Image Forgery Detection

Submitted for

Artificial Intelligence and Machine Learning CSET301

Submitted by:

(E23CSEU2388) HARSHIT CHAUDHARY

(E23CSEU2389) HARSH RAJ

(E23CSEU2393) ARYAN MITTAL

Submitted to

DR. NITIN ARVIND SHELKE

Jan-May 2025

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

INDEX

1. Abstract

Image forgery has become a critical issue in the digital age, where tampered images can easily spread misinformation. This project presents a novel approach for detecting image forgeries by combining **Error Level Analysis (ELA)** with **Convolutional Neural Networks (CNNs)**. ELA enhances the visibility of tampered regions by highlighting compression artifacts, which are then processed by a CNN to classify images as forged or authentic. Using the **CASIA 2.0 dataset**, our model achieved an accuracy of **90.3%**, demonstrating its effectiveness in identifying subtle manipulations. The proposed solution is scalable, adaptable to real-world scenarios, and serves as a foundation for real-time and DeepFake detection.

2. Introduction

In recent years, the rise of image editing tools and AI-generated media has made image forgery detection more important than ever. Fake images can mislead public opinion, compromise legal evidence, and propagate disinformation. Traditional forgery detection methods struggle with complex manipulations. Hence, there is a need for automated, intelligent systems that can detect forged images with high precision.

This project introduces an innovative fusion of **Error Level Analysis (ELA)** and **Convolutional Neural Networks (CNNs)** to automatically detect forgeries. ELA reveals areas of varying compression, a common byproduct of image tampering, while CNNs are adept at learning intricate visual patterns from data. Together, they form a powerful pipeline for forgery detection.

3. Related Work

Numerous techniques have been explored for image forgery detection, including:

- **Passive Detection Methods**: These rely on analyzing inconsistencies in lighting, noise patterns, or metadata.
- **Pixel-Based Techniques**: These inspect differences at the pixel level to identify anomalies in color or texture.
- **Frequency-Domain Techniques**: Such as DCT and wavelet transforms, used for spotting splicing or copy-move forgeries.
- **Machine Learning Approaches**: Some models use SVMs and decision trees, but lack the accuracy of deep learning models.
- **Deep Learning Approaches**: CNNs have recently been used for forgery detection, but combining them with ELA is a relatively novel strategy.

Our work builds on these by preprocessing images with ELA before feeding them to a CNN, improving the accuracy and robustness of the model.

4. Methodology

4.1. Dataset

- **CASIA 2.0 Dataset** from Kaggle
- Contains 12,000+ images with authentic and forged versions
- Includes splicing and copy-move tampering types

4.2. Error Level Analysis (ELA)

- Compresses an image and compares it to the original
- Highlights areas with different compression levels (potential tampering regions)
- Converts subtle artifacts into visual cues for the CNN

4.3. Convolutional Neural Network (CNN) •

Input: ELA-processed image (RGB)

- Layers:
    - Convolution layers to extract spatial features ○ Pooling layers for dimensionality reduction ○ Dropout to prevent overfitting
    - Dense layers for final classification
- Output: Binary classification (Forged / Authentic)

4.4. Training

- Split dataset: 80% training, 20% testing
- Loss Function: Binary Cross-Entropy
- Optimizer: Adam
- Accuracy Achieved: **90.97%**

5. Hardware/Software Required

Hardware:

- CPU: Intel i5 or better (for local development)
- GPU: NVIDIA GPU (Recommended for faster training)
- RAM: 8GB minimum, 16GB preferred

Software:

- **Python 3.x**
- Libraries:
  - TensorFlow / Keras
  - OpenCV
  - NumPy, Pandas, Matplotlib
- Jupyter Notebook / Google Colab for experimentation
- Kaggle Datasets API for dataset access

