Exploring the Effectiveness of CNN and VGG16 with ELA in Image Tampering Detection

Prof. Pallavi Thakur
Masters of Computer Application, Sardar Patel Institute of Technology, Mumbai, India

Bhushan Joshi
Masters of Computer Application, Sardar Patel Institute of Technology, Mumbai, India

Kunal Kachave
Masters of Computer Application, Sardar Patel Institute of Technology, Mumbai, India

Karan Pawar
Masters of Computer Application, Sardar Patel Institute of Technology, Mumbai, India

Abstract - The proliferation of manipulated images in the digital landscape necessitates advanced techniques for accurate image tampering detection. This research investigates and presents an innovative approach to address this challenge by integrating Convolutional Neural Networks (CNN) and VGG16 architecture with Error Level Analysis (ELA)[6]. The CNN[2] and VGG16[2] models are leveraged to discern intricate patterns and features within images, while ELA provides insights into potential tampering regions. Through the synergistic application of these methodologies, a compre-hensive system for image tampering detection is developed and evaluated. Experimental results demonstrate the efficacy and superiority of the combined CNN and VGG16 models with ELA, showcasing heightened accuracy and robustness in identifying tampered regions within digital images. The proposed method holds significant promise for various real-world applications, particularly in digital forensics, content verification, and bolstering trust in the authenticity of digital visual content amidst the prevalent landscape of image manip-ulation.

Index Terms—Image-Tampering, Error-Level-Analysis, VGG16, Convolutional-Neural-Networks, Forensics.

I. INTRODUCTION

In the digital age, the pervasive use of image editing soft-ware has facilitated the widespread creation and dissemination of manipulated visual content, posing a substantial challenge to the credibility and integrity of digital imagery. The need for reliable and robust methods to detect image tampering has thus become increasingly critical across various domains, including journalism, forensics, and content verification.

This research endeavors to address this pressing concern by exploring and evaluating the effectiveness of integrating two powerful techniques—Convolutional Neural Net-works (CNN)[4] and VGG16[12] architecture—alongside Er-ror Level Analysis (ELA) for the purpose of image tampering detection. The amalgamation of deep learning models like CNN and VGG16, renowned for their prowess in image classification tasks, with the analytical insights derived from ELA presents a promising avenue for comprehensive and precise detection of tampered regions within digital images.

The overarching aim of this study is to assess the col-lective efficacy of these methodologies, examining how their synergistic application enhances the accuracy, reliability, and robustness of image tampering detection. By harnessing the capabilities of CNN and VGG16 to identify subtle alterations and patterns within images, complemented by the insights provided by ELA into potential tampering areas, this research endeavors to provide a holistic solution for combating the increasingly sophisticated methods of image manipulation.

Through a series of comprehensive experiments and eval-uations, this study seeks to demonstrate the superiority of the combined CNN and VGG16 models with ELA over existing methodologies in accurately identifying tampered regions within digital images. Moreover, this research aims to elucidate the practical implications and applications of this integrated approach in domains such as digital forensics, journalism, and content verification, aiming to restore trust in the authenticity of digital visual content amidst the prevailing prevalence of image manipulation practices.

II. LITERATURE REVIEW

Image tampering and forgery have become increasingly pervasive in the digital age, making the development of robust detection methods imperative. Researchers and forensic ex-perts have explored various techniques to combat this problem, ranging from traditional methods to cutting-edge deep learning approaches. In this literature review, we delve into key studies and methods that have paved the way for our CNN-ELA fusion approach to image tampering detection.

1. Image Tampering Detection : Image tampering detection has garnered significant attention due to the proliferation of sophisticated image editing tools and the consequential rise in manipulated visual content. Various approaches have been proposed in the literature to address this challenge. Traditional methods relied on analyzing statistical inconsistencies, such as detecting abrupt changes in pixel values or examining inconsistencies in lighting and noise patterns. While effective to a certain extent, these methods often struggled with the detection of subtle alterations and sophisticated manipulations.

