Digital Image Forgery Detection Using Deep Learning

CH. Lakshmi Kumari[1], Bellam Saikumar[2], Tharigoppula Shivaram[3]
[1]*Assistant Professor, Mahatma Gandhi Institute of Technology*
[2,3]*UG Student, Mahatma Gandhi Institute of Technology*

Abstract- **In today's digital age, social media is image-dominated, making visuals a central tool for communication and dissemination of information. This extensive usage of images, however, has also given rise to their malicious exploitation by means of intentional manipulation, mostly in the form of producing fictitious or false content. Image forgery—modifying images to deceive audiences—has emerged as a key problem, perpetuating the spread of misinformation, cyberattacks, and even legal issues. In response to the increasing problem, deep learning algorithms have become an effective solution in identifying fake images. Specifically, Convolutional Neural Networks (CNNs), a branch of deep learning algorithms, are very effective at identifying subtle inconsistencies in images. By conducting pixel-level analysis, CNNs can discover unnatural visual patterns, tiny inconsistencies, and indications of manipulation that go unnoticed by the human eye.**

CNNs are efficient at detecting prevalent forms of image forgery like splicing, copy-move, and retouching, by examining prominent image features such as edges, texture, and lighting irregularities. Out of numerous techniques, the combination of Error Level Analysis (ELA) with CNN structures has emerged with tremendous potential for detecting tampered areas. CNNs can learn from very large datasets including original and tampered images and make accurate classifications and localizations of forged content. Therefore, deep learning—specifically CNNs in conjunction with ELA—has been a very effective solution to counter the growing problem of image forgery, thus improving credibility and reliability in digital media.

Keywords: **Image forgery, Pixel-level analysis, Splicing, Copy-move, Image retouching, Error Level Analysis (ELA), Convolutional Neural Networks (CNNs), Tampering artifacts.**

INTRODUCTION

Images are crucial to modern communication, especially in social media websites. But the increasing number of cases of image forgery is a serious concern. Image forgery is defined as the alteration or manipulation of images in an attempt to mislead or mislead the audience. The ease of availability of sophisticated tools has enabled even amateurs to produce very authentic-looking forged images. Such manipulations have serious implications in fields such as journalism, the legal system, scientific studies, and interpersonal communication.

Digital image forgery usually creates anomalous patterns that disturb the natural features of an image. Detection processes for such fakes are typically classified as active and passive methods.

- Active detection methods use additional information added into the image when it is captured. These methods include digital signatures, introduced towards the end of image capture, and digital watermarking, implemented either during the image capture process or in the post-processing process.
- Passive detection methods, however, are not dependent on any prior embedded data. They examine the image itself and determine inconsistencies based on tampering. These methods can detect such alterations as Copy-Move Forgery (CMF), Splicing Forgery, and Retouching Forgery.
- Copy-Move Forgery (CMF) is the process of copying a section of the image and inserting it somewhere else in the same image. This process is hard to identify because the copied area is of the same nature.
- Splicing Forgery is the process of merging pieces from different images. The final image is altered in lighting, noise, and color to make it look seamless, hence hard to detect.
- Retouching Forgery involves making subtle manipulations of parameters such as brightness, contrast, or color so as to bring out or cover up certain characteristics.

As image editing software continues to evolve, forgeries have grown increasingly sophisticated and challenging to detect using naked eye inspection. This serves to underscore the critical need for robust and automated systems with the ability to precisely detect and localize image manipulations, a feature required to ensure image integrity in a wide range of applications.

A. Problem Statement

Pictures have become a pervasive means of communication and information exchange over different mediums such as social media, news organizations, and messaging services in present times. Easy access to sophisticated image editing software has resulted in an increased use of image manipulation, hence the proliferation of digital image forgery. Forgery is used to spread false information, influence public opinion, defraud people, or even tamper with evidence in court cases.

