YUVARAJ
2022504554

AI Image Forgery Detection

Abstract. Image forgery detection has become an essential task in the modern digital age, especially with the proliferation of AI-generated media that can closely mimic authentic content. This project proposes a deep learning-based approach for detecting forged images using **Error Level Analysis (ELA)** and a **Convolutional Neural Network (CNN)**. ELA is a forensic technique that highlights the discrepancies introduced during image manipulation, allowing for the identification of subtle differences often invisible to the human eye. In this system, images are preprocessed through ELA and then fed into a ResNet50-based neural network for classification. The model distinguishes between authentic and AI-generated (fake) images with high accuracy and provides a confidence score. A user-friendly web interface, developed using **Streamlit**, allows users to upload images and receive real-time verification results. While the initial dataset includes a small number of real and fake images, the methodology sets the foundation for scalable, real-world applications. Despite its limitations in dataset size and generalization, this project demonstrates the potential of combining image forensics with deep learning to build reliable, automated image authentication systems.

Keywords. *Image Forgery Detection, Error Level Analysis (ELA), Deep Learning, ResNet50, Python, Streamlit, Convolutional Neural Network (CNN), AI-Generated Images*

1.Introduction

Image authenticity verification has become increasingly important in the digital era, particularly with the emergence of highly realistic AI-generated content, such as GAN-based images and deepfakes. These forged images are now almost indistinguishable from real ones to the human eye, posing significant challenges for content verification and digital forensics. To address this, artificial intelligence (AI) and deep learning techniques have been employed to automatically detect manipulated images.

This project focuses on detecting forged images using **Error Level Analysis (ELA)** as a preprocessing technique, combined with a **ResNet50-based Convolutional Neural Network (CNN)** for classification. ELA works by exposing compression artifacts introduced during image editing, making subtle manipulations visually

detectable when amplified. The CNN then learns these patterns and predicts whether the image is authentic or fake.
Several related works have explored this domain:

[1] Image Forgery Detection Using ELA and Deep Learning – Chakraborty, A., Rathi, R., & Sahu, A. (2024) – This paper proposes a dual-branch CNN architecture using Error Level Analysis (ELA) and noise residuals for tampering detection. The model achieves over 98% accuracy on the CASIA dataset.
[2] ELA-CNN Based Model for Image Forgery Detection – Kubal, A., Mane, R., & Pulgam, P. (2023) – The authors design an integrated model combining ELA preprocessing with a CNN classifier. The study emphasizes JPEG compression inconsistencies and achieves 92.1% accuracy.
[3] Detecting Image Manipulation with Error-Level and Deep Learning Techniques – Nagm, E., et al. (2024) – A lightweight CNN trained on ELA images is proposed, achieving 99% training accuracy and 94% testing accuracy. Tested on CASIA-2 and public datasets.
[4] Enhancing Image Forgery Detection with CNN and Forensic Preprocessing – More, R., Pawar, P., & Bhoi, V. (2025) – Combines forensic ELA processing with convolutional architectures to detect tampered images in real-time settings, achieving ~94% accuracy.
[5] A Survey on Deep Learning for Image Forgery Detection – He, Z., Dong, L., Yi, S., & Wang, F. (2020) – Surveys CNN-based methods using frequency and ELA domains. Emphasizes generalization challenges and the role of data augmentation.
[6] Boundary-Based Forgery Localization Using Shallow CNN – Zhang, Y., & Cao, X. (2018) – Proposes a fast shallow CNN to localize forged boundaries using compression cues. Effective for detecting small-scale edits in compressed images.
[7] D-UNet: Dual Encoder UNet for Splicing Forgery Detection – Liu, J., & Shi, Y. (2020) – Introduces a deep dual-encoder U-Net architecture optimized for detecting spliced image regions under multiple manipulation formats.
[8] Learning Residual Features for Forensic Analysis via CNNs – Cozzolino, D., Poggi, G., & Verdoliva, L. (2017) – Recasts traditional residual-based descriptors into a CNN-compatible format for robust manipulation classification.
[9] Double JPEG Compression Detection via Deep Convolutional Neural Networks – Amerini, I., et al. (2017) – Focuses on detecting double-compression traces through deep CNNs, aiding in forgery region identification and quality analysis.
[10] Evaluating ELA Compression Factors for Deep Forgery Detection – Sari, D., & Fahmi, H. (2021) – Compares various JPEG quality factors in ELA and their effect on CNN accuracy. A 2.7% accuracy improvement is noted for mid-range compression.