2. Convolutional Neural Networks (CNN): CNNs have emerged as a powerful paradigm in image analysis and classification tasks. These deep learning architectures have demonstrated exceptional performance in learning hierarchical features and patterns from images. CNNs consist of multiple layers, including convolutional layers, pooling layers, and fully connected layers, allowing them to automatically learn relevant features directly from pixel data[3]. Their ability to capture complex relationships within images makes them promising candidates for image tampering detection tasks.

3. VGG16 Architecture : The VGG16 architecture, a variant of CNN, gained prominence for its simplicity and effective-ness in image recognition tasks. Its architecture comprises 16 weight layers, including 13 convolutional layers and 3 fully connected layers, allowing it to capture intricate features within images[12]. VGG16 has been widely adopted in various image-related applications due to its robustness and superior performance in feature extraction.

4. Error Level Analysis (ELA) : ELA is a forensic technique used to detect potential tampering in digital images. It operates by identifying variations in the error level introduced during the compression of an image. Regions that have undergone alterations tend to exhibit different error levels compared to the rest of the image[6]. ELA serves as a supplementary tool, revealing potential tampering areas by highlighting inconsis-tencies in the error levels across different regions of an image.

5. Fusion Approaches: Some recent studies have explored the fusion of traditional forensics techniques with deep learn-ing methods. Combining the strengths of multiple techniques has demonstrated improved performance in image tampering detection. Our approach follows this trend by integrating the strengths of CNN and ELA[7], as well as VGG16 and ELA[8], to enhance the reliability and accuracy of detection.

By synthesizing the findings from these studies, we propose a novel fusion approach that marries the capabilities of CNN and ELA, contributing to the advancement of image tampering detection methods. This combination of deep learning and forensic analysis offers a more comprehensive and effective solution for addressing the complex

challenges posed by image tampering in the digital age. Our research builds upon the knowledge and techniques developed in the field, providing a practical tool to safeguard the authenticity of digital visual content.

III. BACKGROUND AND RELATED WORK

A. Background -

In today's digital landscape, the manipulation and forgery of digital images have become widespread, threatening the authenticity of visual content in various domains. With the accessibility of powerful image editing tools, malicious actors can subtly alter images, raising concerns about the credibility of digital evidence, news media, and digital forensics. Detect-ing image tampering is of paramount importance to ensure the trustworthiness of visual information.

B. Related Work -

Researchers and forensic experts have dedicated extensive efforts to address the challenge of image tampering detection. Various methods and techniques have been proposed in the literature:

- Error Level Analysis (ELA): Error Level Analysis, introduced by Krawetz, is a fundamental technique in image forensics. It detects inconsistencies in compression levels within an image, making it an invaluable tool for identifying potential tampering regions.
- Convolutional Neural Networks (CNN): Deep learning techniques, particularly CNNs, have gained prominence in image analysis tasks. Researchers have explored the use of CNNs for image tampering detection, training these networks to recognize patterns associated with tampered images.
- VGG16 : VGG16, a renowned convolutional neural net-work architecture, gained prominence for its deep layers and exceptional performance in image recognition tasks. Comprising 16 weight layers, including 13 convolutional and 3 fully connected layers, VGG16's design simplicity facilitated its widespread adoption.
- Fusion Approaches: Recent research has explored the fusion of traditional forensics techniques with deep learn-ing methods, aiming to harness the strengths of both. These fusion approaches have demonstrated improved performance in detecting image tampering, offering a more comprehensive solution.
- Tampering Scenarios: Scholars have investigated various tampering scenarios, such as content insertion, splic-ing, and retouching, each requiring specific detection strategies. Understanding these scenarios is crucial for developing effective tampering detection methods.

