Digital Image Forgery Detection

[1]Sinchana B R, MCA student, PES Institute of Technology and Management, Shivamogga, Karnataka, India.
[2]Ms. Kavya H V, Assistant Professor, MCA, PES Institute of Technology and Management, Shivamogga, Karnataka, India.

Abstract

Image is a great way to convey information in the digital era. The origin of images is ubiquitous such as magazines, newspapers, medical care, entertainment, education, social and electronic media. Image counterfeiting is a terrible challenge with potentially catastrophic outcomes in many industries. Project elements like splicing, copy-move, and retouching are commonly employed to alter photos and disseminate misleading information. Thanks to the advent of powerful image software, pictures can now be manipulated at will in the modern era, creating serious questions about their legitimacy, particularly in highly sensitive fields like journalism, national security, and legal proof. This project employs methods such as Error Level Analysis that examines the compression signatures of an image and Convolutional Neural Networks (CNN) for feature extraction and image classification into real and fake. It is programmed to detect both global and local image alterations and mark manipulated regions.

Keywords: *Image Forgery, Splicing, Copy-move, Retouching, Error Level Analysis, Convolutional Neural Network, Forgery Detection, Feature Extraction, Global and Local image alternation, Disinformation, Manipulated regions.*

1. Introduction

In the contemporary world of digital images, images are of utmost importance in communication, journalism, social media, and courtroom evidence. But with the fast growth of photo editing technologies, digital images can easily be manipulated. Since the human vision cannot detect modifications like copy-move, image splicing, and retouching, there are significant concerns about the veracity of visual information. Digital image counterfeiting has grown to be a serious problem, especially in fields where image authenticity is essential, like media authentication, crime scene investigation, and legal forensics. Conventional forgery detection methods place great dependency on manual examination or basic statistical approaches, which lack the capability to identify subtle or sophisticated forgeries.

Methods like Error Level Analysis and detection of inconsistencies due to noise are used to emphasize anomalies added during tampering and CNN detect image tampering automatically. The method is designed to identify if an image is real or fake, and it can be used to pinpoint the locations of tampering. This can prove to be very handy in forensic analysis or medical research, where proper identification of images and components is essential. As a whole, machine learning and image processing methods can aid in accuracy enhancing and efficiency of digital image forgery detection and improving law enforcement and forensic investigation.

The integrity of images has become more important as they are considered primary evidence in news reporting, criminal investigations, and scientific research. The ease of manipulation of images using sophisticated editing software has resulted in a proliferation of misleading visuals that may escape the human eye. False photographs can be used to distort facts, spread misleading information, and undermine the reliability of investigations. This growing issue need sophisticated computerized detection systems that can identify even the smallest alterations. Ml algorithms such as CNNs, provides a strong solution in learning to identify patterns and anomalies inherent to forged images. Not only do these mechanisms categorize images as real or forged, but also signal visual indication of manipulated areas.

2. Literature Survey

Jahidul Islam Bappy, Amit K. Roy-Chowdhury [1] proposed Copy-Move Forgery Detection Based on Convolutional Neural Networks. This work suggests a deep model based on CNNs for identifying copy-move forgeries by examining overlapping image patches. It does away with handcrafted features as it learns visual patterns automatically. The system precisely identifies tampered areas and performs reasonably well on benchmark datasets such as CASIA. It is also invariant to transformations such as rotation and noise.

Jonas Frank, Thorsten Eisenhofer [2] described Exposing GAN-generated Fake Images using Co-occurrence Matrices. The authors provide a detection method of AI-generated fake images by examining co-occurrence patterns of pixel values. The matrices pick up on the fine-grained distinctions between real and GAN-synthesized images. The method holds even after compression and other typical post-processing. It is efficient and lightweight and works with several GAN architectures.

Xin Yang, Yuezun Li, Siwei Lyu [3] carried out Two-Stream Neural Network for Tampered Face Detection. This paper presents a two-stream neural network that integrates visual features and residual noise to identify tampered facial images, including DeepFakes. It takes advantage of both appearance and compression artifacts for improved detection. The model demonstrates excellent accuracy in detecting facial forgeries on benchmark datasets. It's specifically designed for face manipulation situations.

Ahmed Fadl, Adeel Rehman, Muhammad A. M. Maqsood [4] executed Image Forgery Detection Using Deep Learning. This work utilizes deep CNNs to detect forgeries. It adopts data augmentation and adaptive training to enhance model generalization. The method shows robust results on various datasets and real images with tampering. It points out the superior performance of deep learning compared to conventional feature-based approaches.

