

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 11, November 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

## Impact Factor: 8.625

9940 572 462

🕥 6381 907 438

🛛 🖂 ijircce@gmail.com

🙋 www.ijircce.com

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

## Image Forgery Detection using Convolutional Neural Network

#### M.Shafeeka Begum, M.Ashmi Graha

Department of Computer Science and Engineering, Arunachala College of Engineering for Women, Manavilai,

Vellichanthai, India

Department of Computer Science and Engineering, Arunachala College of Engineering for Women, Manavilai,

Vellichanthai, India

**ABSTRACT:** Digital editing software has become more sophisticated and advanced, which will bring along with a higher probability of risk to digital media integrity and authenticity. The identification of covert image altering techniques like copy-move and splicing is extremely challenging, which can substantially undermine public confidence in visual content. A CNN-based framework for image forgery detection is proposed in this paper. The objective is to create an effective algorithm that can automatically identify the type of picture manipulation based on distinctive characteristics and differentiate real photographs from fakes. In order to extract hierarchical features from photos and identify both local and global patterns indicative of forgery, CNNs were used. The CASIA 2.0 Image Tampering Detection Dataset, which includes both real and altered photos, was used to train the algorithm. According to research findings, the suggested CNN framework can identify manipulated images with a 95% accuracy rate and identify subtle changes that conventional forensic techniques detect. The system works effectively with all suspected forging strategies, including object removal, splicing, and copy-move. The approach here proposes a dependable method for image forgeries using a CNN-based framework. Potential uses include content verification, digital security, and media forensics. By maintaining the integrity of digital images, this framework promotes authenticity in digital media and builds confidence across a range of sectors.

**KEYWORDS:** Image forgery detection, Convolutional Neural Network, Digital media authenticity, Copy-move and splicing, Media forensics.

#### **I. INTRODUCTION**

As the use of digital images in media, social networks, and documentation increases, authenticity in digital visual content becomes vital. Sophistications in copy-move, splicing, and retouching grew the marketplace of image forgery, and yet traditional watermarking or cryptographic hashing approaches were limited owing to their incapability of detecting edits without markers [1]. Consequently, response has come to imply image forgery detection, which is of particular concern to identify alterations in images themselves [2]. New approaches have strayed from rule-based methods to much more adaptive ones that are based on machine learning to manipulate complex patterns of data to find tampering [3]. That in itself serves to strengthen the need for automated, dependable systems whereby one can ascertain the authenticity of the images involved [4].

Advanced machine learning and deep learning techniques, particularly CNN, appear to hold great promise in applications of image classification and feature extraction, making them very suitable in forgery detection [5]. Methods like Support Vector Machines and K-Nearest Neighbours that have been used much in classic image analysis pose many difficulties in capturing very subtle, high dimensional forgery patterns [6]. Deep learning models [7], particularly CNNs, offer more advanced feature extraction, yet traditional applications of CNN might miss the fine details concerning the forgeries without a proper preprocess or training to a specific type of manipulation [8]. Although these techniques take the image analysis forward, they are restricted by faults in the detection capabilities of subtle alteration across any forgery type [9].



The proposed framework alleviates these disadvantages since it makes use of CNNs particularized for forgery detection that exploit the enhancements in feature extraction so as to capture local and global patterns indicative of tampering. This focus is mainly on manipulations relating to copy-move, splicing, and object removal, making use of the CASIA 2.0 Image Tampering Detection Dataset. Applying grid search and cross-validation on the training process of the model increases the accuracy with robustness while enabling forgery that would seem to bypass traditional forensic systems. Therefore, the framework will present an even more holistic approach to detection of image forgery, thus promoting secure and authentic digital media content.

A. Key Contributions of the Proposed Framework

- The Image Tampering Detection Dataset trained and validated has multiple samples, authentic as well as tampered images, in total. It exploits Convolutional Neural Networks CNNs to detect image forgery by focusing on learning both local and global features that are representative of forgery.
- The framework captures hierarchical features thus becoming able to detect subtle alterations that often are evaded by traditional methods of detection.
- The present framework can easily detect manipulation techniques, among which include copy-move, splicing, and object removals addressing the most critical gaps as had been thus far registered in prior image forgery detection methods.

#### B. Organization of the Paper

The outline of the paper will thus be as follows: Section I: Summarizes the subject and what the work intends to achieve that is going to be covered within the paper. A comprehensive literature work is included in section II, and the problem statement is there in section III. Section IV is for outlining the methodology, Section V is where the findings are presented, and Section VI gives a summary of the conclusion drawn through research.

