ISSN: 2278-4632
Vol-14, Issue-8, August: 2024

DIGITAL IMAGE FORGERY DETECTION USING DEEP LEARNING

Battula Aarthika
Email: aarthikabattula123@gmail.com
M. Tech, Department of Computer Science and Engineering.
Annamacharya Institute of Technology and Science, Hyderabad, Telangana, India.

Ramesh Babu Varugu
Email: rameshvarugu82@gmail.com
Assistant Professor
& HOD, Department of CSE.
Annamacharya Institute of Technology and Science, Hyderabad, Telangana, India.

Abstract - The advent of digital image manipulation tools has exacerbated the proliferation of image forgeries, necessitating robust solutions for their detection. This project presents a novel approach to address this challenge, utilizing Python and Convolutional Neural Network (CNN) model architecture. The CNN model, employed as the core of our forgery detection system, has exhibited remarkable performance. With a training accuracy of 98% and a validation accuracy of 92%, it showcases its efficacy in distinguishing authentic from tampered images. The dataset utilized in this study comprises 12,615 images, consisting of 7,492 authentic (real) images and 5,123 tampered (fake) images, providing a diverse and extensive testbed for evaluation. Toenhance the precision of our approach, we incorporate Error Level Analysis (ELA) as apreprocessing step. Each image is resized to a standardized 256x256 resolution, after which ELA is applied. ELA aids in the identificationof regions within an image that exhibit varying compression levels. In an untampered image, all regions should exhibit uniform compression. Deviations from this uniformity may indicate digital manipulation. The processed images are stored as numpy arrays for subsequent analysis. Our proposed system leverages the synergy between deep learning through CNNs and the subtleties uncovered by ELA. This combination empowers the model to not only achieve high accuracy but also to provide insights into the specific regions of potential manipulation within an image. By harnessing the capabilities of Python and a well-structured CNN architecture, this project represents a significant stride towards robust digital image forgery detection, with potential applications in various domains where image authenticity is paramount.

Keywords: Deep neural network (DNN), image compression, image forgery detection (IFD), pretrained model, transfer learning.

I. INTRODUCTION

In the digital age, the manipulation and dissemination of visual content have become ubiquitous, giving rise to the critical challenge of ensuring the authenticity and integrity of digital images. The project, "Digital Image Forgery Detection Using CNN and Error Level Analysis (ELA)," addresses this pressing concern by presenting an innovative solutionthat leverages the power of deep learning and advanced image analysis techniques to detect image forgeries.

As the use of image editing tools and the ease of digital content creation continue to expand, the potential for digital image manipulation, alteration, and forgery has never been greater.

ISSN: 2278-4632
Vol-14, Issue-8, August: 2024

This poses serious implications for various domains, including journalism, forensics, social media, and e-commerce, where the credibilityof visual content is paramount. It is imperative to develop robust and reliable methods for detecting manipulated images to safeguard trust and ensure the integrity of digital visual information.

The proposed project takes a multi-faceted approach to digital image forgery detection, combining Convolutional Neural Networks (CNNs), a cutting-edge deep learning architecture, with Error Level Analysis (ELA), a powerful image analysis technique. By synergizing these methods, the system aims to excel in identifying tampered images, distinguishing between authentic and manipulated content with a high degree of accuracy and robustness.

This project's significance extends beyond mere detection; it plays a crucial role in preserving the authenticity of digital images and combating the spread of manipulated or falsified visual information. The ability to identify digital image forgeries has far-reaching implications, from maintaining the credibilityof news sources to aiding law enforcement in forensic investigations and ensuring the trustworthiness of e-commerce product images.

In the following sections, we will delve into the details of the proposed system, including its architecture, methodology, advantages, and potential future directions. By the end of this project, we aim to provide a comprehensive solution that empowers users to confidently assess the authenticity of digital images in an era characterized by the rapid proliferation of digital content.

II. EXISTING SYSTEM

In the realm of digital image forgery detection, the utilization of deep learning techniques has become increasingly prevalent. An existing system, which employed the MobileNetV2 architecture, stands as a testament to the efficacy of this approach in addressing the critical challenge of detecting image forgeries.

