Image Forgery Detection Using CNNs

Abstract

The rapid proliferation of digital media has made image forgery a pressing concern. Addressing this, our project integrates the capabilities of Error Level Analysis (ELA) and Convolutional Neural Networks (CNNs) to create a formidable image forgery detection system. By meticulously analysing JPEG compression artifacts and leveraging the prowess of deep learning, our system stands as a beacon of authenticity verification, adeptly differentiating between genuine and manipulated images.

Introduction

In today's digital era, images play a pivotal role in shaping perceptions, narrating stories, and even influencing decision-making processes. However, the ease of access to advanced image editing tools has made image manipulation a commonplace activity. These doctored images, which often go undetected to the naked eye, have the potential to misrepresent facts, leading to misinformation. Such distortions can have profound implications, especially in critical domains like journalism, digital forensics, legal scenarios, and even social media. Recognizing the gravity of this issue, our project embarks on a mission to provide a reliable solution. By synergizing the traditional technique of ELA, which inspects compression inconsistencies, with the state-of-the-art CNNs, known for their image pattern recognition capabilities, we aim to offer a robust and comprehensive approach to detect image forgeries.

Problem Statement

The core challenge lies in developing an efficient and accurate system that can seamlessly differentiate between authentic and tampered images. Given the sophistication of modern image editing tools, forgeries can often be imperceptible, making the detection task even more challenging. Our objective is to harness the subtle inconsistencies introduced during JPEG compression, which often go unnoticed, and combine this with the deep learning capabilities of CNNs. The goal is to create a system that not only identifies manipulations but also stands resilient against evolving forgery techniques.

Data

The Casia dataset from Kaggle, a widely recognized dataset in digital forensics, was employed. It contains a mix of authentic and forged images, providing a balanced dataset for training and validation.

- Total images - 11129
- Authentic images - 8144
- Forged images – 2985

Methodology

CNN Architecture:

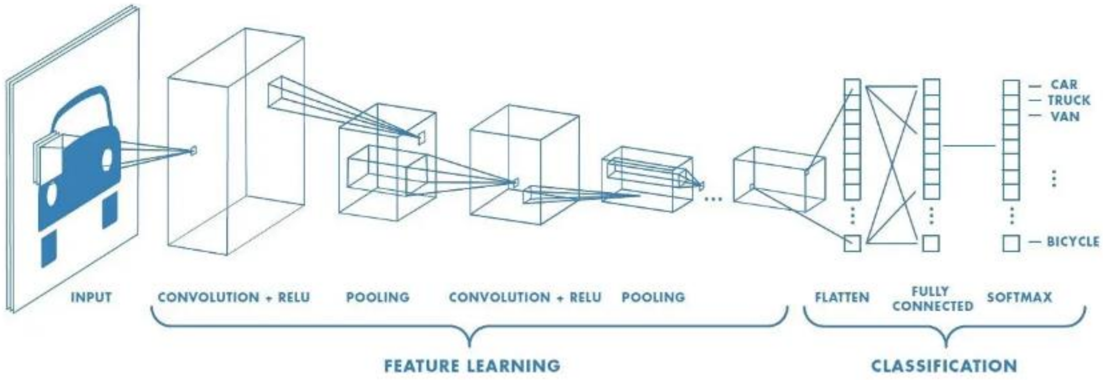
- Input Layer: Accepts ELA-processed images of size 128x128 pixels with 3 colour channels (RGB).

- Convolutional Layers: Multiple layers with 64 filters of size 5x5, using ReLU activation. These layers extract features from the ELA images.

- Pooling Layers: Max-pooling layers with a pool size of 2x2, reducing spatial dimensions and retaining significant features.

- Global Average Pooling: Reduces the spatial dimensions to a flat vector while preserving the depth.

- Output Layer: A dense layer with a single neuron using sigmoid activation for binary classification (authentic or forged).



Error Level Analysis (ELA):

- Principle: ELA highlights inconsistencies in the compression rates of images. Manipulated regions often exhibit different error levels compared to the rest of the image.

- Implementation: The input image is resaved at a predefined JPEG quality. The difference between the original and the resaved image is computed, resulting in the ELA image. This image accentuates discrepancies in compression, revealing potential tampered areas.

