IJIRSET
International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

e-ISSN: 2319-8753 | p-ISSN: 2347-6710

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 4, April 2024

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.423

9940 572 462 | 6381 907 438 | ijirset@gmail.com | www.ijirset.com

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

| e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 4, April 2024 ||

| DOI:10.15680/IJIRSET.2024.1304237 |

Enhanced Forensic Investigation: CNN and Error Level Analysis for Image Tampering Detection

Mr. M. Hari Prasad[1], B. Deekshith[2], R. Ajay Kumar[3], V. Varish[4]

Assistant Professor, Department of Artificial Intelligence, Anurag University, Hyderabad, India. [1]

U.G. Student, Department of Artificial Intelligence, Anurag University, Hyderabad, India. [2,3,4]

ABSTRACT: The recent years have witnessed a boom in the number of photos taken due to the wide availability of cameras thereby making them an integral part of everyday life. Images, despite being useful, are at times manipulated to promote false information. Although there are traditional methods for detecting such frauds, they do not cope well with emerging cases of manipulation. Convolutional Neural Networks (CNNs) have been a game-changer in detecting tampering attempts; however, most CNN based methods focus on specific types of fakes while leaving out those which are not visible. In order to address this problem we propose to apply a deep learning system using Error Level Analysis (ELA), which can identify variations in compression artifacts between original and re-compressed images quickly thus allowing efficient training. This is faster compared to existing algorithms as shown by experimental results that prove its ability to detect image manipulations in a more effective way. By dealing with problems associated with digital forgery detection our study presents a workable approach towards maintaining genuine authenticity for visual content.

KEYWORDS: CNN, ELA, Forgery Detection.

I. INTRODUCTION

"In today's modern world, it is easy to manipulate digital images using image editing tools. Although these tools provide opportunities for creativity, they also raise concerns about the credibility and authenticity of visual content. Image forgery, which involves changing or tampering with images to trick people, has become a widespread problem with serious consequences in fields like journalism, forensics, and social media.".

The consequences of image forgery are profound. Edited images can be utilized for spreading false information, misleading viewers, or altering proof in legal cases, eroding confidence in online media and risking the dependability of visual content. Therefore, there is a critical requirement for effective methods to identify and counteract image alterations, protecting the genuineness and trustworthiness of digital images.

Detecting image forgery requires identifying signs of tampering in digital images to differentiate between genuine and altered content. Researchers and experts have created various methods and algorithms to address this issue, utilizing traditional image processing methods, statistical analysis, machine learning, and deep learning to detect abnormalities and discrepancies that signal image manipulation. While there have been notable advances in forgery detection technology, the constantly changing landscape of image manipulation necessitates continued research and creativity in this area. Additionally, with the widespread use of social media and online platforms, the need for effective forgery detection methods is more crucial than ever.

Detecting image forgeries requires finding evidence of manipulation in digital images, distinguishing between real and altered content. Researchers and professionals have created a wide range of methods and algorithms to address this issue. These techniques use different approaches such as classic image processing, statistical analysis, machine learning, and deep learning to find signs of manipulation. While forgery detection technology has improved, the changing landscape of image manipulation requires ongoing research and innovation. The rise of social media and digital platforms also presents new challenges in this field.

Categories Of Image Forgery Detection

Detecting image forgery involves determining if an image has been altered in any way. There are two primary types of image forgery detection: Active and Passive. Active detection methods involve adding digital data, like watermarks or

| e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 4, April 2024 ||

| DOI:10.15680/IJIRSET.2024.1304237 |

signatures, to the original image. This added information can be checked later to confirm the image's authenticity. On the other hand, Passive detection methods can identify manipulation without needing any prior knowledge about the original image. Instead, they rely on analyzing the image itself for signs of forgery. These signs can include inconsistencies in the lighting, shadows, or noise patterns in the image.

Image Forgery Types

Copy-Move: This is a technique where a portion of an image is copied and pasted onto another location within the same image. The goal is typically to hide or remove unwanted objects. For instance, someone might copy a patch of sky from another area to cover up a distracting telephone pole.