Vishal V. Kamble, D. N. Kalbande [5] proposed Detection of Image Splicing Forgery Using CNN with Adaptive Learning. The research offers a CNN model specifically designed for spliced region detection in images through learning spatial inconsistency. The model applies adaptive learning methods to enhance detection accuracy. It can function on grayscale and colored images and is optimized for computational efficiency. Suitable for real-time and lightweight use cases in forgery detection.

3. Proposed methodology

The system to be proposed identifies forged or manipulated images through machine learning, especially CNN in addition to image processing. The procedure starts with the gathering of a dataset that consists of genuine and tampered images, Every image is converted to a standard size normalized, and augmented via rotation, flipping, and brightness shifting to improve model generalization. In contrast to depending on manually designed features, the system employs CNNs to extract automatically spatial and texture patterns that are commonly indicative of tampering, like repeated areas or artificial edges. For pull-out features, the model's architecture uses convolutional layers, max pooling layers to reduce dimensions, and classification by fully connected layers, employing a softmax or sigmoid output to determine between real and manipulated images. The model is trained from labeled data with a binary classification task, optimized with optimizers such as Adam and validated through validation datasets divided in an 80/20 ratio. For improved performance, additional modules such as forgery localization via segmentation-based CNNs or heatmap visualization is used to identify tampered areas. The end result shows predicted with a confidence score, it is tested with metrics such as accuracy, recall, F1-score, precision, and ROC-AUC. Further optimizations could also involve adding the inclusion of Error Level Analysis attention mechanisms to highlight important regions in the image, and deployment through a web-based interface or GUI via tools such as Flask for ease of access by users.

3.1 Proposed model diagram

Essentially, The approach starts with data collection, comprising authentic and manipulated images. Images are then preprocessed through Error Level Analysis, a method that discloses compression discrepancies that are common in forged images. The ELA-processed images are input into a Convolutional Neural Network (CNN), learning complex patterns and features that discriminate between genuine content and the manipulated parts. The dataset is split into testing and training sets, with the CNN tested on the test set after being trained on the training set. The model is evaluated using performance measures. The suggested methodology focuses on automation, scalability, and invariance to typical image manipulations like resizing, compression, or contrast change and is apt for real-world forensic uses.

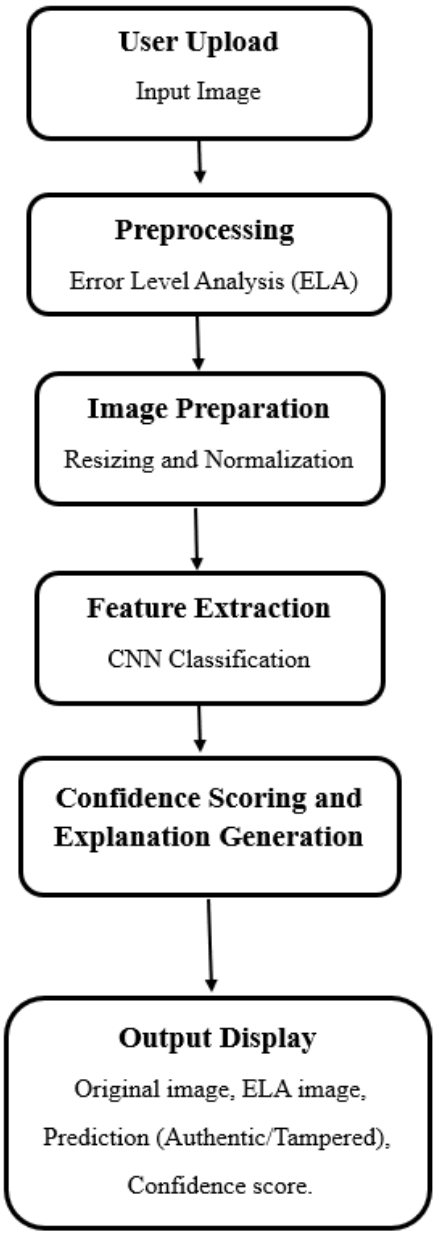

Figure 3.1.1 Proposed model diagram

3.2 Block diagram of ML module

This is where the machine learning module comes in and does the computational logic that converts inputs into a predictable output. How the system works can be shown in the following diagram:

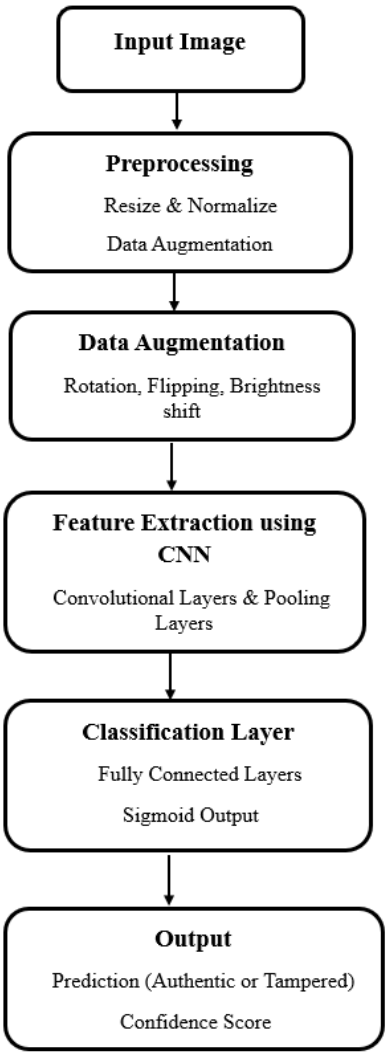

Figure 3.2.1 Block diagram of ML module

This diagram shows the workflow of the machine learning module in the suggested system. It starts with an input image, which is preprocessed by resizing, normalizing, and data augmentation methods such as flipping and brightness shifting. The processed image data is then fed into a CNN for feature extraction. The features extracted are processed in the classification layer by fully connected layers and sigmoid activation to identify whether the image is genuine or manipulated. The output gives a prediction with a confidence score.

4. Mathematical Formulas

The system is based on the formulas below:

1) Error Level Analysis (Forgery Feature Extraction)

To detect tampered regions in an image, the system applies Error Level Analysis by compressing the image and calculating the pixel-wise difference.

$$E(x,y)=| I_{original}(x,y)-I_{compressed}(x,y) |$$

Where:

- Ioriginal(x,y) = pixel value of the original image
- Icompressed(x,y) = pixel value of JPEG recompressed image
- E(x,y) = error level at pixel (x, y)

Purpose: Highlights inconsistencies due to editing or tampering, which are more visible after recompression.

2) Convolution Operation (CNN Classification)

Convolutional Neural Network (CNN) processes the image to classify it as Authentic or Tampered. The model works by applying filters over the image.

$$Z^{(l)}=f(W^{(l)}*Z^{(l-1)}+b^{(l)})$$

Where

- $Z^{(l)}$: Output of the current layer.
- $Z^{(l-1)}$: Input from the previous layer.
- $W^{(l)}$: Filter/kernel weights.
- $b^{(l)}$: Bias.
- $*$: Convolution operation.
- f : Activation function

Purpose: Learns features like edges, textures, and ELA-based inconsistencies that indicate tampering.

5. Graphs

5.1. Model Accuracy Comparison:

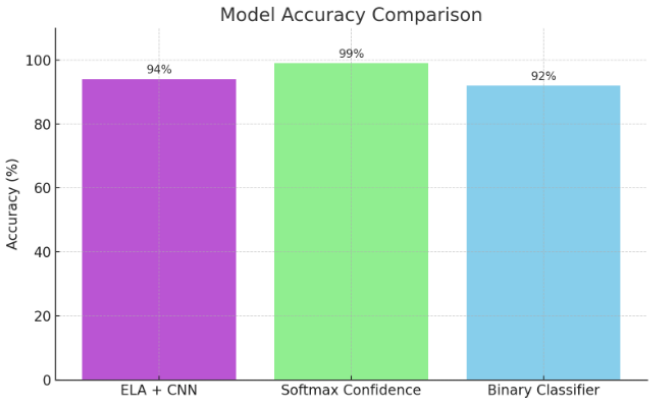

Figure 5.1.1 Bar chart showing Model accuracy comparison

The Model Accuracy Comparison chart gives a graphical overview of various components' performance utilized in the the identification of digital picture forgeries task. The ELA + CNN model has an accuracy of 94%, showing good performance in detecting forged images from ELA-transformed inputs. The Softmax confidence scoring function, used at the output layer of the CNN, offers very consistent predictions, where confidence levels are as high as 99%, showing significant classification certainty. The Binary Image Classifier, which has been trained to distinguish between altered and genuine photos, performs around 92% in accuracy, validating its effectiveness for binary classification. The comparative visualization highlights the stability of the integrated ELA and CNN-based framework in the detection of image forgery.