The implications of image forgery in actual application are significant. Artificial images have the potential to drive the dissemination of false information, ruin reputations, and determine political and social outcomes. In legal and forensic applications, manipulated images have the potential to mislead investigations and hinder the pursuit of justice. Moreover, the quick proliferation of manipulated material on social sites further magnifies its effect, complicating real-time authentication and content management in the process.

EXISTING SYSTEM

Modern digital image forgery detection systems heavily depend on deep learning to improve accuracy as well as processing speed. Such methods are specially designed to counter the challenges involved in detecting different types of forgeries—splicing and copy-move—under real-world, practical scenarios. The basis of the suggested method is detecting inconsistencies resulting from image compression. This is achieved by creating a difference image that emphasizes differences between an original image and a recompressed image. This created image is then passed through pre-trained deep learning models that have been optimized to identify images as real or fake.

Eight pre-trained neural network models were tested in this work, namely VGG16, VGG19, ResNet50, ResNet101, ResNet152, MobileNetV2, Xception, and DenseNet. Among them, MobileNetV2 was notable for having the highest classification accuracy of around 95% and being computationally lightweight. To suit the task of forgery detection, the original classification layers of the models were replaced with custom layers fine-tuned for binary classification.

This strategy greatly surpasses conventional techniques by lowering computational loads, efficiently coping with post-processing artifacts, and identifying various types of tampering within a single system. The results highlight the strength of transfer learning in enhancing the performance and reliability of digital image forgery detection methods.

LITERATURE SURVEY

This paper introduces novel detection of the digital image forge, focusing precisely on two splicing and two copy-move images. The developed approach relies upon a deep-learning model with extended transfer learning which enhances the chances of better image detection accuracy. The key idea is to analyse the difference in compression qualities between the forged and authentic regions of an image, which is typically undetectable by the human eye but can be detected through deep learning methods. They employ eight pre-trained models—VGG16, VGG19, ResNet50, ResNet101, ResNet152, MobileNetV2, Xception, and DenseNet—adapted for binary classification after fine-tuning. The results demonstrate that MobileNetV2 provides the highest detection accuracy (around 95%) with fewer training parameters, leading to faster training times and reduced computational costs. The proposed method significantly outperforms previous state-of-the-art techniques in terms of accuracy, precision, recall, F1 score, and AUC. Furthermore, the technique is designed to be lightweight, making it suitable for deployment in environments with limited computational resources. The study concludes the combination of transfer learning with analysis of image compression as offering robust solutions for image forgery detection in real-time, and thus future work lies in generalization to unseen data and incorporation of localization of the forged areas.[1]

The authors proposed a hybrid approach to detect SURF, A-KAZE, and DBSCAN clustering-based copy-move forgery. This approach identifies the tampered region clearly and is robust to rotation, scaling, and post-processing. Comparing the original image with the affine-transformed version, it will detect with great accuracy areas forged. When compared with datasets like Ardizzone and CoMoFoD, it stands tall in recognizing the tampered parts, mostly on smooth surfaces that other algorithms could not mark out. The technique greatly minimizes the complexity of computations at the expense of high precision, recall, and F1 scores. The effectiveness, efficiency, and resilience against complex manipulations make this an important tool in digital forensics.[2]

This hybrid framework of the Reptile Search Algorithm combined with deep learning for the purpose of detecting copy-move forgery is proposed by the authors. RSA optimizes feature selection with a reduced dimensionality without compromising the important information. Then, these features are passed to the deep learning model for the purpose of forgery classification and localization. The above approach shows greater robustness against occlusion, complex forgeries, and post-processing artifacts. Tested on benchmark datasets, this outperforms existing methods in precision, recall, and F1 score. However, the present model lacks runtime analysis and resilience against adversarial attacks; hence, a further exploration is needed for real-world application.[3]