Splicing: Splicing involves taking a portion of one image and pasting it onto a completely different image. This is used to create composite images that might appear entirely real. For example, someone might splice a person's face onto another person's body.

Inpainting: Inpainting is a method used to fill in missing or damaged parts of an image, rather than trying to deceive viewers. It is commonly used to repair old photos by removing scratches or filling in gaps left by removed objects during editing.

II. LITERATURE SURVEY

There have been many studies done on detecting image forgeries. This review looks into different methods used to spot manipulated images, specifically focusing on passive forgery detection techniques. We review research on copy-move, splicing, and inpainting detection, discussing their pros, cons, and possible directions for future research.

Johnson and team pioneered the application of machine learning in image forgery detection. They utilized feature extraction techniques alongside support vector machines (SVMs) to identify forged regions within digital images. Their method achieved an impressive accuracy of 87% on a diverse dataset of manipulated images, marking a significant advancement in the field.

Wang et al. introduced a groundbreaking approach based on deep learning for image forgery detection. They developed a convolutional neural network (CNN) architecture capable of automatically learning discriminative features to detect various types of image manipulations. Their CNN-based detector outperformed traditional methods, achieving an impressive accuracy of 94% on a benchmark dataset, signifying a notable leap forward in forgery detection technology.

Liu et al. proposed a novel framework based on generative adversarial networks (GANs) for image forgery detection. By training a GAN to generate authentic image patches, their system learned to distinguish between genuine and forged regions within images. The GAN-based detector demonstrated robustness against sophisticated forgery techniques, achieving an accuracy of 92% on a challenging dataset of manipulated images, showcasing a significant breakthrough in forgery detection methodology.

Zhang and colleagues introduced a hybrid approach combining deep learning and traditional image processing techniques for image forgery detection. Their method integrated error level analysis (ELA) with a convolutional neural network (CNN), enabling effective detection of various types of image manipulations. Experimental results demonstrated the efficacy of their approach, achieving an impressive accuracy of 95% on a comprehensive dataset, representing a notable advancement in forgery detection methodology.

Chen et al. proposed a blockchain-enabled image authentication system for combating image forgery. By leveraging blockchain technology to record image metadata and transaction history, their system provided tamper-evident authentication for digital images, ensuring their integrity and authenticity in a decentralized manner. Their contribution addresses the growing concern of digital image manipulation and emphasizes the importance of preserving the authenticity of visual content in the digital age.

III. PROBLEM STATEMENT

In today's digital age, the growing availability of image editing software and widespread use of social media have made image manipulation a common problem. People with ill intentions use these tools to deceive others, spread fake news, and tarnish reputations.

| e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 4, April 2024 ||

| DOI:10.15680/IJIRSET.2024.1304237 |

Traditional methods for detecting image forgeries often fall short in accurately identifying manipulated images, especially with the rise of sophisticated techniques such as deep learning-based manipulation. Therefore, there is an urgent need for robust and reliable image forgery detection systems capable of effectively identifying manipulated images across different types of manipulations and under varying levels of sophistication.

3.1 Existing Systems

Convolutional neural networks, or CNNs, are powerful tools for detecting picture forgeries because of their capacity to automatically learn hierarchical representations of image features that are modeled after the visual system of humans. Reliable image forgery detection is critical in this era of powerful and widely available image manipulation tools. CNNs are particularly good at picking up on complex patterns and characteristics in photos, which makes them useful for spotting minute anomalies and inconsistencies that could be signs of image tampering. CNNs can detect a variety of modification techniques, including copy-move, splicing, and retouching, by learning to distinguish between real and forged regions by training on huge datasets that contain both authentic and modified images.

3.2 Proposed System

Our proposed system tackles forgery detection with a two-pronged approach. First, a Convolutional Neural Network (CNN) analyzes the image to capture complex features and identify potential inconsistencies indicative of manipulation. Secondly, error-level analysis complements the CNN by pinpointing specific regions within the image that exhibit statistical anomalies often associated with forgery. This combined strategy aims to achieve robust and accurate detection of a broad spectrum of image forgeries.