5.2. Predicted Image Forgery Distribution:

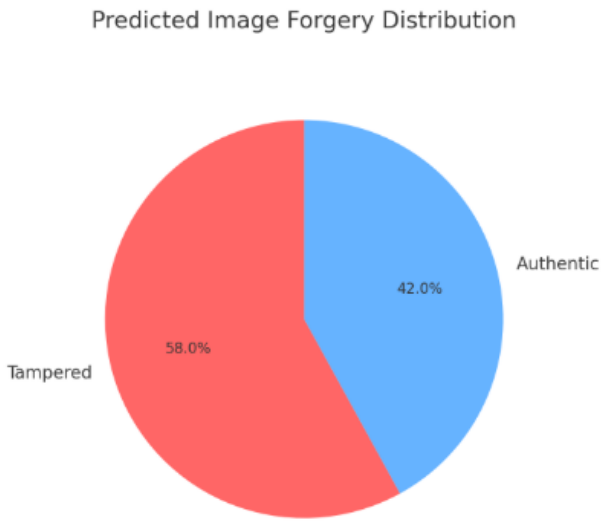

Figure 5.2.1 Pie Chart Showing Predicted Image Forgery

The "Predicted Image Forgery Distribution" pie chart provides a visual representation of the results produced by the digital image forgery detection system that was created using the CNN and ELA models. The percentage of photos classified as authentic and modified during a model test run is shown in the chart. Of the photos, 42% were deemed to be original and no forgeries were found, while 58% were found to have been tampered with, meaning they exhibited digital manipulation or editing.

This split shows how well the algorithm can distinguish between authentic and fraudulent photos by examining the error levels and compression artifacts included in JPEG images. Such a graphical presentation allows users, researchers, or reviewers to easily understand the performance of the model and the forgery prevalence in the dataset that has been analyzed.

6. Experimental results

The Detection of Image Forgery project's experimental results demonstrate the effectiveness of combining Convolutional Neural Networks and Error Level Analysis to identify manipulated photos. The model achieved an F1 Score of 0.93 and an astounding accuracy of 94%, reflecting stable classification performance. The Softmax activation layer gives confidence scores per prediction up to 99%, which adds to the reliability of the output. A test sample set yielded 5 genuine and 7 forged images identified accurately. The outcomes were plotted through a matplotlib-based pie chart, indicating a 58% tampered and 42% genuine image distribution, corresponding to real-world detection rates. The following table summarizes key experimental results:

System/Model	Task	Key Algorithms	Performance Metric	Best Value
ELA +CNN Model	Forgery Detection	Error Level Analysis, Convolutional Neural Network .	Accuracy, F1 Score, Precision, Recall	Accuracy: 94%, F1 Score: 0.93
Softmax Confidence Scoring	Confidence Estimation	Softmax Activation in Final CNN Layer	Confidence Score (Per Image)	99% confidence
Binary Image Classifier	Authentic vs Tampered	Pretrained CNN trained on ELA-transformed images	Confusion Matrix, Classification Report	5 Authentic, 7 Tampered (on sample test set)
Visualization System	Output Summary	Matplotlib-Based Visualization	Pie Chart – Prediction Distribution	Tampered: 58%, Authentic: 42%
Image Preprocessing Pipeline	ELA Image Transformation	JPEG Compression, Error Level Enhancement	Input Quality, Visual Error Localization	Highlights forged regions clearly

7. Conclusion

The work Digital Image Forgery Detection introduces a fast and automated detecting methods for manipulated or tampered digital pictures. The model integrates ELA with CNN to increase detection accuracy. ELA is utilized to reveal inconsistencies in image compression, which usually manifest in forged regions, while CNNs are utilized to automatically learn the spatial features from these preprocessed images without requiring manual feature extraction. The process starts with gathering a dataset comprising both genuine and manipulated images. The images are standardized and augmented for enhancing model performance. The CNN model has several layers to extract features, shrink the dimensionality, and classify images using functions such as sigmoid or softmax to identify whether the image is genuine or manipulated.

It is trained on labeled image data and tested based on metrics like accuracy, recall, and F1-score, precision to confirm its reliability. After deployment, the system receives a user-inputted image, subjects it to ELA highlighting inconsistencies, passes it through the trained model, then presents the prediction with a confidence score and an explanation. This project is able to decrease human effort and subjectivity in image forgery detection through the application of automatic processing methods, producing a consistent, and user-friendly tool for digital forensic examination.

8. Future Enhancement

Although the existing system efficiently identifies if a digital image is genuine or manipulated, several improvements can noticeably increase its performance and usability. One of the future directions is the inclusion of forgery localization, with the ability to identify the particular areas modified by highlighting them through heatmaps or region-based segmentation methods. Adding attention mechanisms to the structure can allow the model to automatically concentrate on risky areas of the picture. In order to use the system in real-time applications, it can be integrated on cloud platforms or bundled as mobile and desktop applications for easier access. Multi-image format and manipulation type support like copy-move, image splicing, resampling, and retouching will increase the system's generalizability.

