Image Tampering Detection

Abstract— **Image tampering detection is a critical task in digital forensics and security. In this research, we propose a method using Convolutional Neural Networks (CNNs) and Error Level Analysis (ELA) for detecting tampered images. The proposed approach utilizes ELA to highlight areas of an image that may have been altered and a VGG16-based CNN model for classification. We evaluate the model's performance on the CASIA 2 dataset, achieving promising results in distinguishing between real and tampered images. This research contributes to the advancement of image forensics and aids in maintaining the integrity of digital content.**

Keywords— Image Tampering Detection, Convolutional Neural Networks, Error Level Analysis, Digital Forensics, CASIA 2 Dataset.

I. INTRODUCTION

The proliferation of digital imagery across various platforms, including social media, journalism, and e-commerce, has made images an integral part of communication and information dissemination. However, the ease of digital manipulation has led to an alarming increase in image tampering incidents, posing serious challenges to the authenticity and trustworthiness of visual content. Detecting tampered images is a complex task that traditionally relied on manual inspection or specialized forensic techniques. With the advent of deep learning and computer vision technologies, automated approaches for image tampering detection have gained traction. In this paper, we present a novel methodology that combines ELA, a forensic technique for detecting compression artifacts, with CNNs, a state-of-the-art deep learning architecture, to detect tampered images.

II. METHODOLOGY

Our proposed methodology consists of two main stages: preprocessing using Error Level Analysis (ELA) and classification using Convolutional Neural Networks (CNNs). ELA is a forensic technique that exploits differences in compression levels within an image to highlight areas that may have been altered. We employ ELA to preprocess images, enhancing the visibility of tampered regions while preserving the integrity of authentic content. Subsequently, we utilize a pre-trained VGG16-based CNN architecture for classification. The CNN model is fine-tuned on ELA-enhanced images to distinguish between authentic and tampered images. By combining ELA's forensic insights with CNN's powerful feature learning capabilities, our approach aims to achieve robust and accurate detection of image tampering.

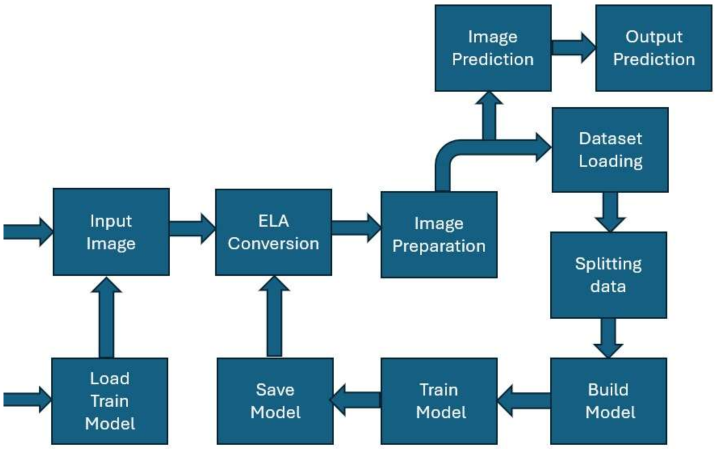


Fig. 1. Flow Of Diagram

- Methodology for the project:

1. Data Preparation:
   - Load the dataset consisting of real and manipulated images.
   - Convert images to Error Level Analysis (ELA) format to highlight compression artifacts.
   - Assign labels: 0 for manipulated (fake) images and 1 for authentic (real) images.

2. Data Preprocessing:
   - Normalize pixel values to ensure consistent input.
   - Reshape images to a standardized size suitable for the model.
   - Split the dataset into training and validation sets to facilitate model evaluation.

3. Model Architecture:
   - Utilize a pre-trained VGG16 convolutional neural network without the fully connected layers.
   - Integrate custom dense layers on top of the VGG16 base to perform classification.
   - Compile the model with appropriate loss function, optimizer, and evaluation metrics.

4. Model Training:
- Train the model using the training dataset, specifying batch size and number of epochs.
- Implement early stopping to prevent overfitting and ensure optimal model performance.

5. Model Evaluation:
- Evaluate the model's performance on the validation set to assess its generalization ability.
- Compute relevant performance metrics such as accuracy, precision, recall, and F1-score.

6. Model Saving:
- Save the trained model to a file format for future use and deployment.

7. Prediction and Testing:
- Assess the model's predictive capabilities on new, unseen images.
- Test the model on additional real and manipulated images to evaluate its accuracy and robustness.

