WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Detection of Digital Image Forgery Using Deep-Learning

G. Prasad[1], G Rakesh Reddy[2], M. Abhishek Guptha[3], B. Harinath[4], N. Kamal Reddy[5]

[1,2,3,4,5] Computer Science & Engineering (B. Tech, JNTUH), Sphoorthy Engineering College (JNTUH)
prasadcs@sphoorthyengg.ac.in[1], grakeshreddy@sphoorthyengg.ac.in[2], abhishekguptha2002@gmail.com[3], b.harinath2511@gmail.com[4], kamalreddy243@gmail.com[5]
DOI: https://doi.org/10.55248/gengpi.5.0524.1376

ABSTRACT

In contemporary times, digital images serve as a primary medium for disseminating information across social media platforms. However, malicious software can manipulate these images to spread false information. So, it's crucial to identify these forgeries. The literature addresses this issue through a range of digital image forgery detection techniques. However, most of these techniques are limited to detecting only one type of forgery, such as image splicing or copy-move, which is not practical for real-life applications. This paper presents an approach to improve digital image forgery detection by utilizing deep learning techniques through transfer learning, aiming to simultaneously uncover two types of image forgery. The proposed technique relies on discovering the compressed quality of the forged area, which normally differs from the compressed quality of the rest of the image. A deep learning-based model is proposed for detecting digital image forgery by calculating the difference between the original image and its compressed version. This process generates a featured image, which serves as input to the pre-trained model. The model is then trained by removing its original classifier and adding a new, fine-tuned classifier. A comparison of eight different pre-trained models adapted for binary classification has been conducted. The experimental results indicate that utilizing this technique with eight different adapted pre-trained models surpasses state-of-the-art methods. This conclusion is based on a comparison of the resulting evaluation metrics, charts, and graphs. Additionally, the results indicate that employing the technique with the pre-trained MobileNetV2 model achieves the highest detection accuracy rate (approximately 95%) with fewer training parameters, resulting in faster training times.

INTRODUCTION

The tampering of digital images, known as digital image forgery, often goes undetected by the naked eye. These altered images are a major source of fake news and misinformation, particularly on social media platforms like Facebook and Twitter. Free editing software with advanced features, such as GNU, GIMP, and Adobe Photoshop, can easily create these forgeries. To identify such manipulations, digital image forgery detection algorithms and techniques are employed, especially in situations where the original content is unavailable, playing a crucial role in image security. Digital image forgery involves introducing unusual patterns into original images, resulting in heterogeneous variations in image properties and an atypical distribution of image features. Figure 1 illustrates the classification of digital image forgery. Active approaches require critical information about the image for verification. This embedded information is used to detect any alterations. There are two types of active approaches: digital signatures and digital watermarking. Digital signatures involve adding extra data derived from the image at the end of the acquisition process. Digital watermarking can be embedded into images either during the acquisition phase or during the processing phase.

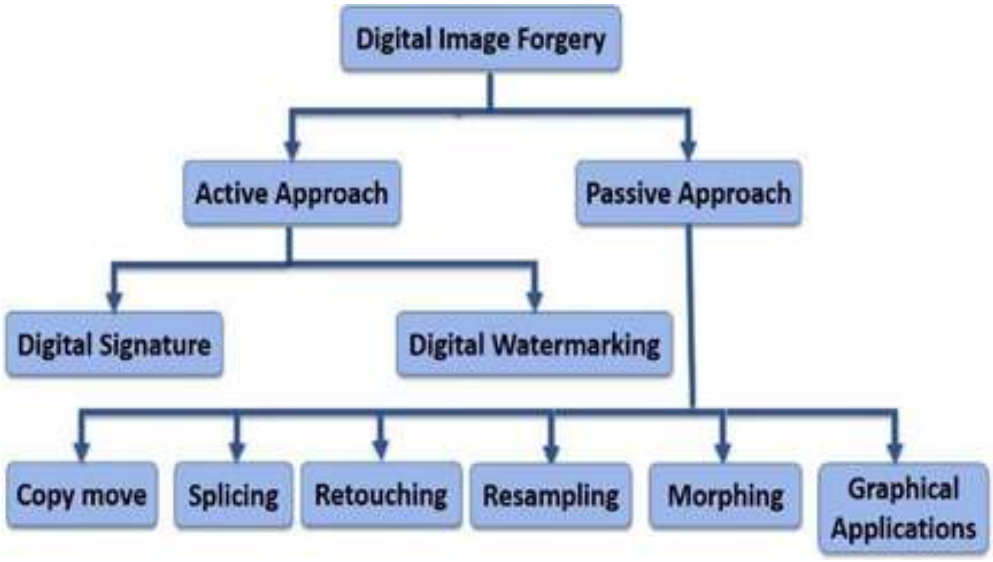

FIGURE 1. Digital image forgery classification.