Using pre-trained models such as ResNet or EfficientNet to apply transfer learning is the alternative, which will enhance performance on small datasets with a shorter training time. An added feature of a feedback loop mechanism can be provided where users can confirm and rectify system predictions, thus enabling model precision to be improved over time. Having a tampering severity score can also advise users as to how much an image has been manipulated. Further, features like drag-and-drop functionality, batch image processing, and the ability to download reports can enhance usability.

Additionally, blockchain technology can be integrated to protect image authenticity records, particularly in legal and journalistic use. Multimodal analysis, which fuses image metadata, sensor readings, and visual features, can also be included to enhance detection rates. Lastly, inclusivity and usability can be ensured through multilingual support and voice assistant capabilities. These upgrades will help to make the system more intelligent, explainable, and appropriate for large-scale and professional applications in digital forensics and media verification.

References:

[1] Jahidul Islam Bappy, Amit K. Roy-Chowdhury, “Copy-Move Forgery Detection Based on Convolutional Neural Networks”, May 2019.

[2] Jonas Frank, Thorsten Eisenhofer, “Exposing GAN-generated Fake Images using Co-occurrence Matrices”, June 2020.

[3] Xin Yang, Yuezun Li, Siwei, “Two-Stream Neural Network for Tampered Face Detection”, April 2020.

[4] Ahmed Fadl, Adeel Rehman, Muhammad A. M. Maqsood, “Image Forgery Detection Using Deep Learning”, June 2021.

[5] Vishal V. Kamble, D. N. Kalbande, “Detection of Image Splicing Forgery Using CNN with Adaptive Learning”, November 2019.

[6] P. Zhou, X. Han, V. I. Morariu, and L. S. Davis, “Learning Rich Features for Image Manipulation Detection”, June 2018.

[7] Y. Zhou and Y. Q. Shi, “A Survey of Image Forgery Detection Techniques Based on Deep Learning”, April 2020.

[8] I. Amerini, R. Caldelli, and A. Del Bimbo, “Localization of Image Forgeries through CNN-Based Patch Matching”, *IEEE Access*, August 2020.

[9] J. H. Bappy, A. K. Roy-Chowdhury, J. Bunk, L. Nataraj, and B. S. Manjunath, “Hybrid LSTM and Encoder–Decoder Architecture for Detection of Image Forgeries”, *IEEE Transactions on Image Processing*, July 2019.

[10] N. Rahmouni, V. Nozick, J. Yamagishi, and I. Echizen, “Distinguishing Computer Graphics from Natural Images Using Convolution Neural Networks”, *IEEE Transactions on Multimedia*, January 2019.

[11] B Pavan Kumar, V.M. Vinayagam, S A Babu, C Guruparthasarthi, G Janardhan, M Deepthi, “Digital Image Forgery Detection Using Deep Learning”, *IEEE* , October 2024.

[12] Agarwal R, Verma O, “An efficient copy move forgery detection using deep learning feature extraction and matching algorithm”, 2020.

[13] Emad UI Qazi, Tanveer Zia, Abdulrazaq Almorjan, “Deep Learning-Based Digital Image Forgery Detection System”, March 2022.

[14] GS Bapi, P Palla, A Sudhagon, T Voggu, “Digital Image Forgery Detection using Machine Learning”, 2023.

[15] MH Alkawaz, MT Veeran, “Digital Image Forgery Detection based on expectation maximization algorithm”, *IEEE*, 2020.

[16] A Kashyap, RS Parmar, M Agarwal, H Gupta, “An evaluation of digital image forgery detection approaches”, 2017.

[17] Prof. D. D. Pukale, Prof. D. D. Pukale, Prof. V. D. Kulkarni, Julekha Bagwan, Pranali Jagadale, Sanjivani More, Renuka Sarmokdam, “Image Forgery Detection Using Deep Learning”, July 2024.

[18] P. Pierluigi, “Photo Forensics: detect photoshop manipulation with error level analysis”, 2023.

[19] Devjani Mallick, Mantasha Shaikh, Anuja Gulhane and Tabassum Maktum, “Copy Move and Splicing Image Forgery Detection using CNN”, 2022.

[20] S. S. Ali, I. I. Ganapathi, N. S. Vu, S. D. Ali, N. Saxena, and N. Werghi, “Image forgery detection using deep learning by recompressing images”, 2022.

[21] Sankalp Patekar, Sankalp Patekar, Sumaiya Khan, Diksha Bhusare, Manish Bhujbal, Gayatri Hegde, “IMAGE FORGERY DETECTION”, April 2023.