Training:

- Optimizer: Adam optimizer with a learning rate decay strategy.

- Loss Function: Binary cross-entropy, suitable for binary classification tasks.

- Early Stopping: Monitors validation accuracy to prevent overfitting, halting training if accuracy doesn't improve over ten epochs.

Dataset Preparation:

- Images undergo ELA processing and are resized to the required input size of 128x128 pixels.

- Normalization is applied to scale pixel values between 0 and 1.

- The dataset is partitioned into training, validation, and testing sets, ensuring a diverse representation of images in each set.

Evaluation Metrics:

- Confusion Matrix: Provides a visual representation of the model's performance, highlighting true positives, true negatives, false positives, and false negatives.

- Classification Report: Offers precision, recall, and F1-score metrics, giving a holistic view of the model's classification capabilities.

Deployment:

- A user-friendly web application is developed using Streamlit. Users can upload images, which are then processed and classified by the trained model in real-time.

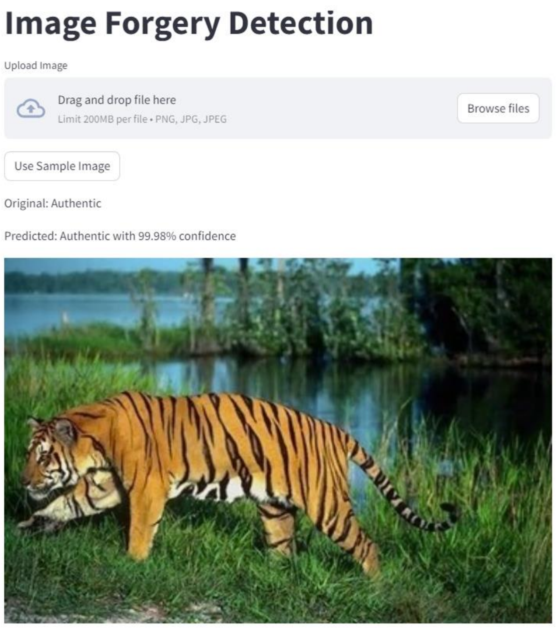


Image Forgery Detection

Original: Authentic

Predicted: Authentic with 99.98% confidence

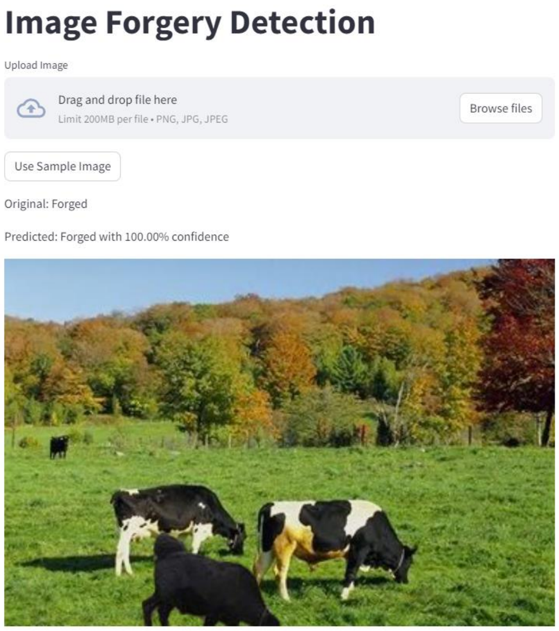


Image Forgery Detection

Original: Forged

Predicted: Forged with 100.00% confidence

Results:

The system's efficacy is evident from its high accuracy on the test dataset. The combination of ELA and CNN ensures that the model captures both subtle and prominent signs of tampering.

Conclusion:

The fusion of traditional image processing techniques like ELA with advanced deep learning models like CNNs offers a promising solution to the challenge of image forgery detection. The system's accuracy and efficiency make it a valuable tool in the digital forensics toolkit.

Future Scope:

- Incorporate Advanced Models: Integrate architectures like ResNet or VGG for enhanced feature extraction.

- Real-time Video Analysis: Extend the system to detect forgeries in videos, addressing deepfakes.

- Transfer Learning: Utilize pre-trained models to improve accuracy and reduce training time.

- User Feedback Mechanism: Implement a feedback loop in the web application, allowing users to correct misclassifications and continuously improve the model.