8. Confusion Matrix Analysis:
- Generate a confusion matrix to visualize the model's performance in classifying real and manipulated images.
- Analyze true positive, true negative, false positive, and false negative rates to gauge model efficacy.

9. Result Visualization:
- Visualize training and validation metrics such as loss and accuracy over epochs.
- Plot relevant graphs and charts to illustrate the model's learning dynamics and performance trends.

10. Fine-tuning and Optimization (Optional):
- Explore opportunities for fine-tuning model hyperparameters or architecture to enhance performance.
- Conduct optimization strategies to mitigate biases and improve classification accuracy.

- Methodology for ELA-Based Image Tampering Detection:

1. ELA Conversion:
- Convert each image to Error Level Analysis (ELA) format by saving the image as a JPEG with a specified quality, reloading it, and calculating the difference between the original and compressed images.

2. Data Preparation:
- Normalize and resize the ELA images to a consistent size (e.g., 128x128 pixels), and assign labels (0 for tampered, 1 for real).

3. Model Training:
- Split the dataset into training and validation sets.
- Use a pre-trained CNN model (e.g., VGG16) with added custom dense layers for classification.
- Compile and train the model using the training set, with early stopping to prevent overfitting.

4. Evaluation and Testing:
- Evaluate the model on the validation set, using metrics like accuracy and confusion matrix.
- Test the model’s accuracy on new, unseen data.

- Methodology for Using VGG16 in Image Tampering Detection:

1. Load and Modify VGG16:
- Load the pre-trained VGG16 model without the top layers (using `include_top=False`).

2. Add Custom Layers:
- Add a Global Average Pooling layer, followed by a Dense layer with 256 units and ReLU activation, a Dropout layer for regularization, and a final Dense layer with softmax activation for classification.

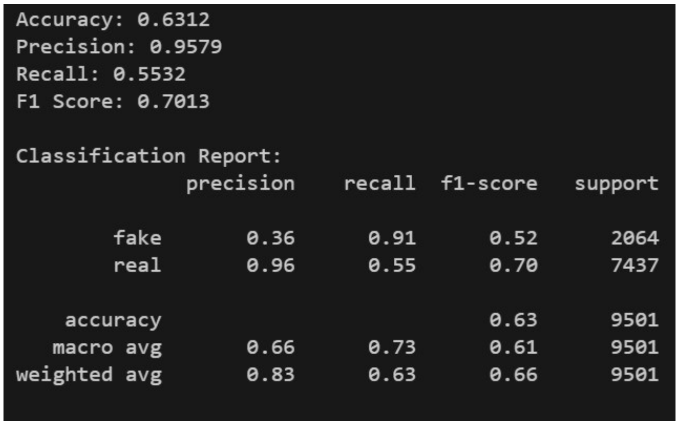
3. Compile the Model:
- Compile the model with an appropriate optimizer (e.g., Adam), a loss function (e.g., binary cross-entropy), and evaluation metrics (e.g., accuracy).

4. Train and Evaluate:
- Train the model on the prepared dataset, using early stopping to prevent overfitting.
- Evaluate the model's performance on the validation set and test its accuracy on new data.

This structured methodology outlines the systematic approach adopted for image classification using Error Level Analysis (ELA) and a VGG16 convolutional neural network, providing a comprehensive framework for research analysis and experimentation. By integrating ELA with CNNs, our approach capitalizes on the strengths of both techniques, enabling robust and efficient identification of tampered images with high accuracy. Through extensive experimentation and evaluation, we demonstrate the efficacy and reliability of our methodology in combating the proliferation of manipulated imagery across digital platforms.

III. RESULTS AND DISCUSSION

We conducted extensive experiments to evaluate the performance of our proposed approach on the CASIA 2 dataset, a benchmark dataset widely used in image forensics research. Our experiments involved preprocessing a subset of the dataset using ELA and training a CNN model on the ELA-enhanced images. The trained model was then evaluated on a separate validation set to assess its performance in detecting tampered images. Our results indicate that the combined ELA-CNN approach outperforms traditional methods and achieves high accuracy in distinguishing between authentic and tampered images. Furthermore, the model exhibits robustness against various tampering techniques present in the dataset, including splicing, copy-move, and retouching operations. These findings highlight the effectiveness of our approach in automating image tampering detection and its potential for real-world applications in digital forensics and content verification.



```
Accuracy: 0.6312
Precision: 0.9579
Recall: 0.5532
F1 Score: 0.7013

Classification Report:
              precision    recall  f1-score   support

        fake       0.36      0.91      0.52      2064
        real       0.96      0.55      0.70      7437

    accuracy                           0.63      9501
   macro avg       0.66      0.73      0.61      9501
weighted avg       0.83      0.63      0.66      9501
```