MobileNetV2 is a state-of-the-art neuralnetwork architecture that has been specifically designed for mobile and embedded visionapplications. Its lightweight design and computational efficiency make it an attractive choice for tasks where resource constraints area concern. In the context of digital image forgery detection, MobileNetV2 provides astreamlined and effective solution.

The existing system, utilizing MobileNetV2, demonstrated impressive results in distinguishing between authentic and tampered images. Through a process of feature extraction and hierarchical learning, the model was able to capture intricate patterns and nuances withinthe images, thereby achieving high accuracy in forgery detection.

III. PROPOSED SYSTEM

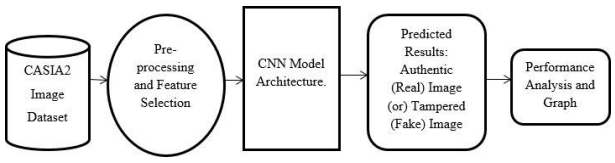
The proposed system for digital image forgery detection is a comprehensive approach that combines Convolutional Neural Network (CNN) architecture with Error Level Analysis (ELA) to achieve accurate and robust detection of manipulated images. This system addresses the critical need for reliable forgery detection in the era of digital image manipulation.

ISSN: 2278-4632
Vol-14, Issue-8, August: 2024

The proposed system employs a CNN model as its core. The CNN is designed to perform feature extraction and classification of images efficiently. The model is trained on a diverse dataset comprising both authentic and tampered images. During training, it learns to identify distinctive patterns, features, andinconsistencies that indicate imagemanipulation. Various hyper parameters, suchas the number of layers, filter sizes, and learning rates, are fine-tuned to optimize themodel's performance. Appropriate activation functions, such as ReLU (Rectified Linear Unit), are used to introduce non-linearity into the model.

Before feeding images into the CNN, they undergo preprocessing using ELA. All images are resized to a standardized resolution (e.g., 256x256 pixels) to ensure consistency. ELA is applied to each image. This involves saving the image at a specific compression level and then subtracting this compressed image from the original image. The resulting ELA image highlights areas with differing compression levels, potentially indicating regions of manipulation.

IV. System Design & Development



DATA FLOW DIAGRAM:

The DFD is also called as bubble chart. It is a simple graphical formalism that can be used to represent a system in terms of input data to the system, various processing carried out on this data, and the output data is generated by this system.

The data flow diagram (DFD) is one of themost important modeling tools. It is used tomodel the system components. These components are the system process, the dataused by the process, an external entity that interacts with the system and the information flows in the system.

DFD shows how the information moves through the system and how it is modified by a series of transformations. It is a graphical technique that depicts information flow and the transformations that are applied as data moves from input to output.

DFD is also known as bubble chart. A DFD may be used to represent a system at any level of abstraction. DFD may be partitioned into levels that represent increasing information flow and functional detail.

IMPLEMENTATION MODULES:

- Dataset
- Importing the necessary libraries
- Retrieving the images
- Splitting the dataset
- ELA image analysis
- Building the model
- Apply the model and plot the graphs for accuracy and loss
- Accuracy on test set
- Saving the Trained Model

V. LITERATURE SURVEY

Multiple image splicing dataset (MISD): A dataset for multiple splicing.

K. D. Kadam, S. Ahirrao, and K.Kotecha.

Image forgery has grown in popularity due to easy access to abundant image editing software. These forged images are so devious that it is impossible to predict with the naked eye. Such images are used to spread misleading

ISSN: 2278-4632
Vol-14, Issue-8, August: 2024

information in society with the help of various social media platforms such as Facebook, Twitter, etc. Hence, there is an urgent need for effective forgery detection techniques. In order to validate the credibility of these techniques, publically available and more credible standard datasets are required. A few datasets are available for image splicing, such as Columbia, Carvalho, and CASIA V1.0. However, these datasets are employed for the detection of image splicing. There are also a few custom datasets available such as Modified CASIA, AbhAS, which are also employed for thedetection of image splicing forgeries. A study of existing datasets used for the detection of image splicing reveals that they are limited to only image splicing and do not contain multiple spliced images. This research work presents a Multiple Image Splicing Dataset, which consists of a total of 300 multiple spliced images. We are the pioneer in developing the first publicly available Multiple Image Splicing Dataset containing high-quality, annotated, realistic multiple spliced images. In addition,we are providing a ground truth mask for these images. This dataset will open up opportunities for researchers working in this significant area.

