efforts to address the evolving challenges of digital image manipulation.

8. REFERENCE:

1. Ali, S.S.; Ganapathi, I.I.; Vu,N.-S.; Ali, S.D.; Saxena, N.; Werghi, N., "Image Forgery Detection Using Deep Learning by Recompre -ssing Images," Electronics 2022, 11, 403.

2. J.Malathi, B.Narasimha Swamy, Ramgopal Musunuri, "Image Forgery Detection by using Machine Learning, International Journal of Innovative Technology and Exploring Engineering (IJITEE)ISSN: 2278-3075, Volume-8, Issue- 6S4, April 2019.

3. F. Matern, C. Riess and M. Stamminger, "Gradient-Based Illumination Description for Image Forgery Detection," in IEEE Transactions on Information Forensics and Security, vol. 15, pp. 1303-1317, 2020, doi:10.1109/TIF S.2019.2935913.

4. Z. J. Barad and M. M. Goswami, "Image Forgery Detection using Deep Learning: A Survey," 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), 2020, pp. 571-576, doi: 10.1109/ICACCS48705.2020.9074408.