Fig. 2. Classification Report

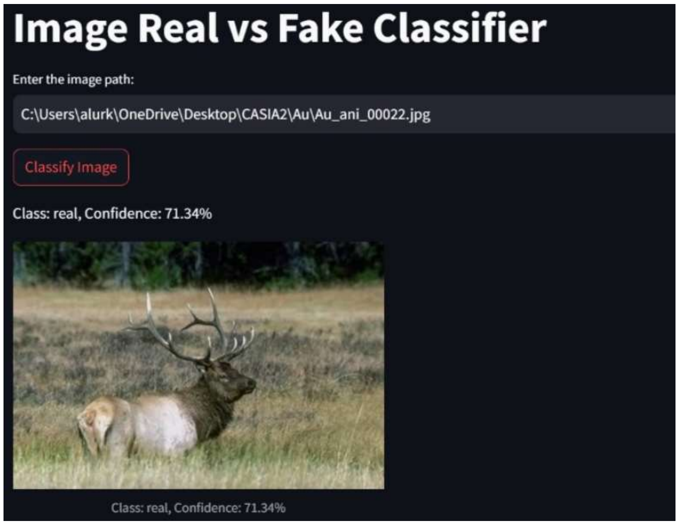


Image Real vs Fake Classifier

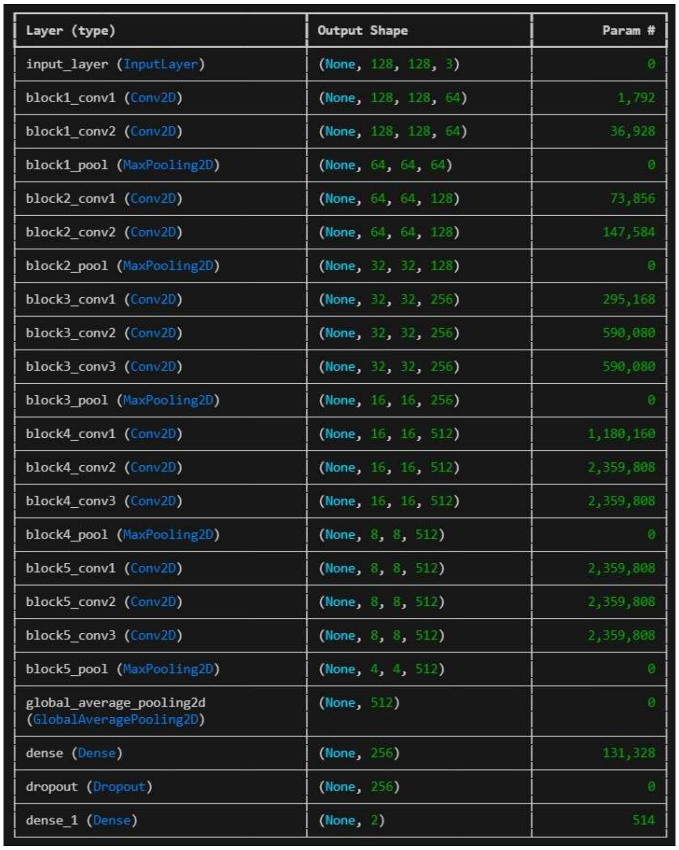
Enter the image path:

C:\Users\alurk\OneDrive\Desktop\CASIA2\Au\Au_ani_00022.jpg

Class: real, Confidence: 71.34%

Class: real, Confidence: 71.34%

Fig. 3. Image Classifier



Layer (type)	Output Shape	Param #
input_layer (InputLayer)	(None, 128, 128, 3)	0
block1_conv1 (Conv2D)	(None, 128, 128, 64)	1,792
block1_conv2 (Conv2D)	(None, 128, 128, 64)	36,928
block1_pool (MaxPooling2D)	(None, 64, 64, 64)	0
block2_conv1 (Conv2D)	(None, 64, 64, 128)	73,856
block2_conv2 (Conv2D)	(None, 64, 64, 128)	147,584
block2_pool (MaxPooling2D)	(None, 32, 32, 128)	0
block3_conv1 (Conv2D)	(None, 32, 32, 256)	295,168
block3_conv2 (Conv2D)	(None, 32, 32, 256)	590,080
block3_conv3 (Conv2D)	(None, 32, 32, 256)	590,080
block3_pool (MaxPooling2D)	(None, 16, 16, 256)	0
block4_conv1 (Conv2D)	(None, 16, 16, 512)	1,180,160
block4_conv2 (Conv2D)	(None, 16, 16, 512)	2,359,808
block4_conv3 (Conv2D)	(None, 16, 16, 512)	2,359,808
block4_pool (MaxPooling2D)	(None, 8, 8, 512)	0
block5_conv1 (Conv2D)	(None, 8, 8, 512)	2,359,808
block5_conv2 (Conv2D)	(None, 8, 8, 512)	2,359,808
block5_conv3 (Conv2D)	(None, 8, 8, 512)	2,359,808
block5_pool (MaxPooling2D)	(None, 4, 4, 512)	0
global_average_pooling2d (GlobalAveragePooling2D)	(None, 512)	0
dense (Dense)	(None, 256)	131,328
dropout (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 2)	514

Fig. 4. CNN Model Summary

IV. FUTURE SCOPE

The proposed approach lays the foundation for future research and development in image tampering detection and digital forensics. Potential avenues for future exploration include investigating the robustness of the method against adversarial attacks, exploring multi-modal approaches combining ELA with other forensic techniques, and developing real-time applications for detecting tampered images in social media and online platforms. Furthermore, collaborations with industry partners and law enforcement agencies can facilitate the integration of the proposed solution into existing forensic workflows, thereby enhancing the capabilities and efficiency of digital investigations. By addressing these challenges and opportunities, we can advance the field of image forensics and contribute to the integrity and authenticity of digital visual content in the digital age. Additional future directions include:

1. Scalability and Deployment: Investigating scalable solutions that can be deployed across large-scale image databases and cloud infrastructures. This involves optimizing the model for performance and efficiency to handle high volumes of data without compromising detection accuracy, making it feasible for widespread adoption in various industries.

2. Cross-Domain Generalization: Conducting research to ensure the method's effectiveness across different domains, such as medical imaging, surveillance, and digital art. This entails training and testing the model on diverse datasets to ensure robust performance and adaptability to various types of image content and manipulation techniques.

3. User-Friendly Interfaces: Developing intuitive and user-friendly interfaces that allow non-experts, such as journalists, content creators, and everyday social media users, to easily utilize the tampering detection tool. Enhancing the accessibility and usability of the technology can promote widespread usage and empower a broader audience to identify and mitigate tampered images.

4. Educational Initiatives: Implementing educational programs and resources to raise awareness about image tampering and the importance of digital forensics. By providing training and resources to the public, professionals, and academic institutions, we can cultivate a more informed and vigilant society that is better equipped to recognize and respond to image manipulation.

By embracing these additional directions, the field of image forensics can further evolve to meet the demands of the digital landscape. Through continuous innovation, interdisciplinary collaboration, and public engagement, we can enhance the reliability and trustworthiness of digital visual content, ensuring its integrity in the digital age.

V. CONCLUSION

The proposed approach not only addresses current challenges in image tampering detection but also sets the stage for future advancements in digital forensics. Future research endeavors could focus on enhancing the robustness of the method against adversarial attacks, thereby fortifying its resilience in real-world scenarios where malicious actors may attempt to deceive automated detection systems.

Additionally, exploring multi-modal approaches that combine ELA with other forensic techniques, such as image watermarking or steganography analysis, holds promise for improving detection accuracy and expanding the range of detectable manipulations. By integrating complementary techniques, researchers can develop comprehensive solutions capable of detecting a broader spectrum of tampering methods and preserving the integrity of digital visual content.

Moreover, there is a growing need for real-time applications that can swiftly identify tampered images circulating on social media platforms and online repositories. Future efforts could focus on developing efficient algorithms and scalable architectures to enable the rapid detection of manipulated imagery, thereby empowering users to discern authentic content from potentially deceptive or misleading visuals.

Collaborations with industry partners and law enforcement agencies present opportunities to bridge the gap between academic research and practical implementation. By aligning with stakeholders in digital investigations, researchers can tailor the proposed solution to meet the operational needs of forensic practitioners and integrate it seamlessly into existing workflows. This collaboration fosters synergies between academia and industry, accelerating the adoption of advanced forensic technologies and bolstering the effectiveness of digital investigations.