With this paper, an active methodology capable of detecting image forgeries, especially from the perspective of social media, is proposed. The authors deal with the problems arising from compressed, low-quality images that such digital platforms commonly used. A modified U-NET model architecture is adopted here, which is further optimized by the Grasshopper Optimization Algorithm (GOA) that enhances segmentation performance. The U-NET model is primarily used in the biomedical image segmentation field and has been adapted to highlight the forged regions in the digital images. Some of the modifications included here are the addition of a few convolutional layers to both encoder and decoder pipelines, batch normalization, and better weight connections to enhance its accuracy and stability. The GOA optimizes hyper-parameters such as learning rates and mini-batch sizes to attain maximum performance. The authors show the robustness of their algorithm with the CASIA dataset concerning copy-move and splicing forgeries. The experimental results show that their proposed method performed better than existing models in terms of precision, recall, and F1 scores with better accuracy and segmentation results. The study merits the possibility of further validation on different datasets and real-time applications. The paper contributes greatly to the field by using

deep learning together with optimization techniques for forgery detection.[4]

A robust trained system proposed for image forgery detection through deep learning techniques, especially splicing manipulation. Authors made use of ResNet50v2 architecture for their training and used transfer learning with pre-trained weights from a YOLO CNN model. This made the system capable of extracting meaningful features effectively, thus reducing training time and computational complexity. The proposed system was trained and tested on two benchmark datasets, CASIA_v1 and CASIA_v2, having labelled examples of authentic and forged images. After a series of experiments, the authors verified that the system achieved an elegant 99.3% accuracy on the CASIA_v2 dataset. Such performance is unprecedented when compared to results without transfer learning, thus reinforcing the usefulness of pre-trained models. The results of this study summarize the efficiency and reliability of the proposed method in spliced image detection. However, they also emphasize the need for further validation in quite diverse forgery scenarios and data sets. Their work sets the stage for further development of the system toward various types of digital image manipulation.[5]

The authors introduce a new two-stage hybrid approach for the detection of copy-move forgeries in digital images. The authors combined CNN architectures with the CenSurE keypoint detection algorithm to enhance the robustness and accuracy of forgery detection. The approach handles challenges related to geometric transformations, such as scaling and rotation, and post-processing operations, such as JPEG compression, noise addition, and brightness adjustments. The first step uses the CenSurE keypoint detector and FREAK descriptors to detect possible forged regions. RANSAC is then used to filter out outliers using the Random Sample Consensus. In the second step, the CNN model extracts image features with a deep learning-based classifier to enhance the detection and localization performance. Large-scale experiments were carried out on seven benchmark datasets, obtaining better results in various forgery scenarios. The hybrid model achieved high F1 scores while maintaining computational efficiency, processing images faster than existing methods. The findings highlight the approach's ability to detect forgeries in images with smooth, dense, or self-similar textures, making it a robust solution for multimedia forensics. Future work could explore further real-time scalability and adversarial robustness.[6]

PROPOSED SYSTEM

The proposed methodology for digital image forgery detection presents a strong and wise system integrating the power of Convolutional Neural Networks (CNNs) and Error Level Analysis (ELA). The hybrid system improves both accuracy and precision in forgery detection by leveraging ELA's ability to expose areas of compression anomalies and CNNs' good feature extraction ability. ELA operates by examining how heavily each component of an image has been compressed, the hypothesis being that regions of tampering tend to exhibit varying levels of compression relative to intact areas. These areas are then mapped visually onto ELA-processed images, which are then transformed into NumPy arrays and used as input for CNN models.

The experimental configuration incorporates a number of well-known CNN models like MobileNet, VGG, and DenseNet, among others, pre-trained CNN architectures to establish the best model for image classification as authentic or forged. These models are then fine-tuned for a particular task of forgery detection so that the system not only detects tampering has taken place but is also able to specify the precise regions of tampering. The integration of ELA's localized expertise with CNNs' pattern recognition based on deep learning strongly enhances the detection capability of splicing and finer pixel-level forgery that typical methods tend to overlook.