#### **II. RELATED WORKS**

Image Forgery Detection using Support Vector Machines SVM [10] has been used for the classification of features that were extracted from tampered images, and it mainly categorizes in the category of copy-move forgery detection. SVM performs very well on binary classification and succeeds relatively well under controlled circumstances. However, it fails to handle high-dimensional complex data in images, which diminishes its effectiveness in highly diversified datasets to identify slight changes. The use of KNN classifier [11]When local picture features were compared to detect duplicate areas in copy-move forgeries, forgery detection was employed. KNN is great for small datasets due to its interpretability and simplicity, but it suffers from wasteful processing and lower accuracy in large datasets, rendering real-time detection unreliable. Using random forest [12], the tampered areas of the image were classified while applying an ensemble of decision trees that reduces overfitting. It works well with structured datasets but isn't good at capturing complex patterns in a high-resolution image and when done with alterations as small as possible, it completely flops hence not adept at catching subtle or sophisticated forgeries.

PCA [13] is combined with SVM for the reduction in dimensionality of feature space and enhanced the computational efficiency while detecting forgery in an image. In this process, the identification of manipulated portions occurs by focusing on principal components with information loss in the process of dimensionality reduction. K-Means clustering [14] was used for the detection of copy-move forgery by identifying duplicated clusters in images. It segments the images based on similar features for locating the forgery regions. Though useful for simple segmentation, K-Means is very sensitive to initial conditions and usually fails with complex or large-scale image data.

#### **III. PROBLEM STATEMENT**

The existing machine learning methods like SVM [15]and KNN [16]suffer from weak points caused by big image data and slight manipulations. SVM with feature selection shows the problem in dealing with big data and low robustness to manipulation. The proposed CNN-based framework outstrips the aforementioned shortcomings by leveraging deep learning to extract hierarchal both at local and global automatically features for better enhancement in the subtle case of detection of forgery and overall, more robustness. Overall, CNNS are capable of providing greater generalization across



various manipulations of an image compared to the existing techniques. In turn, this leads to better performance on a larger range of imagery forgeries.

#### IV. PROPOSED CNN FRAMEWORK FOR DETECTING FOGERY IMAGES

The work for this proposed CNN-based framework for image forgery detection works based on well-understood stages to discern and decide whether the actual images are genuine or manipulated. The first step under a data preprocessing stage is that all images, whether real or forged, originating from the Image Tampering Detection Dataset, will be sent for processing. This stage enhances the image quality with functions like downscaling, denoising, and contrast, so the CNN attention can be paid to significant visual features about the images. Then, the images with extracted features are fed into the feature extraction stage of the CNN. Several learned convolutional layers were employed; these begin from simple patterns of edges, and proceed to more complex textures and shapes which, based on these are only indicative and evidence for forgery. These layers capture local details such as fine textures and global structures of inconsistencies in shapes. Together, these layers form an all-inclusive representation of the content of the image.Once features are extracted, the framework uses the fully connected layers to perform pattern recognition and classification to differentiate between authentic and tampered images. This stage permits the detection of subtle irregularities caused by certain forgery methods like copy-move, splicing, and object removal.



Fig. 1. Proposed CNN Framework in Detecting Forgery Images

This improves the performance of the framework further, based on optimization and validation, carried out by using a grid search and cross-validation process for the selection of the optimal hyperparameters, including the learning rates and layer configuration. The model attains accuracy and generalization performance. By consequence, the accuracy level attained by the CNN-based framework is 95%. It also detects minor alterations that may not be noticed using conventional forensic techniques. This is a comprehensive structure that makes it a robust tool for applications in media forensics, digital security, and content verification, in efforts to ensure trust in various domains over digital media [17]. The system learns to differentiate between authentic and manipulated images based on the characteristics while the model is in its training process. In this case, the CNN uses that feature of identifying any subtle discrepancy that can lead to manipulation during the training process. After training, the model can classify new images as authentic or manipulated. Classifications use learned features for predictions in high accuracy rates. Finally, the output will be shown: whether the image is forged or not.



#### A. Dataset Description

The proposed framework utilizes Image Tampering Detection Dataset [18], which have both authentic and manipulated images. This dataset, extensively used for forgery detection, incorporates different forgery methods: copy-move, splicing, and object removal. It contains thousands of images with different resolutions and manipulations. Therefore, it's a diversified dataset, making it easy for the model to learn and detect minor changes in images. Authentic as well as tampered images availability make it suitable for training and evaluation.

#### B. Dataset Pre-processing

In order to maintain uniformity throughout preprocessing, the photos are shrunk to a uniform dimension. Data augmentation techniques include rotation, flipping, cropping, and others to make the training data more variable. This aids in the model's generalization. Additionally, normalization is used to ensure that the values of the pixels are within the acceptable range.