Deep learning leverages computational models with multiple processing layers to learn data representations at various levels of abstraction. These techniques have significantly advanced the state-of-the-art in fields such as speech recognition, visual object recognition, object detection, drug discovery, and genomics. By employing the backpropagation algorithm, deep learning identifies complex structures in large datasets, guiding machines on how to adjust their internal parameters to compute representations in each layer based on the previous layer. Deep convolutional networks have revolutionized the processing of images, video, speech, and audio, while recurrent networks have excelled in handling sequential data like text and speech.

Deep learning methods are a type of representation learning that utilizes multiple layers of representation. These layers are formed by combining simple yet non-linear modules, each of which transforms the representation at one level (beginning with the raw input) into a higher, slightly more abstract representation.

Supervised learning

The most common form of machine learning, whether deep or not, is supervised learning. Suppose we aim to develop a system that can classify images as containing a house, car, person, or pet. We start by gathering a large dataset of images, each labeled with its respective category. During the training process, the machine is shown an image and generates an output in the form of a vector of scores, one for each category. Ideally, the correct category

should have the highest score, although this is unlikely to occur before training. We then compute an objective function to measure the error (or discrepancy) between the output scores and the desired scores. The machine then adjusts its internal parameters to minimize this discrepancy.

The paper is organized as follows: A literature review is covered in section II. Section III outlines the proposed approach and provides a detailed presentation of the proposed architecture. Section IV outlines the module's discussion, Section V has the algorithms, and Section VI concludes with future work.

LITERATURE REVIEW

In the field of image forgery detection, a variety of approaches have been proposed. Traditional methods typically involve extracting hand-crafted features, followed by classification techniques such as feature matching to distinguish between authentic and forged images. Machine learning approaches may utilize classifiers such as Support Vector Machines and Naïve Bayes classifiers for the classification process.

In contrast, more recent techniques leverage convolutional neural networks (CNNs) and deep neural networks (DNNs). Some of these methods utilize pre-trained models and harness the power of transfer learning. The discussion will focus on CNN and deep learning-based techniques, exploring the use of different pre-trained

Models.

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

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

The prevalence of image forgery has surged due to the widespread accessibility of image editing software. These manipulated images are crafted with such sophistication that they defy detection by the human eye alone. They are utilized to propagate misleading information across various social media platforms like Facebook and Twitter, underscoring the critical need for effective forgery detection techniques. However, validating the reliability of these techniques necessitates access to publicly available and reputable standard datasets. Existing datasets tailored for image-splicing detection, such as

Columbia, Carvalho, and CASIA V1.0, primarily cater to the detection of image-splicing forgeries. Although some custom datasets like Modified CASIA and AbhAS also serve this purpose, they are limited in scope and do not encompass multiple spliced images. Recognizing this gap, our research introduces the multiple-image splicing Dataset comprising a total of 300 multiple-spliced images. This initiative marks the inception of the first publicly available Multiple Image Splicing Dataset, offering high-quality, annotated, and realistic multiple-spliced images. Furthermore, we provide ground truth masks for these images. This dataset holds promise in fostering advancements in research within this significant domain, providing a valuable resource for researchers engaged in forgery detection studies.

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

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

Protecting information from manipulation is a crucial challenge in today's world, with digital images being a widely used form of information representation across various domains such as military, social media, security, intelligence, legal evidence, and journalism. Digital image forgeries involve adding deceptive patterns to original images, altering their properties in a heterogeneous manner. Among these forgeries, copy-move forgery poses a significant challenge in detection. It involves duplicating a part of an image and inserting it into another location within the same image. Forgery detection algorithms play a vital role in image security, particularly when the original content is unavailable.

This paper presents a novel approach for Copy Move Forgery Detection (CMFD) based primarily on deep learning techniques. The proposed model relies on Convolutional Neural Network (CNN) and Convolutional Long Short-Term Memory (ConvLSTM) networks to extract image features through a sequence of Convolution (CNV) layers, ConvLSTM layers, and pooling layers. These features are then matched to identify and detect copy-move forgery. The model is evaluated on four publicly available databases: MICC-F220, MICC-F2000, MICC-F600, and SATs-130. Additionally, new datasets are created by combining existing datasets to facilitate generalization testing and address overfitting issues. Furthermore, the study compares the performance of the hybrid ConvLSTM and CNN model using CNN alone, showcasing the advantages of the proposed approach.