The combined approach addresses key shortcomings of current forgery detection methods through enhanced precision as well as dependability. The ELA preprocessing step enables the model to generalize more accurately across various kinds of manipulations, from trivial edits such as brightness changes to intricate changes such as region cloning and image splicing. The comparative analysis over various CNN models enables optimization over accuracy, rates of false positives, and processing costs, and MobileNet tends to be a preferred choice as it balances between performance and utilization.

In contrast to traditional detection approaches that often depend on hand-designed features or do not generalize well across various tampering methods, this system minimizes error rates by taking advantage of the complementary strengths of ELA and CNNs. Its modular and flexible design makes it applicable to a broad variety of applications, such as digital forensics, media authentication, legal investigations, and research in academia. Finally, the method proposed above is a well-rounded, scalable, and cost-effective solution for addressing the burgeoning requirement for foolproof image forgery detection in today's highly-digital world.

SYSTEM ARCHITECTURE

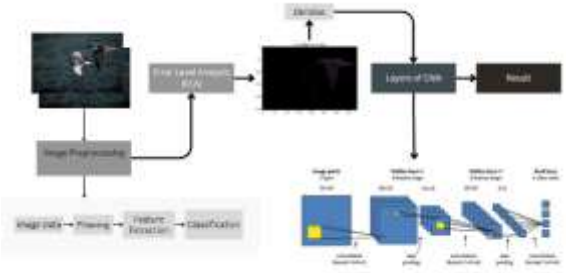

Fig: System Architecture

The system employs a Convolutional Neural Network (CNN) integrated with Error Level Analysis (ELA) for image forgery detection. The process begins by converting the input image into an ELA image, which highlights potential tampered regions by exposing compression inconsistencies.

This ELA-transformed image is then passed through the CNN. The architecture consists of two convolutional layers (Conv1, Conv2) with 32 filters each, followed by a max pooling layer to reduce dimensionality and a dropout layer to prevent overfitting. The feature maps are flattened and fed into two fully connected layers—FC1 with 150 neurons and FC2 with 2 output neurons—representing "Authentic" and "Tampered" classes. A SoftMax function is used for final classification. This architecture effectively captures pixel-level anomalies and ensures robust forgery detection.

RESULTS

The suggested digital image forgery detection system, developed on a deep learning framework augmented with Error Level Analysis (ELA), has shown robust performance in correctly classifying manipulated vs. original images. ELA is instrumental in preprocessing by amplifying compression artifacts, allowing the model to identify pixel-level discrepancies that are usually imperceptible to the human eye. Such disparities, when expressed as the ELA difference image, offer the critical localized insight that assists the deep learning network in making exact determinations

Fig: Viewing the Uploaded Image is Authentic/ Original Image

Fig: Viewing the Uploaded Image is Forge/ Tampered Image

When being used in practice, the application shows both the original image as well as the ELA difference image side-by-side. When presented with legitimate images, the ELA response is evenly varied and slight in appearance, devoid of any indicia of manipulations. The model correctly classifies such images as "Authentic," as indicated by the green overlay text on the screen. In contrast, in forged images, even though they appear to be original at first glance, the system correctly identifies them based on the evaluation of minute differences in compression artifacts. The forged areas are clearly demarcated in the ELA image with glowing spots or color noise, representing manipulated areas. This feature highlights the ability of the system to detect such forgeries as splicing, object insertion, or local editing.