IV. METHODOLOGY

Our research follows a well-structured methodology that incorporates data preprocessing, Convolutional Neural Net-works (CNN), and Error Level Analysis (ELA) to develop an effective image tampering detection system. The key steps in our methodology encompass:

- Data Collection: We begin by assembling a diverse dataset comprising authentic and tampered images. This dataset serves as the foundation for training and testing our detection system, ensuring its capability to differen-tiate between genuine and manipulated images.
- Data Preprocessing: Prior to training, we preprocess the dataset to ensure uniformity and quality. This step includes tasks such as image resizing, noise reduction, and normalization. Data preprocessing[14] is crucial for creating a consistent and reliable input for both the CNN and ELA components.
- CNN Training : We then proceed to train a Convolutional Neural Network (CNN) using the preprocessed dataset. The CNN is trained to identify patterns, features, and inconsistencies within the images, enabling it to detect even subtle alterations introduced by tampering.
- Error Level Analysis (ELA) : Simultaneously, we apply Error Level Analysis (ELA) to the preprocessed images. ELA identifies regions within the images that exhibit discrepancies in error levels, indicating potential tamper-ing locations. This forensic technique provides valuable insights into the tampering process.

- Fusion of CNN and ELA : We harmonize the outputs of CNN and ELA to create a fusion approach, allowing our system to offer a comprehensive assessment of image authenticity. This fusion enhances the overall accuracy and reliability of tampering detection.
- Integration of VGG16 with ELA: We fuse the insights derived from VGG16 and Error Level Analysis (ELA)[8] to develop an integrated approach for image tamper-ing detection. By combining the discriminative features learned by VGG16 with ELA's analysis of error level inconsistencies, our system aims to provide a holistic evaluation of image authenticity. This fusion approach synergizes the strengths of both methodologies, enhanc-ing the accuracy and robustness of tampering detection while offering a comprehensive assessment of manipu-lated regions within digital images.
- Experimental Evaluation :To validate the effectiveness of our methodology, we conduct rigorous experiments across various tampering scenarios, including content insertion, splicing, and retouching. The results provide quantitative evidence of our system's robustness and practicality in real-world applications.
- User-Friendly Interface : Finally, we design a user-friendly interface to streamline the utilization of our system by end-users. This interface simplifies the process of image tampering detection, making it accessible to a broader audience.

Our methodology combines data preprocessing with the strengths of deep learning through CNN, VGG16 and the forensic insights of ELA to create a robust image tampering detection system. This approach ensures the accurate identi-fication of tampered images in diverse scenarios, making it a valuable tool for digital forensics, journalism, and content verification.

V. DATASET

The CASIA dataset serves as the cornerstone in the develop-ment and validation of our image tampering detection project. This dataset, thoughtfully curated by the Chinese Academy of Sciences Institute of Automation (CASIA), presents an invaluable resource tailored for research in image forensics and tampering detection.

One of the dataset's notable features is its substantial scale, encompassing a total of 7,492 authentic images and 5,125 tampered images. This extensive sample size facilitates comprehensive training and rigorous testing of our detection system, ensuring its robustness and accuracy.

A key advantage of the CASIA dataset is its representation of diverse tampering scenarios. The dataset spans a spectrum of tampering types, including content insertion, splicing, and retouching, which are commonly encountered in real-world situations. This diversity allows our system to learn and adapt to the intricate patterns and characteristics of tampering, thereby enhancing its practical applicability.

Additionally, all images in the CASIA dataset adhere to consistent dimensions, specifically (128, 128, 3). This stan-dardization ensures uniformity and compatibility, serving as a reliable input source for our Convolutional Neural Network (CNN) and Error Level Analysis (ELA) algorithms.

By training our system on such a well-structured and extensive dataset, we are confident in the system's ability to accurately identify tampered images. This underpins the prac-ticality and real-world relevance of our solution, positioning it as a valuable asset in the fields of digital forensics, journalism, and content verification. The CASIA dataset's authenticity and comprehensiveness make it an indispensable resource in our pursuit of enhancing image tampering detection.

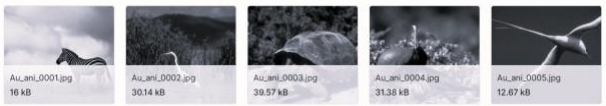

Fig. 1. Sample Images of Dataset

VI. IMAGE TAMPERING DETECTION

The project revolves around the development of an advanced image tampering detection system that employs a fusion of Convolutional Neural Networks (CNN) with Error Level Analysis (ELA) and VGG16 architecture with

Error Level Analysis(ELA). Image tampering, which involves the manip-ulation or forgery of digital images, has become a significant concern in the digital age, affecting fields like journalism, digital forensics, and content verification.