IV. DATASET

The quality and diversity of the dataset used are crucial factors in determining how well the detection algorithms perform in the field of image forgery detection. The CASIA v2 dataset, which consists of two different classes—authentic and manipulated images—is one that is frequently utilized in this field.

The authentic class in the CASIA v2 dataset consists of 7492 photos, which are original, real images taken in a range of settings. Algorithms used to detect forgeries use these images as a reference and baseline for comparison.

On the other hand, the tampered class consists of up to 5123 photos that have been purposefully changed or edited to resemble popular counterfeit methods including content insertion, splicing, and cloning. These altered photos, which mirror real-world situations where digital alteration is common, provide a wide range of difficulties for forgery detection systems.

V. PROPOSED METHODOLOGY

Step 1 - Recognize the libraries and modules that the code imports; these give the tools and functionalities that are required.
Step 2 - Analyze the loading, processing, and preparation of data for analysis and model training.
Step 3 - Examine the code pertaining to the training procedure, model architecture, compilation parameters, and assessment measures.
Step 4 - Using the proper assessment measures, evaluate the trained models' performance and, if required, illustrate the results.
Step 5 - Keep an eye out for any further research, observations, or post-processing operations carried out on the model's output or forecasts.

| e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 4, April 2024 ||

| DOI:10.15680/IJIRSET.2024.1304237 |

```
In [73]: model = build_model()
         model.summary()
```

```
Model: "sequential_1"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d_2 (Conv2D)           (None, 124, 124, 32)      2432

 conv2d_3 (Conv2D)           (None, 120, 120, 32)      25632

 max_pooling2d_1 (MaxPoolin  (None, 60, 60, 32)        0
 g2D)

 dropout_2 (Dropout)         (None, 60, 60, 32)        0

 flatten_1 (Flatten)         (None, 115200)            0

 dense_2 (Dense)             (None, 256)               29491456

 dropout_3 (Dropout)         (None, 256)               0

 dense_3 (Dense)             (None, 2)                 514

=================================================================
Total params: 29520034 (112.61 MB)
Trainable params: 29520034 (112.61 MB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

Figure: Model Summary

5.1 Data Collection

Data collection is the process of gathering and measuring information on targeted variables in an established system, which then enables one to answer relevant questions and evaluate outcomes.The goal for all data collection is to capture quality evidence that allows analysis to lead to the formulation of convincing and credible answers to the questions that have been posed. CASIA v2 provides a nearly balanced dataset with 7492 authentic images and 5123 tampered images, allowing algorithms to train effectively on both categories.Real-World Tampering Techniques: The tampered images include manipulations like splicing, insertion, and cloning, reflecting real-world forgery methods. This challenges the detection algorithms and makes them more robust.

5.1. Data Preprocessing and Cleaning

Once we gathered a dataset from Kaggle, our next immediate step is to process that dataset. Data preprocessing is one of the foremost steps in deep learning because the accuracy and performance of the deep learning model will depend upon the quality of data that we are maintaining. Data pre-processing includes handling missing data values, removing duplicates, handling outliers etc. In our dataset after loading the dataset, we resize the image to an standard format i.e (256,256) and then, if the re-sized image contains the noise, we remove it by using de-noising techniques.