The proposed algorithm demonstrates high accuracy, reaching up to 100% for certain datasets with minimal Testing Time (TT) of nearly 1 second for some datasets, particularly when the number of epochs is set to 100.

TABLE 1 summarizes existing image forgery detection techniques based on deep learning, highlighting previous research efforts. While significant progress has been made in detecting image splicing forgery with high accuracy rates, detecting copy-move forgery remains challenging. It's noteworthy that few studies have addressed the simultaneous detection of both splicing and copy- move forgeries, with lower accuracy rates recorded compared to other techniques.

Forgery Type	Reference	Year	Features Extraction technique	Classification technique	Dataset	
Splicing	[10]	2020	CNN-Based Local Descriptor Construction	SVM	CASIAv2 DVMM DSO-1	**Accuracy**: CASIAV2= 96.97%, DVMM= 97.04%, DSO-1= 97.5%
	[11]	2023	CNN	CNN	CASIAv1 CASIAv2 CUISDE	**Accuracy:** CASIAV1= 99.1%, CASIAV2= 99.3%, CUISDE= 100%
Splicing, Copy-Move Separately	[12]	2022	RGB stream + noise streams	End-to-end fully CNN + (TDSA)	NIST16 CASIA COLUMBIA	**Accuracy:** NIST16= 98.4%, COLUMBIA=97.7%
	[13]	2021	Hybrid Encoding+ Decoding CNN	Hybrid features and semantic reinforcement network HFSRNet	NIST16 COVERAGE CASIAv1	**Accuracy:** NIST16= 98.86%, COVERAGE= 92.76%, CASIAV1= 93.21%
Copy-Move	[14]	2022	DCNN	SD-Net: (super-[BPD) +DCNN	USCISI CoMoFoD CASIAv2	CoMoFoD **P**=59.11 **R**=57.69 **F**=50.77 CASIAv2: **P**=57.48 **R**=51.25 **F**=48.06
	[15]	2021	CNN (Encoder+ decoder)	CNN	CoMoFoD CMFD	**Accuracy**: CoMoFoD= 98.39%, CMFD= 98.78%
Splicing + Copy-Move Together	[16]	2022	CNN	CNN	MICC-F2000	**Accuracy**= 97.52%
	[9]	2021	Regularizing U-Net	Regularizing U-Net	CASIAv2	**F1-Score** = 0.9486
	[17]	2022	Difference Compression Quality +CNN	CNN	CASIAv2	**Accuracy**= 92.23%

TABLE 1. Summary of deep learning-based image forgery detectiontechniques.

In their study, the authors emphasized the utility of CNNs in detecting image forgery, a task often challenging for the human eye due to residual artifacts left by the forgeries. Notably, the disparity in source between the forged region and the background images facilitates detection through compression differences. This distinction served as the basis for training the CNN-based model to identify image forgery. However, the experimental results revealed a detection accuracy rate of 92.3%, indicating room for improvement. Additionally, the model's substantial parameter count necessitates a reduction to optimize CPU and memory consumption. Furthermore, performance metrics such as F1 score, recall, precision, true positive rate (TPR), and true negative

rate (TNR) require enhancement by increasing their values. Moreover, the model exhibits a high false positive rate (FPR) and false negative rate (FNR), highlighting the need for mitigation. These evaluation metrics will be elaborated upon in the subsequent section.

PROPOSED APPROACH

The rise of digital image manipulation tools has exacerbated the spread of image forgeries, highlighting the need for robust detection solutions. This project introduces a fresh approach to tackle this challenge, utilizing Python and Convolutional Neural Network (CNN) architecture. The CNN model, serving as the cornerstone of our forgery detection system, has demonstrated impressive performance. With a training accuracy of 98% and a validation accuracy of 92%, it effectively discerns authentic images from tampered ones. The dataset employed comprises 12,615 images, including 7,492 authentic and 5,123 tampered images, providing a diverse testbed for evaluation.

To refine our approach, Error Level Analysis (ELA) is integrated as a preprocessing step. Each image is standardized to a 256x256 resolution and undergoes ELA. This process aids in identifying regions with differing compression levels, which may indicate digital manipulation. The resulting images are stored as numpy arrays for further analysis.