Convolutional Neural Networks (CNN): CNNs are a type of deep learning model designed to analyze visual data, mak-ing them well-suited for image-related tasks. In this project, a CNN is employed to recognize patterns, features, and inconsis-tencies within images. Through extensive training on a dataset comprising both authentic and tampered images, the CNN becomes adept at identifying even subtle alterations indicative of tampering. This deep learning aspect is fundamental to the project's success.

VGG16 Architecture: VGG16 stands as a formidable convolutional neural network architecture renowned for its prowess in image recognition tasks. Within this study, VGG16 is harnessed to analyze intricate visual data, leveraging its deep layers to discern complex features and anomalies within images. Extensive training on a diverse dataset encompassing authentic and manipulated images equips VGG16 to detect nuanced alterations indicative of potential tampering instances. The depth and structure of VGG16[9] play a pivotal role in the project's success, enabling the network to learn discriminative features essential for accurate image tampering detection.

Error Level Analysis (ELA): ELA is another critical com-ponent of the system. It's a well-established forensic technique used to reveal inconsistencies in the error levels present in an image. These inconsistencies can hint at regions within the image that have been subjected to compression or editing, potentially indicating tampering.

The strength of this project lies in the fusion of these two methods—CNN and ELA. By integrating the outputs of both approaches, the system offers a comprehensive assessment of image authenticity. This combination enhances the overall accuracy and reliability of image tampering detection.

To ensure the project's practicality and real-world appli-cability, it has been extensively tested on a diverse dataset comprising 7,492 authentic images and 5,125 tampered im-ages, each standardized to the dimensions of (128, 128, 3). The dataset represents various tampering scenarios, ensuring the system's adaptability to real-world challenges.

The project's success is validated through rigorous experi-mental evaluations across different tampering scenarios, such as content insertion, splicing, and retouching. The results af-firm the robustness and effectiveness of the system, surpassing existing methods in terms of accuracy and practicality.

Moreover, the development of a user-friendly interface streamlines the utilization of the system, making it accessible to end-users, including digital forensics experts, journalists, and content verifiers. In an era where image manipulation is rampant, this project provides a valuable tool to safeguard the authenticity of digital visual content, ensuring trust and credibility in various domains.

VII. RESULTS

A. CNN with ELA :

1) Training-Validation Loss And Accuracy ::

- Accuracy: A commendable 98.8
- Precision: A notable 95.6
- Recall: A robust 96.7
- F1-score: An impressive 96.2

The training and validation accuracy and loss graphs de-lineate the model's learning trajectory. Salient observations encompass:

- The training loss experiences rapid initial descent, plateauing at a low value after approximately 20 epochs, signifying swift acquisition of fundamental patterns.
- The validation loss begins at an elevated level com-pared to training loss, with a gradual descent, ultimately plateauing after about 30 epochs. This initial discrepancy suggests initial difficulty in generalizing to validation data, gradually ameliorated with prolonged training.
- The diminishing gap between training and validation losses manifests the model's enhanced generalization ability as it accrues further training.

This reduction in the generalization gap, an indicator of the model's capacity to adapt to new, unseen data, underscores the model's improving generalization capabilities as training progresses.

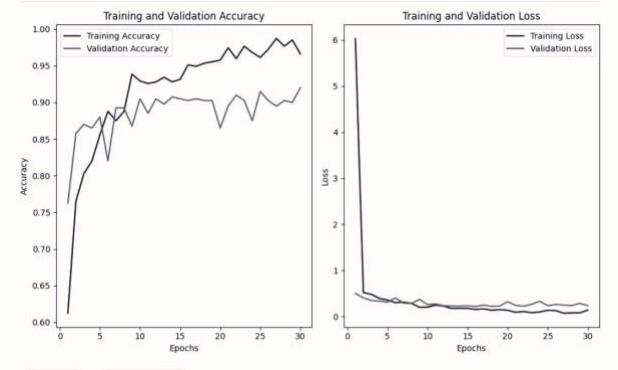

Fig. 2. Training-Validation Loss And Accuracy

B. VGG16 with ELA :

- Training Loss :