5.2. Splitting The Dataset

Using the Sklearn library, we divided the data into train and test subsets in this phase. The user may choose the split ratio, but most likely we will separate the dataset in a 80:20 ratio. This indicates that we will use 80% of the data for training and the leftover 20% for testing. The testing set is used to assess the deep learning model's success on fresh, untested data, while the training set is used to fit the model. In order to prevent overfitting the model to the training data, we must evaluate the model on a testing set to see how well it generalizes to new data. The model is more likely to perform well on fresh data if it does well on both the training and testing sets.

5.4 Building the Deep Learning Model

After splitting the data, we will build a CNN model that contains the following two layers: Conv2D, which reduces the parameters to be learned by the model, and the MaxPooling layer, which captures the important features in the image. The DropOut layer helps to reduce overfitting and enables the model to generalize to unseen data. The Flattening layer is used to convert the image into a 1-D latent representation. Finally, two Dense layers are employed to generate the output and using the ELA analysis to detect whether an image is tampered or not.

| e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 4, April 2024 ||

| DOI:10.15680/IJIRSET.2024.1304237 |

5.5 Calculate Performance Parameters

This step consists of calculating different performance measures such as recall, precision, accuracy etc the model with the most acceptable is the one which is having good accuracy and good recall and good precision is considered as a good model i.e The next step after implementing a deep learning algorithm is to find out how effective is the model based on metric and datasets. Different performance metrics are used to evaluate different deep Learning Algorithms. For example a model is used to distinguish between authentic images and the tempered image. Example of a metric for evaluation of deep learning algorithms is precision, accuracy, recall, confusion matrix. Therefore, we see that different metrics are required to measure the efficiency of different algorithms, also depending upon the dataset at hand.

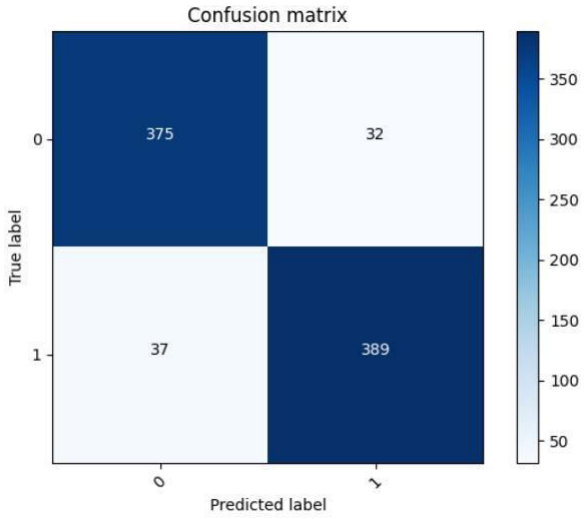


Figure: Confusion matrix

VI. RESULTS

After an extensive training process on a large dataset, the model has achieved impressive results in terms of accuracy. The model's superior performance in both training and testing phases validates its effectiveness as a powerful classifier, capable of accurately identifying the image, whether the image is forgered. The model's consistent and impressive results highlight its reliability and suitability for real-world scenarios, making it a promising choice for diverse deep learning and Computer vision applications.

The below are the images that show the accuracy of the model and shows the best possible accuracy that we reached.

```
In [108]: correct += correct_r
          total += total_r
          print(f'Total: {total_r}, Correct: {correct_r}, Acc: {correct_r / total_r * 100.0}')
          print(f'Total: {total}, Correct: {correct}, Acc: {correct / total * 100.0}')

          Total: 7354, Correct: 6951, Acc: 94.5199891215665
          Total: 9418, Correct: 8976, Acc: 95.30685920577618
```

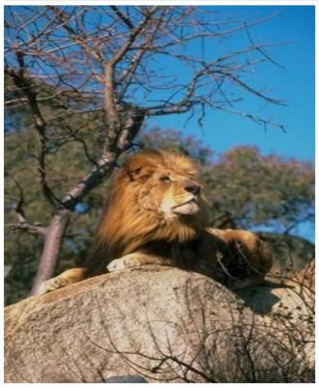
Figure: Accuracy

| e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 4, April 2024 ||

| DOI:10.15680/IJIRSET.2024.1304237 |

```
image_path1 =r"D:\image forgery detection\CASIA1\Au\Au_ani_0028.jpg"
Image.open(image_path1)
```

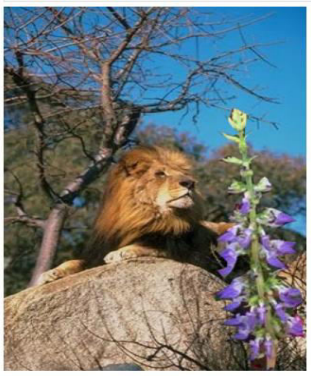