.

2.Methodology

The methodology for image forgery detection using AI in Python involves several key steps, from data preprocessing and ELA conversion to model training and deployment. Below is the detailed workflow for developing this system using Error Level Analysis (ELA) and Convolutional Neural Networks (CNNs).

2.1 Data Collection and Preprocessing

Data Collection:

- A dataset of real and AI-generated (fake) images is curated manually.
- Real images are captured or collected from authentic sources.
- Fake images are generated using AI tools or sourced from public datasets with synthetic content.

Image Preprocessing:

- **Resizing:** All images are resized to 224×224 pixels for model compatibility.
- **Error Level Analysis (ELA):**
  – Each image is resaved with a slight compression (JPEG quality = 90).
  – The difference between the original and compressed image is computed.
  – Contrast enhancement is applied to highlight forgery artifacts.
- **Normalization:** ELA images are normalized to a [0,1] pixel range for efficient model training.
- **ELA Image Storage:** Processed ELA images are saved in a separate folder for training and testing.

2.2 Feature Extraction

ELA-Based Feature Generation:

- Forgery patterns are amplified using ELA, highlighting tampered regions based on recompression inconsistencies.

- These enhanced regions serve as implicit feature maps for training the CNN..

Deep Learning Feature Extraction

- A pre-trained ResNet50 network (without top layers) is used to automatically extract hierarchical visual features.
- GlobalAveragePooling is applied to condense spatial features into a flat vector representation.

2.3 Model Selection

- **Deep Learning Models**:
  Convolutional Neural Network (CNN):

– ResNet50 is selected due to its strong feature extraction capability and robustness.
– A custom dense layer with sigmoid activation is added for binary classification (real vs fake).
– This transfer learning approach allows effective training on small datasets.

2.4 Model Training

Training:

- The dataset is split into training and testing sets (80/20 split).
- Binary labels are assigned (0 = real, 1 = fake), inferred from image filenames.
- Loss Function: Binary Cross-Entropy.
- Optimizer: Adam with a learning rate of 1e-4.
- Epochs: The model is trained for 5 epochs.

Supervised Learning:

- The CNN learns to map visual patterns in ELA images to authenticity labels.
- The trained model is saved as ela_model.h5 for deployment.

2.5 Forgery Detection

- **Input Processing:** Uploaded images are converted into ELA format in real time.
- **Model Inference:** The preprocessed ELA image is passed through the trained CNN.
- **Output:**
  - The model predicts a value between 0 and 1.
  - Values > 0.5 are classified as forged, while ≤ 0.5 are real.
- **Confidence Score:** The prediction value is used to display confidence (e.g., 0.83 = 83% sure the image is fake).

2.6 Evaluation and Optimization

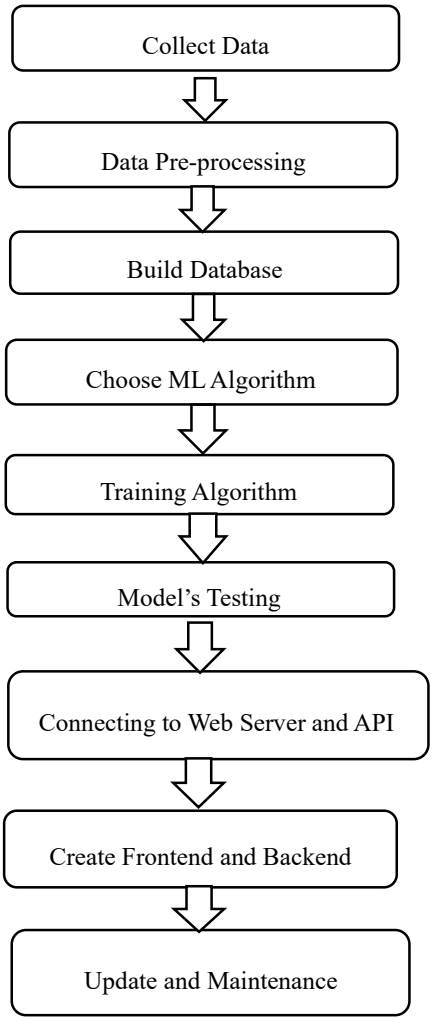
Evaluation Metrics:

- **Accuracy**: Percentage of correctly classified images.
- **Precision and Recall**: Particularly important when minimizing false alarms in real/fake classification.
- **Confusion Matrix:** Visual breakdown of true/false predictions.