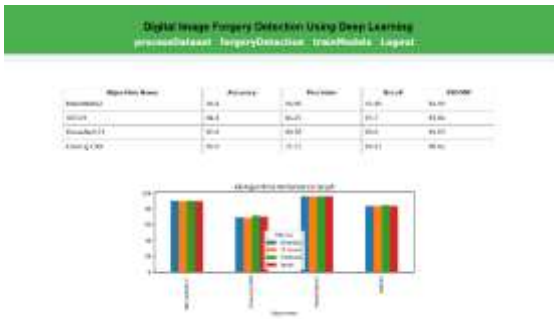

Fig: Viewing or comparing the results of different models

The system was extensively tested by comparing the accuracy of four various CNN models: MobileNetV2, DenseNet121, VGG19, and an Existing CNN model. Among them, the most precise and reliable model proved to be MobileNetV2, which has a tremendous accuracy of 96%, combined with high precision, recall, and F1-score values. Its performance bars consistently high in all assessment parameters reflect its strength in identifying spoofed images and reducing misclassifications. DenseNet121 came in second, achieving 90% accuracy. Although its recall wasn't as high (89.6%), it still reflected strong detection performance with little loss of integrity. VGG19 reached a modest accuracy of 84%, where precision (84.25%) was marginally higher than recall (83.7%), suggesting a fairly conservative model that does not produce false positives at the expense of losing some forgeries. In

contrast, the Existing CNN model fell far behind, with low accuracy (69%) and an F1-score of 68.62%, demonstrating its poor capacity for balancing precision and recall.

The above results were quantitatively confirmed by applying key performance indicators:

- Accuracy, calculated as the proportion of correctly identified authentic and forged images, reflects the overall effectiveness.
- Precision evaluates how many of the detected forgeries were actually forged, important for reducing false positives.
- Recall measures the ability to detect all instances of forgeries, indicating the sensitivity of the model.
- F1-score, a harmonic mean of precision and recall, provides a balanced assessment of the model's performance.

CONCLUSION

The proposed system for Digital Image Forgery Detection using Deep Learning effectively deals with the emerging challenge of image tampering in the digital age. The integration of Convolutional Neural Networks (CNNs) with Error Level Analysis (ELA) enhances the accuracy, robustness, and adaptability of detection across various forgery techniques. The use of ELA as a preprocessing step helps identify compression inconsistencies, and CNNs extract complex features for reliable and precise detection of manipulations, such as copy-move, splicing, and retouching forgeries.

This method overcomes the limitations of current approaches because it provides localized insight into the regions that are tampered with, reduces false positives, and gives a much more comprehensive and automated solution. The experimental evaluation confirmed the system's high performance in detecting sophisticated forgeries and therefore is an effective application for use in media, forensics, and digital security.

Future improvements may focus on optimizing computational efficiency, addressing adversarial attacks, and expanding dataset diversity for real-world deployment. Overall, this system contributes significantly to digital forensics and image authenticity verification, strengthening trust in visual media.

REFERENCE

[1] H. Khalil, A. Z. Ghalwash, H. A. -G. Elsayed, G. I. Salama and H. A. Ghalwash, "Enhancing Digital Image Forgery Detection Using Transfer Learning," in *IEEE Access*, vol. 11, pp. 91583-91594, 2023, doi: 10.1109/ACCESS.2023.3307357

[2] Fu, G.; Zhang, Y.; Wang, Y. Image Copy-Move Forgery Detection Based on Fused Features and Density Clustering. *Appl. Sci.* 2023, *13*, 7528.

[3] M. Maashi *et al*., "Modeling of Reptile Search Algorithm with Deep Learning Approach for Copy Move Image Forgery Detection," in *IEEE Access*, vol. 11, pp. 87297-87304, 2023, doi: 10.1109/ACCESS.2023.3304237.

[4] Ghannad, N.; Passi, K. Detecting Image Forgery over Social Media Using U-NET with Grasshopper Optimization. *Algorithms* 2023, *16*, 399.

[5] Qazi, E.U.H.; Zia, T.; Almorjan, A. Deep Learning-Based Digital Image Forgery Detection System. *Appl. Sci.* 2022, *12*, 2851.

[6] Diwan and A. K. Roy, "CNN-Keypoint Based Two-Stage Hybrid Approach for Copy-Move Forgery Detection," in *IEEE Access*, vol. 12, pp. 43809-43826, 2024, doi: 10.1109/ACCESS.2024.3380460.