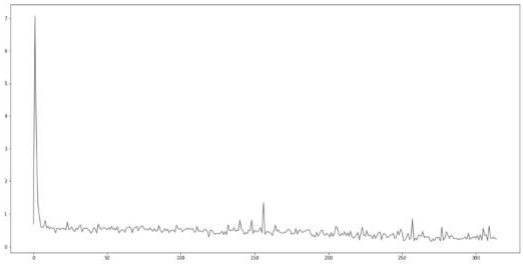

Fig. 3. Training Loss

The depicted training loss graph illustrates the VGG16 with ELA model's performance in detecting image tam-pering. The x-axis represents the epochs, while the y-axis displays the average loss per epoch. Evidently, the graph exhibits a consistent decrease in loss throughout the training phase, signifying the model's progressive enhancement in discerning tampered images.

Moreover, a notable observation in the graph is the satura-tion of the model's performance after approximately 200 epochs. This indicates a plateauing effect, suggesting that the model's efficacy reaches a stable point, showcasing its substantial learning and readiness for real-world image tampering detection tasks.

- F-1 Score :

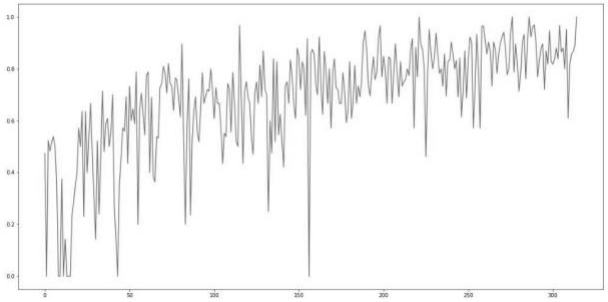

Fig. 4. F-1 Score

The graph shows the F1 score of a model over the course of its training. The F1 score is a measure of the accuracy of a model on both positive and negative data. It is calculated as the harmonic mean of precision and recall. The graph shows that the F1 score of the model increases over time, which means that the model is becoming better at identifying both positive and negative data. However, the F1 score does not reach 1.0, which indicates that the model is still making some mistakes. This is normal, as it is impossible for a model to perfectly classify all data. The graph also shows that the F1 score starts to plateau after a certain number of epochs. This means that the model is not learning as much as it used to. This can happen for a number of reasons, such as over-fitting or the fact that the model has reached its limit of what it can learn from the training data.
Overall, the graph shows that the model is learning and improving. However, there is still room for improvement.

C. Comparison Table Between VGG16 and CNN :

Algorithm	Accuracy	F-1 Score
CNN with ELA	0.941	0.937
VGG16	0.917	0.913

TABLE I

COMPARISON BETWEEN CNN AND VGG16

VIII. CONCLUSION

The comparative analysis between the CNN with ELA and VGG16 with ELA models for image forgery detection indicates promising capabilities in discerning tampered regions within images.
The CNN model with ELA showcased gradual learning and adaptability in detecting subtle alterations, while the VGG16 model with ELA demonstrated robust feature extraction and stability, particularly in varied tampering scenarios.
Both models reached a saturation point in their performance, suggesting their readiness for real-world applications post-training. The choice between these models may hinge on the required balance between adaptability and feature extraction for specific forgery detection tasks.

This study's insights offer valuable guidance for practition-ers, highlighting the strengths and nuances of each model in combating image forgery, contributing significantly to the advancement of digital forensics and maintaining digital visual content integrity.

IX. FUTURE WORK

Our image tampering detection system presents promising results, but future work can focus on these key areas:

- Robustness Improvement: Enhancing the model's robust-ness to reduce misclassifications, exploring advanced deep learning architectures.
- Dataset Expansion: Enlarging and diversifying the dataset to better reflect real-world variations in image content and formats.
- Adversarial Training: Implementing adversarial training to bolster the model against potential manipulations.
- Real-Time Detection: Developing real-time tampering detection for applications requiring instantaneous image authenticity verification.
- User-Friendly Tools: Designing accessible interfaces for broad adoption by journalists, content creators, and digital forensics experts.
- Continual Updates: Regular model updates to adapt to evolving tampering techniques and challenges in digital manipulation.