Optimization:

- **Data Augmentation**: More fake and real samples can be added (e.g., rotation, lighting changes) to enhance robustness.
- **Model Fine-Tuning**: Additional training epochs or freezing/unfreezing deeper ResNet layers can be explored.
- **Cross-Validation:** Helps in mitigating overfitting, especially with small datasets.

2.7 Flowchart



3.Results and Discussions

The proposed AI-based image forgery detection system demonstrates effective performance in identifying manipulated images using a combination of Error Level Analysis (ELA) and deep learning. The ResNet50-based model successfully learns visual inconsistencies introduced during image tampering and provides accurate classification between real and forged images.

The ELA preprocessing enhances pixel-level differences, enabling the model to detect subtle alterations that may not be visible to the naked eye. Real-time results are displayed through a user-friendly Streamlit interface, offering predictions along with confidence scores.

While the system shows promising results, further enhancements can include expanding the range of forgery types, improving model robustness, and integrating additional forensic features such as metadata or frequency-domain analysis to improve reliability across diverse image sources.

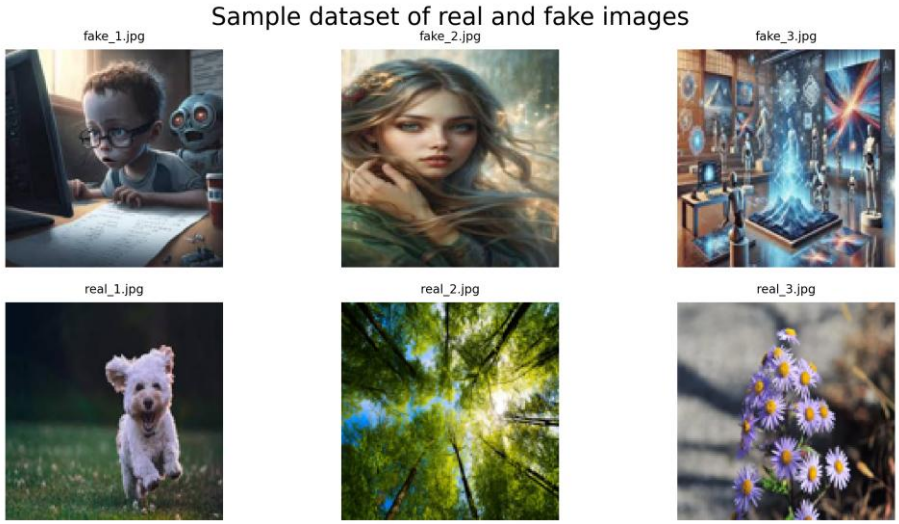


Fig. 1. Sample dataset of real and fake images

Fig. 2. Original Input Image Uploaded by the User

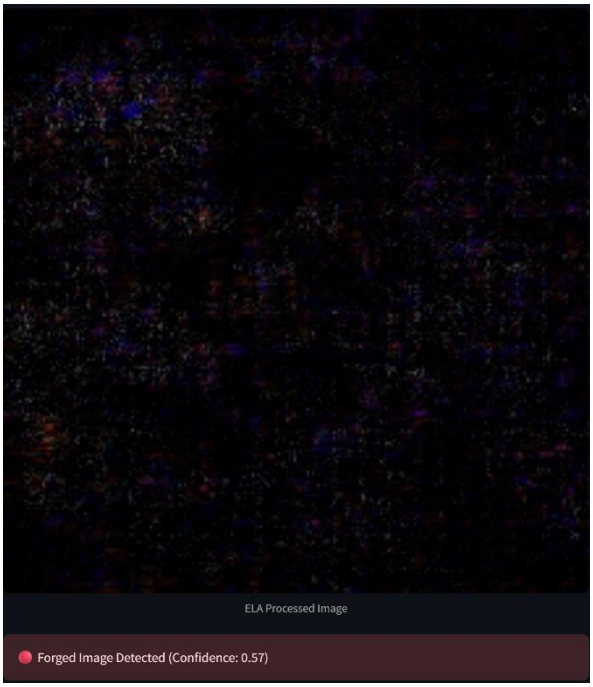


Fig. 3. ELA-processed image and the corresponding forgery detection result with confidence score.