#### C. Working of CNN

In the proposed framework, the CNN is used to identify hierarchically significant features from input images to separate genuine and manipulated images [19]. The process begins with convolutional layers through which filters or kernels move over a space in an input image to detect low-level features of edges, textures, and colours. Mathematically, the convolution operation for an input image I and a filter F can be expressed in the Eqn (1):

$$I_{out} = I * F + b \tag{1}$$

where  $I_{out}$  represents the output feature map, denotes the convolution operation b is the bias term. Convolution Convolutional, this lays the guide kernel over on the input image to create a new set of values resulting from a dot product that reflect specific features in an image. The result of convolution is to extract the low-level features from the images with regard to edges, gradients, and textures [20]. The output of convolution now travels through ReLU activation function in order to introduce nonlinearity to the network to provide it with the capabilities of learning complex featurescan be expressed in the Eqn (2):

$$ReLU(x) = max(0, x)$$
(2)

This step eliminates the negative features and keeps only positive ones, thus the network captures only significant patterns important for the classification. Pooling retains the most important information while reducing the workload of the network during computations. This step helps one avoid overfitting and instead enable the network to generalize very well across the various images. When progress is made down through the different layers of the network, higher levels of convolutional layers are applied than in earlier layers of the network, which capture more abstract and complex patterns, like shapes and objects as well as other salient features pertinent for the detection of image manipulation.

Deeper layers enable the CNN to have much better content understanding in an image. After passing through a couple of convolution and pooling layers, a 1D vector; these vectors are then interpreted by the network to classify the image by extracting the hierarchical features. In this layer, the SoftMax function is utilized to output whether it is an authentic or manipulated image.

#### V. RESULTS AND DISCUSSION

The performance evaluation of the suggested model's Python implementation is included in the results section. The effectiveness of the model has been assessed using the F1-score, accuracy, and recall criteria. In this context, the CNN-based framework has achieved exceptional results when compared to more conventional models such as SVM, Random Forest, and K-NN. This is due to the CNN model's ability to learn hierarchical characteristics of a picture, which gives it a significant advantage over previous models that have been shown to be unable to detect subtle forgeries and object removal. In this sense, it can be said that the CNN framework demonstrates a significant level of effectiveness and strength in content verification and digital forensics.



Analysis of Dataset Pre-processing

This phase involves preprocessing the dataset to ensure that its quality and suitability for training this model will be enhanced. These include cleaning of data wherein missing or incorrect values are addressed and normalization wherein it scales the data for better algorithms as shown in the Fig. 2.



Original Image

**Resized Image** 

Normalized Image

#### Fig. 2. After Pre-processing- Normalized Image

Techniques for data augmentation are used to expand the dataset's size and diversity in a way that will guarantee its resilience in identifying different picture alterations. Feature extraction is then carried out, where main attributes of the images, such as edges and texture, are captured as well as color histograms. Lastly, the dataset is split into training and testing sets, which then places the model in the position of being tested on blind data.

Visualization of Forgery Image Detection

An approach to Forged Image Detection, meant to track A variety of picture manipulation techniques have been introduced. The three primary categories of manipulations are object removal, splicing, and copy-move forgery. Specifically, as illustrated in Fig. 3, manipulations such as Copy-Move Forgery depend on copying a portion of the image and then pasting it inside the original image. The regions of content duplication are in general "invisible" to the traditional detection algorithms. Splicing is compositing elements where different images are combined together, parts of one image are inserted inside the other that, basically alters the image and thereby poses problems in their forensic validation. The object removal is manipulation where some objects or elements in an image are simply erased and replaced with some other object, but that leaves subtle inconsistencies in the texture or the lighting of the image. CNN architecture; its architecture has the capability to automatically learn hierarchical features from images and capture both local as well as global visual patterns.



Orginal Image



Copy Move Image

Fig. 3. Copy-Move Forgery Image

Image Tampering Detection Dataset trains the CNN model, which consists of various examples, including authentic as well as manipulated ones. It learns these different patterns and then can identify the unique patterns indicated forgery. Such pre-processing techniques as rescaling and augmentation enhance the model in robustness and in its generalization capability. During training, grid search of hyperparameters is applied and cross-validation on the performance of models under different scenarios. The model is highly accurate for real versus tampered images, since it would be able to identify such fine details that may have relationships with each form of forgery. This framework



offers a powerful and automated method for image tampering detection with potential uses in media forensic and digital content verification applications.

#### Distribution of Tampering Types

The Fig. 4 depicts the frequency of different types of tampering in the dataset. Here, x-axis is the various types of tampering. The y-axis gives the number of occurrence times for each type of tampering. From the chart, it can be noticed that the most occurring type is 'None' for which it has been seen thrice. 'Splicing' and 'Copy-Move' are seen twice, whereas 'Object Removal' happens only once. This would show that most of the images in the dataset are authentic, with only a few having been tampered with.