Deep learning based algorithm (ConvLSTM) for copy move forgery detection.

M. A. Elaskily, M. H. Alkinani, A. Sedik, and M. M. Dessouky.

Protecting information from manipulation is important challenge in current days. Digital images are one of the most popular information representation. Images could be used in several fields such as military, social media, security purposes, intelligence fields, evidences in courts, and newspapers. Digital image forgeries mean adding unusual patterns to the original images that cause a heterogeneity manner inform of image properties. Copy move forgery isone of the hardest types of image forgeries to be detected. It is happened by duplicating part or section of the image then adding again in the image itself but in another location. Forgery detection algorithms are used in image security when the original content is not available. This paper illustrates a new approach for Copy Move Forgery Detection (CMFD) built basically on deep learning. The proposed model is depending on applying (Convolution Neural Network) CNN in addition to Convolutional Long Short-Term Memory (CovLSTM) networks. This method extracts image features by a sequence number of Convolutions (CNVs) layers, ConvLSTM layers, and pooling layers then matching features and detecting copy move forgery. This model had been applied to four aboveboard available databases: MICC- F220, MICC-F2000, MICC-F600, and SATs-130. Moreover, datasets have been combined to build new datasets for all purposes of generalization testing and coping with an over-fitting problem. In addition, the results of applying ConvLSTM model only have been added to show the differences in performance between using hybrid ConvLSTM and CNN compared with using CNN only. The proposed algorithm, when using number of epoch's equal 100, gives high accuracy reached to 100% for some datasets with lowest Testing Time (TT) time nearly 1 second for some datasets when compared with the different previous algorithms.

VI. SYSTEM TESTING

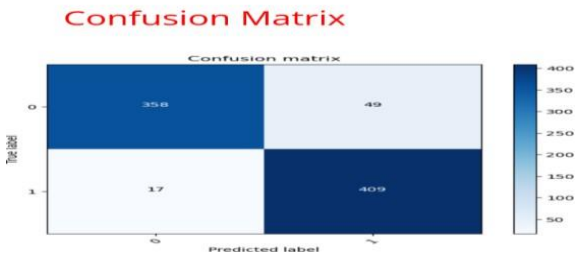
Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs

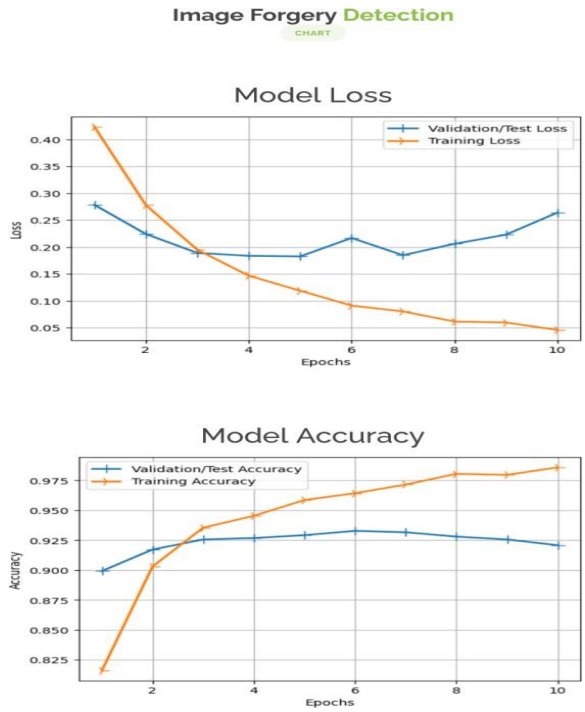
ISSN: 2278-4632
Vol-14, Issue-8, August: 2024

produce valid outputs. All decision branchesand internal code flow should be validated. It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performsaccurately to the documented specifications and contains clearly defined inputs and expected results.

Unit testing is an essential practice in software development that involves testing individual units or components of a software application in isolation. Each unit, typically a small piece of code or a function, is tested to ensure that it functions correctly and produces expected outcomes. Unit testing plays a pivotal role in maintaining code quality, catching bugs early, and facilitating efficient debugging and maintenance.

VII. RESULT





VIII. CONCLUSION

In the ever-evolving landscape of digital media, ensuring the authenticity and integrity of images is a critical concern. The project,"Digital Image Forgery Detection Using CNN and Error Level Analysis (ELA)," presents acomprehensive and effective solution to addressthis challenge.