This program takes an uploaded image, processes it using Error Level Analysis (ELA), and passes the result to a trained deep learning model. The model analyzes the image and predicts whether it is real or forged, displaying the classification result along with a confidence score.

4. Helpful Hints

- Upload images in .jpg, .jpeg, or .png format for best compatibility.
- The system performs better on original or lightly compressed images; avoid those heavily edited or shared through messaging apps.
- JPEG images are preferred, as Error Level Analysis (ELA) relies on compression artifacts to highlight tampering.
- Ensure uploaded images are clear and not blurry, with visible objects or faces.
- For optimal results, use images with standard resolution (e.g., 224×224 or larger).
- This tool is built as a proof of concept; model performance can be enhanced by using a larger and more diverse dataset.
- Retraining the model periodically with new image samples helps maintain accuracy and relevance.

5.Future Scope

- **Dataset Expansion and Diversity**: The system can be significantly improved by using larger and more diverse datasets containing various types of forgeries including deepfakes, splicing, copy-move, and AI-generated images. This would improve the model's generalization to real-world tampered content.

- **Advanced Forgery Detection Techniques**: Future models can move beyond Error Level Analysis (ELA) and incorporate additional image forensic techniques such as chromatic aberration analysis, sensor noise patterns, and frequency-domain detection to improve detection accuracy.

- **Temporal and Sequential Media Verification**: Forgery detection can be extended to videos and sequential image frames, allowing the system to identify inconsistencies across time and detect frame-level tampering in news clips, surveillance footage, or deepfake videos.

- **Integration into Content Platforms**: This system can be embedded into social media or content publishing platforms to automatically verify the authenticity of uploaded images, thereby preventing the spread of manipulated or misleading content.

- **Lightweight Deployment for Real-Time Use** With model optimization techniques like pruning, quantization, or conversion to TensorFlow Lite, the detection system can be deployed on low-resource or mobile devices for on-the-fly verification.

In conclusion, the future of AI-based image forgery detection lies in enhancing detection robustness, enabling real-time deployment, and scaling to a wide variety of tampering techniques. With continued advancements in deep learning and forensic analysis, such systems will play a vital role in securing digital content integrity.

6.Conclusion

AI-based image forgery detection has significantly improved the ability to authenticate visual content in a world increasingly affected by synthetic media. By leveraging techniques like Error Level Analysis (ELA) and deep learning with Convolutional Neural Networks, this project demonstrates an effective pipeline for detecting manipulated images. The system provides real-time predictions through a simple and interactive interface, making it both practical and accessible. While challenges such as generalization to diverse forgery types and dependency on input quality remain, the continuous evolution of AI models and forensic techniques will drive the development of more reliable and adaptable forgery detection systems. Future improvements, including larger datasets, multi-modal analysis, and lightweight deployment, will further enhance the system's applicability across fields such as journalism, legal forensics, social media verification, and digital content protection.

7. Acknowledgement

I would like to express my sincere gratitude for the opportunity to work independently on this project, which allowed me to explore and apply advanced concepts in image forensics and deep learning. Undertaking the development of an AI-based image forgery detection system has been a valuable learning experience that enhanced my understanding of both theoretical and practical aspects of artificial intelligence.

This project was self-initiated and executed entirely on my own, from dataset preparation and model training to interface development and testing. I am thankful for the availability of open-source tools, libraries, and community forums that provided support and guidance throughout the process

8. References

[1] A. Chakraborty, R. Rathi, and A. Sahu, "Detection of Image Tampering Using Deep Learning, Error Levels and Noise Residuals," *Neural Processing Letters*, Springer, 2024.

[2] A. Kubal, R. Mane, and P. Pulgam, "EACN: Error Analysis and Convolutional Neural Network for Image Forgery," *International Journal of Intelligent Systems and Applications in Engineering (IJISAE)*, vol. 11, no. 2, 2023.

[3] E. Nagm et al., "Detecting Image Manipulation with Error Level and Deep Learning Techniques," *PeerJ Computer Science*, 2024.

[4] Z. He, L. Dong, S. Yi, and F. Wang, "A Survey on Deep Learning for Image Forgery Detection," *Scientific Research Publishing*, 2020.

[5] R. More, P. Pawar, and V. Bhoi, "Enhancing Image Forgery Detection with Convolutional Neural Networks and Error Level Analysis," *ResearchGate*, 2025.