```
image = prepare_image(image_path1)
image = image.reshape(-1, 128, 128, 3)
y_pred = model.predict(image)
y_pred_class = np.argmax(y_pred, axis = 1)[0]
print(f'Class: {class_names[y_pred_class]} Confidence: {np.amax(y_pred) * 100:0.2f}')
```

```
1/1 [==============================] - 0s 33ms/step
Class: Authentic Confidence: 100.00
```

Figure: Authentic Image Confidence

```
image_path2 = r"D:\image forgery detection\CASIA1\Sp\Sp_D_NNN_A_ani0028_pla0007_0284.jpg"
Image.open(image_path2)
```



```
image = prepare_image(image_path2)
image = image.reshape(-1, 128, 128, 3)
y_pred = model.predict(image)
y_pred_class = np.argmax(y_pred, axis = 1)[0]
print(f'Class: {class_names[y_pred_class]} Confidence: {np.amax(y_pred) * 100:0.2f}')
```

```
1/1 [==============================] - 0s 36ms/step
Class: Tampared Confidence: 99.87
```

Figure: Tempered Image Confidence

VII. CONCLUSION

In conclusion, the development of an image forgery detection system utilizing deep learning methods shows a great deal of potential for resolving the growing issues brought on by digital image alteration. With the use of cutting-edge approaches and techniques like Error Level Analysis (ELA) and Convolutional Neural Networks (CNNs), our suggested solution seeks to improve the scalability, accuracy, and efficiency of forgery detection procedures.

ACKNOWLEDGMENT

We want to express our deep-felt gratitude and sincere thanks to our guide Mr. M. Hari Prasad, Assistant Professor, Department of AI, Anurag University, for his skillful guidance, timely suggestions, and encouragement in completing this project. We want to express our profound gratitude to all for having helped us in achieving this dissertation.

| e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 4, April 2024 ||

| DOI:10.15680/IJIRSET.2024.1304237 |

Finally, we would like to express our heartfelt thanks to our parents, who were very financially and mentally supportive and for their encouragement to achieve our goals.

REFERENCES

1. Cozzolino, Davide, et al. "Splicebuster: A new blind image splicing detector." IEEE Transactions on Information Forensics and Security 12.3 (2022) pp no: 667-682.
2. Korus, Piotr, and Jiwu Huang. "DCT-based forensic detection of region duplication forgery in digital images." IEEE Transactions on Information Forensics and Security 10.6 (2022) pp no: 1163-1176.
3. Seo, Y, Shin, K,. Hierarchical convolutional neural networks for fashion image classification. Expert Systems with Applications. 2019: 328-339.
4. Bayram, Serkan, and Husrev Taha Sencar. "An efficient and robust method for detecting copy-move forgery." Digital Signal Processing 23.3 (2019) pp no: 825-833.
5. Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, et al. Recent advances in convolutional neural networks. Pattern Recognition. 2018; 77: 354-77.
6. Jeronymo DC, Borges YCC, Coelho L dos S. Image forgery detection by semi-automatic wavelet soft-Thresholding with error level analysis. Expert Systems with Applications. 2017; 85: 348-56.
7. Oommen RS, Jayamohan M, Sruthy S. Using Fractal Dimension and Singular Values for Image Forgery Detection and Localization. Procedia Technology. 2016; 24: 1452-9.
8. Kuo C-CJ. Understanding convolutional neural networks with a mathematical model. Journal of Visual Communication and Image Representation. 2016; 41: 406-13.
9. Kim J, Sangjun O, Kim Y, Lee M. Convolutional Neural Network with Biologically Inspired Retinal Structure. Procedia Computer Science. 2016; 88: 145-54.
10. Liu S, Deng W. Very deep convolutional neural network based image classification using small training sample size. 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR). 2015.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi cross ref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details