6. Experimental Results

Accuracy: 90.97% on test data

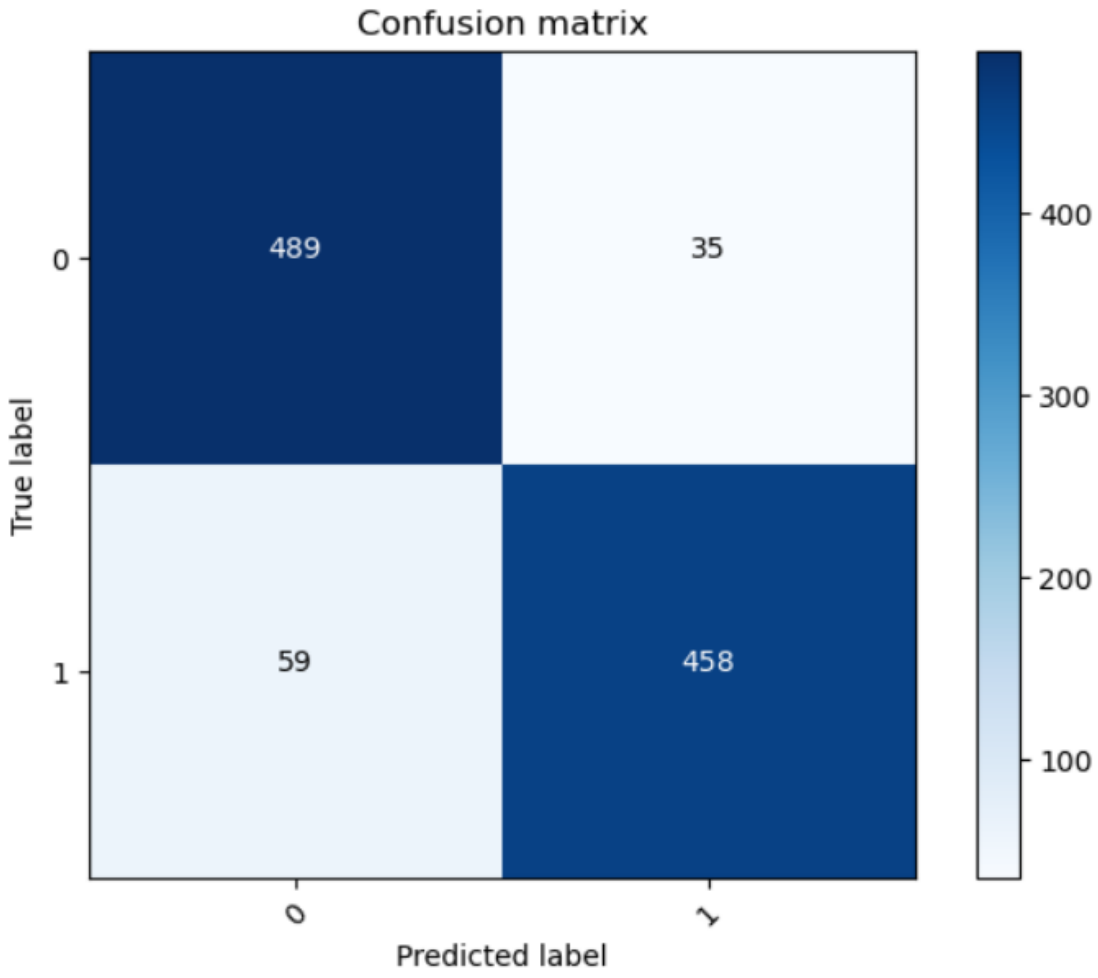
Precision: 92.89%

Recall: 88.66%

F1 Score: 90.72%

Confusion Matrix: Shows balanced performance between true positives and true negatives

Visualizations: ELA images highlight regions of forgery clearly



7. Conclusions

This project successfully demonstrates the potential of combining **Error Level Analysis** with **Convolutional Neural Networks** for image forgery detection. The system achieves high accuracy while maintaining generalizability across various forgery types. ELA enhances the model's ability to detect subtle tampering, making it a valuable pre-processing step. The pipeline proves to be scalable, accurate, and ready for further improvements, showing promise for deployment in media, law enforcement, and digital security applications.

8. Future Scope

- **Real-Time Detection: Optimize the system for live platforms like social media, using lightweight CNN architectures or quantization.**
- **Transfer Learning: Apply pretrained models for faster training and better feature extraction.**
- **DeepFake Detection: Extend the system to include GAN-generated content analysis.**
- **Mobile and Web Deployment: Build cross-platform applications (Android/iOS/Web) for public use.**
- **Explainability: Integrate Grad-CAM or SHAP to explain why a certain region was detected as forged.**

1. GitHub Link of Your Complete Project

https://github.com/LegacyHarsh/Image_Forgery_Detection