Through the fusion of Convolutional Neural Network (CNN) model architecture and Error Level Analysis (ELA), this project has demonstrated a robust system capable of accurately detecting digital image forgeries. By leveraging the strengths of both techniques, the system achieves high accuracy and adaptability, making it well-suited for a wide range of forgery detection scenarios.

The utilization of a diverse dataset containing authentic and tampered images ensures the system's ability to generalize and perform

ISSN: 2278-4632
Vol-14, Issue-8, August: 2024

effectively in real-world applications. Its real-time implementation potential further enhances its utility, allowing it to be seamlessly integrated into various platforms and applications where immediate forgery detectionis paramount.

With its ability to identify both simple and complex forgeries, the proposed system contributes to the preservation of image authenticity and the prevention of digital manipulation. It offers a practical solution for forensic analysts, content moderators, and individuals seeking to verify the credibility of digital visual content.

In conclusion, the "Digital Image Forgery Detection Using CNN and ELA" project represents a significant advancement in the field of digital image forensics. Its robustness, adaptability, and real-time capabilities position it as a valuable tool in the ongoing battle to maintain the trustworthiness of digital images in an era of digital manipulation and misinformation..

IX. FUTURE SCOPE

While the "Digital Image Forgery Detection Using CNN and Error Level Analysis (ELA)" project has made significant strides in the realm of digital image forensics, there are several avenues for future work and enhancements to further improve the system's capabilities:

Deep Learning Architectures: Explore and experiment with more advanced deep learning architectures beyond CNNs. Techniques suchas recurrent neural networks (RNNs), attention mechanisms, or transformer-based models may offer improved performance and context understanding.

Transfer Learning: Investigate the potential benefits of transfer learning by pretraining on a large, diverse dataset and fine-tuning on the specific forgery detection task. This approach can expedite model convergence and potentially improve accuracy.

X. REFERENCES

[1] K. D. Kadam, S. Ahirrao, and K.Kotecha, ‘‘Multiple image splicing dataset (MISD): A dataset for multiple splicing,’’ Data, vol. 6, no. 10, p. 102, Sep. 2021.

[2] R. Agarwal, O. P. Verma, A. Saini, A. Shaw, and A. R. Patel, ‘‘The advent of deep learning-based,’’ in Innovative Data Communication Technologies and Application. Singapore: Springer, 2021.

[3] M. A. Elaskily, M. H. Alkinani, A. Sedik, and M. M. Dessouky, ‘‘Deep learning based algorithm (ConvLSTM) for copy move forgery detection,’’ J. Intell. Fuzzy Syst., vol. 40, no. 3, pp. 4385–4405, Mar. 2021.

[4] A. Mohassin and K. Farida, ‘‘Digital image forgery detection approaches: A review,’’ in Applications of Artificial Intelligence in Engineering. Singapore: Springer, 2021.

[5] K. B. Meena and V. Tyagi, Image Splicing Forgery Detection Techniques: A Review. Cham, Switzerland: Springer, 2021.

[6] S. Gupta, N. Mohan, and P. Kaushal, ‘‘Passive image forensics using universal techniques: A review,’’ Artif. Intell. Rev., vol. 55, no. 3, pp. 1629–1679, Jul. 2021.

ISSN: 2278-4632
Vol-14, Issue-8, August: 2024

[7] W. H. Khoh, Y. H. Pang, A. B. J. Teoh, and S. Y. Ooi, ‘‘In-air hand gesture signature using transfer learning and its forgery attack,’’ Appl. Soft Comput., vol. 113, Dec. 2021, Art. no. 108033.

[8] Abhishek and N. Jindal, ‘‘Copy move and splicing forgery detection using deep convolution neural network, and semantic segmentation,’’ Multimedia Tools Appl., vol.80, no. 3, pp. 3571–3599, Jan. 2021.

[9] M. M. Qureshi and M. G. Qureshi, Image Forgery Detection & Localization Using Regularized U-Net. Singapore: Springer, 2021.

[10] Y. Rao, J. Ni, and H. Zhao, ‘‘Deep learning local descriptor for image splicing detection and localization,’’ IEEE Access, vol. 8, pp. 25611–25625, 2020.