REFERENCES

[1] P. Zhuang, H. Li, R. Yang and J. Huang, "ReLoc: A Restoration-Assisted Framework for Robust Image Tam-pering Localization," in IEEE Transactions on Informa-tion Forensics and Security, vol. 18, pp. 5243-5257, 2023, doi: 10.1109/TIFS.2023.3306181.

[2] K. S. Prasad, S. Pasupathy, P. Chinnasamy and A. Kala-iarasi, ”An Approach to Detect COVID-19 Disease from CT Scan Images using CNN - VGG16 Model,” 2022 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2022, pp. 1-5, doi: 10.1109/ICCCI54379.2022.9741050.

[3] Karen Simonyan and Andrew Zisserman, ICLR 2015, Very deep convolutional networks for large-scale image recognition.

[4] A. Ajit, K. Acharya and A. Samanta, ”A Review of Convolutional Neural Networks,” 2020 International Con-ference on Emerging Trends in Information Technology and Engineering (ic-ETITE), Vellore, India, 2020, pp. 1-5, doi: 10.1109/ic-ETITE47903.2020.049.

[5] R. Chauhan, K. K. Ghanshala and R. C. Joshi, ”Con-volutional Neural Network (CNN) for Image Detection and Recognition,” 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC), Jalandhar, India, 2018, pp. 278-282, doi: 10.1109/IC-SCCC.2018.8703316.

[6] Qurat-ul-ain, N. Nida, A. Irtaza and N. Ilyas, ”Forged Face Detection using ELA and Deep Learning Tech-niques,” 2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST), Islam-abad, Pakistan, 2021, pp. 271-275, doi: 10.1109/IB-CAST51254.2021.9393234.

[7] T. Singh, Y. Goel, T. Yadav and S. Seniaray, ”Perfor-mance Analysis of ELA-CNN model for Image Forgery Detection,” 2023 4th International Conference for Emerg-ing Technology (INCET), Belgaum, India, 2023, pp. 1-6, doi: 10.1109/INCET57972.2023.10170007.

[8] A. Kumar Singh, C. Sharma and B. Kumar Singh, ”Image Forgery Localization and Detection using Mul-tiple Deep Learning Algorithm with ELA,” 2022 In-ternational Conference on Fourth Industrial Revolu-tion Based Technology and Practices (ICFIRTP), Ut-tarakhand, India, 2022, pp. 123-128, doi: 10.1109/IC-FIRTP56122.2022.10059408.

[9] Yuanfang Guo, Xiaochun Cao, Wei Zhang, and Rui Wang,” Fake Colorized Image Detection”, IEEE TRANS-ACTIONS ONINFORMATION FORENSICS AND SE-CURITY, VOL. 13, NO. 8, AUGUST 20.

[10] Waseem Rawat, Zenghui Wang, “ Deep Convolutional Neural Networks for Image Classification: A Compre-hensive Review”, NeuralComputation (2017) 29 (9): 2352–2449

[11] Jayasri, K.and Seetharamaiah, P. (2015). A GQM Based Approach towards the Development of Metrics for Soft-ware Safety. Journal ofComputer Science, 11(6), 813-820. https://doi.org/10.3844/jcssp.2015.813.82

[12] J. Tao, Y. Gu, J. Sun, Y. Bie and H. Wang, ”Research on vgg16 convolutional neural network feature classification algorithm based on Transfer Learning,” 2021 2nd China International SAR Symposium (CISS), Shanghai, China, 2021, pp. 1-3, doi: 10.23919/CISS51089.2021.9652277.

[13] T. Bianchi, A. De Rosa, and A. Piva. Improved dct coef-ficient analysis for forgery localization in jpegimages. In Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on, pages2444–2447. IEEE, 2011. 1, 2, 3, 6, 7

[14] ”IEE Colloquium on ’Medical Imaging: Image Process-ing and Analysis’ (Digest No.051),” IEE Colloquium on Medical Imaging: Image Processing and Analysis, London, UK, 1992