Our proposed system capitalizes on the combination of deep learning through CNNs and the insights gleaned from ELA. This synergy not only yields high accuracy but also offers insights into potential manipulation regions within images. By harnessing Python and a meticulously designed CNN architecture, this project represents a significant advancement in robust digital image forgery detection, with broad applications in domains where image authenticity is critical.

FIGURE 2. Set of images created in the proposed work.

The set of images created in the proposed work can be shown in Figure 2. The initial image, denoted as (a), depicts the unaltered original image, while (b) portrays forged image that is denoted as F, (c) represents the compressed version of (b)that is denoted as Fcomp, (d) represents the mathematical difference between F and Fcomp denoted as Fdiff.

MODULES

Dataset Collection:

In the initial phase of Digital Image Forgery Detection, the primary focus was on obtaining the input dataset. The process of collecting data is pivotal as it sets the foundation for the subsequent development of the machine learning model. The quality and quantity of data collected significantly impact the performance of the model. The more comprehensive and high-quality the dataset, the better the model's performance is likely to be.

Library Import:

Python serves as the language of choice for this project. The essential libraries imported include Keras for constructing the main model, sklearn for splitting the data into training and testing sets, PIL for converting images into arrays of numerical data, along with other indispensable libraries such as pandas, numpy, matplotlib, and TensorFlow.

Image Retrieval:

This phase involves retrieving images from the dataset and preprocessing them to prepare for training and testing the model. Tasks include reading the images, resizing them to a standardized dimension (e.g., 200x200 pixels), and normalizing the pixel values. Additionally, image labels are retrieved and associated with the respective images. The images are then converted into numpy arrays for further processing.

ELA Image Analysis:

Error Level Analysis (ELA) is employed as a technique to identify manipulated images by comparing images stored at different compression levels. This analysis aids in detecting alterations in the images.

Dataset Splitting:

The dataset is divided into training and testing sets in this phase. The division typically follows an 80- 20 split, with 80% of the data allocated for training and the remaining 20% for testing. This split allows for training the model on a subset of the data, validating its performance, and evaluating its accuracy on unseen data.

Model Construction:

The model is built using the sequential model from the Keras library. Convolutional Neural Network (CNN) layers are added to the model architecture. The initial layers consist of two Conv2D layers with 32 filters each and a kernel size of (5,5). MaxPool2D layers are incorporated with a pool size of (2,2) to down sample the image dimensions. Dropout layers with a dropout rate of 0.25 are used to prevent overfitting by randomly removing 25% of neurons. The process is repeated with adjustments in parameters. A Flatten layer is added to convert the 2D data to a 1D vector, followed by dense layers, dropout layers, and another dense layer. The final dense layer outputs 2 nodes, utilizing the softmax activation function to predict the probability of brain tumor presence or absence.

ALGORITHMS

Convolutional neural networks

Convolutional Neural Networks (ConvNets) are designed to process data represented as multiple arrays, such as a color image composed of three 2D arrays corresponding to pixel intensities in the three color channels. Many data modalities come in the form of multiple arrays: 1D for signals and sequences, including language;2D for images or audio spectrograms; and 3D for video or volumetric images. ConvNets leverage four key principles that take advantage of the properties of natural signals: local connections, shared weights, pooling, and the use of multiple layers.

A typical ConvNet architecture is structured as a series of stages. The initial stages consist of two types of layers: convolutional layers and pooling layers. Units in a convolutional layer are organized into feature maps, where each unit connects to local patches in the feature maps of the previous layer through a set of weights called a filter bank. The outcome of this local weighted sum is then passed through a non-linearity function, such as ReLU. All units in a feature map share the same filter bank, while different feature maps in a layer use different filter banks.

This architecture is designed for two main reasons. Firstly, in array data like images, local groups of values tend to be highly correlated, forming distinctive local patterns that can be easily detected. Second, the local statistics of images and other signals are invariant to location; a motif appearing in one part of an image can appear anywhere, hence the use of shared weights across different locations to detect the same pattern in different parts of the array. The filtering operation performed by a feature map is a discrete convolution, giving ConvNets their name. While convolutional layers detect local conjunctions of features from the previous layer, pooling layers merge semantically similar features into one. Since the relative positions of features forming a motif can vary, reliably detecting the motif is achieved by coarse-graining the position of each feature.