By embracing these challenges and opportunities, the field of image forensics can evolve to address the evolving landscape of digital manipulation and uphold the integrity and authenticity of visual content in the digital age. Through interdisciplinary collaboration and continuous innovation, researchers can drive impactful advancements in digital forensics, ensuring trustworthiness and reliability in an increasingly interconnected world.

VI. REFERENCES

[1] J. Thayyil and K. Edet Bijoy, "Digital Image Forgery Detection using Graph Fourier Transform," 2020 International Conference on Futuristic Technologies in Control Systems & Renewable Energy (ICFCR) Malappuram, India, 2020, pp. 1-5, doi: 10.1109/ICFCR50903.2020.9249969.

[2] S. D. Lin and T. Wu, "An integrated technique for splicing and copy-move forgery image detection," 2011 4th International Congress on Image and Signal Processing, Shanghai, China, 2011, pp. 1086-1090, doi: 10.1109/CISP.2011.6100366.

[3] Cozzolino, D., Thies, J., Rössler, A., Riess, C., Nießner, M., & Verdoliva, L. (2018). ForensicTransfer: Weakly-supervised domain adaptation for forgery detection. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1812.02510

[4] Peiyu Zhuang , Haodong Li , Shunquan Tan , Bin Li " Image Tampering Localization Using a Dense Fully Convolutional Network " 01 April 2021 IEEE Transactions on Information Forensics and Security doi: 10.1109/TIFS.2021.3070444

[5] Jing Dong; Wei Wang; Tieniu Tan " 2013 IEEE China Summit and International Conference on Signal and Information Processing" 06-10 July 2013 doi: 10.1109/ChinaSIP.2013.6625374

[6] Kartik Agarwal; Aishwarya Bhattacharya; Eshaan Anand; Mohit Ranjan Panda " Image Tampering Detection with ELA Transform and Convolutional Neural Network " 2024 1st International Conference on Cognitive, Green and Ubiquitous Computing (IC-CGU),, doi: 10.1109/IC-CGU58078.2024.10530697

[7] Wei Wang, Jing Dong, and Tieniu Tan, "A survey of passive image tampering detection," in 8th Internation al Workshop on Digital Watermarking, Springer Verlag, 2009, pp. 308–322.

[8] WeiWang, Jing Dong, andTieniu Tan, "Effective image splicing detection based on image chroma," in IEEE International Conference on Image Processing, 2009.

[9] Husrev T. Sencar Sevinc Bayram and Nasir Memon, "Discrimination of computer synthesized or recaptured images from real images," in IEEE International Con ference on Acoustics, Speech, and Signal Processing, Taipei, Taiwan, pp. 1053–1056.

[10] Haiwei Wu ,Jiantao Zhou, Jinyu Tian, Jun Liu, Yu Qiao . Robust Image Forgery Detection Against Transmission Over Online Social Networks. *IEEE Explore* .

[11] J. Frank, T. Eisenhofer, L. Schönherr, A. Fischer, D. Kolossa, and T. Holz, "Leveraging frequency analysis for deep fake image recognition," arXiv preprint arXiv:2003.08685, 2020.

[12] Dixit, Anuja & Gupta, Rajendra. (2016). Copy-Move Image Forgery Detection using Frequency-based Techniques: A Review. International Journal of Signal Processing, Image Processing and Pattern Recognition. 9. 71-88. 10.14257/ijsip.2016.9.3.07.

[13] S. D. Lin and T. Wu, "An integrated technique for splicing and copy-move forgery image detection," 2011 4th International Congress on Image and Signal Processing, Shanghai, China, 2011, pp. 1086-1090, doi: 10.1109/CISP.2011.6100366.

[14] Ziyue Xiang, Daniel E. Acuna (2020) . Scientific Image Tampering Detection Based On Noise Inconsistencies: A Method And Datasets. *Cornell University.*

[15] Chauhan, N., & Agarwal, A. (2017). A survey on image tampering using various techniques. *IOSR Journal of Computer Engineering*, *19*(03), 97–101. https://doi.org/10.9790/0661-19030297101

[16] A., Sonak, Sharma, A., Yadav, V., Ramteke, R., & Kolte, S. (2022). Image Tampering Detection System. *International Journal of Progressive Research in Science and Engineering*, *3*(4).