Recurrent neural networks

- When backpropagation was first introduced, its most exciting application was in training recurrent neural networks (RNNs). RNNs are particularly effective for tasks involving sequential inputs, such as speech and language, because they process an input sequence one element at a time. They maintain a 'state vector' in their hidden units, which implicitly contains information about the history of all past elements in the sequence. By considering the outputs of the hidden units at different discrete time steps as if they were the outputs of different neurons in a deep multilayer network, we can apply backpropagation to train RNNs. Although RNNs are powerful dynamic systems, training them is challenging because the backpropagated gradients can either grow or shrink at each time step, often leading to exploding or vanishing gradients over many time steps.
- Achieving high accuracy rates compared to the state-of-the-art results found in the literature is crucial. Utilizing a pre-trained model and leveraging the power of transfer learning, the developed lightweight model, with a small number of parameters, is well-suited for environments with memory and CPU limitations, adding significant value to the proposed architecture.
- The performance of eight different pre-trained models, including VGG16, VGG19, ResNet50, ResNet101, ResNet152, MobileNetV2, Xception, and DenseNet, is evaluated.
- A comparative analysis of these eight pre-trained models and state-of-the-art results is presented.
- The CASIAV2 dataset, one of the best benchmark datasets, is used for evaluation. This dataset presents a significant challenge as it contains two main types of image forgery (splicing and copy-move) in various sizes and formats (TIFF, JPEG, BMP). Additionally, the cropped parts in the forged images have undergone processing, including distortion, rotation, and scaling, to create realistic-looking images, often blurring the edges of the spliced regions, which complicates the detection process.

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)," offers a comprehensive and effective solution to address this challenge. By combining Convolutional Neural Network (CNN) architecture with Error Level Analysis (ELA), the project has developed a robust system capable of accurately detecting digital image forgeries. Leveraging the strengths of both techniques, the system achieves high accuracy and adaptability, making it well- suited

for various forgery detection scenarios. The use of a diverse dataset containing both authentic and tampered images ensures the system's ability to generalize and perform effectively in real-world applications. Its potential for real-time implementation further enhances its utility, allowing seamless integration into various platforms and applications where immediate forgery detection is crucial. Capable of identifying both simple and complex forgeries, the proposed system helps preserve image authenticity and prevent digital manipulation. It provides 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 digital image forensics. Its robustness, adaptability, and real-time capabilities make it a valuable tool in the ongoing effort to maintain the trustworthiness of digital images in an era of digital manipulation and misinformation.

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, in "The Emergence of Deep Learning-Based Techniques," in *Innovative Data Communication Technologies and Applications.*

[3] M. A. Elaskily, M. H. Alkinani, A. Sedik, and M. M. Dessouky, in "Utilizing Convolutional Long Short-Term Memory (ConvLSTM) for Copy-Move Forgery Detection," in the Journal of Intelligence.

[4] A. Mohassin and K. Farida, in "Review of Approaches for Digital Image Forgery Detection," 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] 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.

[7] 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.

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

[9] 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.

[10] S. Gupta, N. Mohan, and P. Kaushal, in "Review of Passive Image Forensics Using Universal Techniques," in the journal Artificial Intelligence. Intell. Rev., vol. 55, no. 3, pp. 1629–1679, Jul. 2021.

[11] F. Li, Z. Pei, W. Wei, J. Li, and C. Qin, ‘‘Image forgery detection using tamper-guided dual self- attention network with the multiresolution hybrid feature,’’ *Secur. Commun. Netw.*, vol. 2022, pp. 1–13, Oct. 2022.

[12] A. K. Jaiswal and R. Srivastava, ‘‘Detection of copy-move forgery in digital image using multi- scale, multi-stage deep learning model,’’*Neural Process. Lett.*, vol. 54, no. 1, pp. 75– 100, Aug.2021.

[13] K. M. Hosny, A. M. Mortda, N. A. Lashin, and M. Fouda, in "A Novel Approach for Splicing Image Forgery Detection Using Convolutional Neural Networks," in the journal Applied. Sci., vol. 13, no. 3, p. 1272, Jan. 2023.

[14] C. Haipeng, C. Chang, S. Zenan, and L. Yingda, in "Hybrid Features and Semantic Reinforcement Network for Image Processing," published in the journal Multimedia Systems, volume 28, issue number undisclosed.no-2, pp. 363–374, 2021.

[15] Q. Li, C. Wang, X. Zhou, and Z. Qin, in "Detection and Localization of Image Copy- Move Forgery Based on Superpixel-Based Segmentation and Deep Convolutional Neural Networks," published in the journal Science. Rep., vol. 12, no. 1, Sep. 2022, Art. no. 14987.