The bar chart would, in fact, reflect this imbalance, such that most of it shows non-tampered data. It is handy to know the general trend between tampered and non-tampered images. This might be beneficial when analyzing further or training models to detect tampering. The 'None' category greatly outweighs the categories of tampered images and portrays how rare the act of tampering.

#### Performance Metrics

The key metrics are used to evaluate the performance of this image forgery detection system. Accuracy measures the overall percentage of correct predictions; Precision gives the proportion of true positives among all identified forgeries represented in Eqn (3). Recall evaluates how good the model is at detecting all forgeries represented in Eqn (4), and F1-Score gives a balanced measure of both precision and recall represented in the Eqn (5) and (6).

$$Accuracy = \frac{\text{True Negative+True Positive}}{\text{TruePositive+FalsePositive+TrueNegative+FalseNegative}}(3)$$

$$Precision = \frac{\text{T*p}}{\text{T*p+F*n}} \qquad (4)$$

$$Recall = \frac{\text{T*p}}{\text{T*p+F*n}} \qquad (5)$$

$$F1 \text{ score} = \frac{2\text{T*p}}{2\text{T*p+F*p+F*n}} \qquad (6)$$

Besides, ROC Curve and Area Under the Curve are used to test the balance between true positives and false positives. The use of metrics makes sure that the model is good at identifying all types of tampered images.

#### Performance Evaluation of the Proposed Framework

The Fig. 5 represents the major performance measures of the proposed CNN-based image forgery detection framework. Each measure highlights the key aspect of the validity of the respective model. While the overall correctness is very high at an Accuracy of 95%, the framework can correctly identify genuine and forged images 95% of the time. Accuracy is 93%, which means the model's doing a very good job of classifying images as forgeries, so pretty high specificity when it comes to detecting.





#### Fig. 5. Performance Metrices of the proposed Framework

Recall is 94%, meaning the model's pretty good at catching the real forgeries; meaning it really catches a lot of manipulated images. The F1 Score at 93.5%. In combination, these high metrics indicate the strength and efficiency of the proposed framework in reporting a good number of techniques with reliable detection, and the application of such a framework fits well in those applications that are primarily involved in media forensics and digital content verification.

#### Performance Comparison

The proposed CNN framework significantly outperforms the existing methods, including SVM, Random Forest, and K-NN, in all key performance metrics. It achieves 95% accuracy, 93% precision, 94% recall, and 93.5% F1-Score, surpassing the SVM's 85%, Random Forest's 87%, and K-NN's 80% as shown in the Table 1.

Method	Accuracy	Precision	Recall	F1-Score
(SVM)	85%	83%	84%	83.50%
(Random Forest)	87%	85%	86%	85.50%
(K-NN)	80%	78%	79%	78.50%
Proposed Framework	95%	93%	94%	93.50%

 TABLE 1: PERFORMANCE COMPARISON WITH EXISTING FRAMEWORK

The CNN model outperforms at detecting subtle forgeries such as object removal, which the other models fail to capture. Improvement is due to the CNN ability to learn hierarchical features and detect complex patterns within images that traditional methods, such as SVM and Random Forest, fail to do.

#### Discussion

The superior performance of CNN framework highlights the necessity of deep learning strategies to handle complex activities such as image forgery. While SVM and Random Forest would have the capacity to detect low-level manipulation they are not capable of unlocking more complicated forgery activity, such as object removal. This ability of CNN frameworks to learn from large sets enables them to generalize even better, and, hence be reliable in operational settings. This indicates a trend toward more complex and data-intensive systems for the detection of digital forgeries.



#### VI. CONCLUSION AND FUTURE WORKS

The image forgery detection framework on CNNs shows its superiority over traditional methods, providing an identification solution for different manipulations occurring in images. Its high accuracy, recall, and precision make it a promising tool for media forensics, digital security, and content verification. The future work will be the optimizations of the CNN model in order to reduce its computational complexity for direct use in real-time applications. Finally, one of the robustness enhancements using transfer learning and fine-tuning on various datasets in order to be able to detect new and unseen forgery types.

#### REFERENCES

[1]F. Z. Mehrjardi, A. M. Latif, M. S. Zarchi, and R. Sheikhpour, "A survey on deep learning-based image forgery detection," Pattern Recognition, p. 109778, 2023.

[2]X. Guo, X. Liu, Z. Ren, S. Grosz, I. Masi, and X. Liu, "Hierarchical fine-grained image forgery detection and localization," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 3155–3165. Accessed: Nov. 13, 2024. [Online]. Available: http://openaccess.thecvf.com/content/CVPR2023/html/Guo Hierarchical Fine-

Grained Image Forgery Detection and Localization CVPR 2023 paper.html

[3]Z. N. Khudhair, F. Mohamed, and K. A. Kadhim, "A review on copy-move image forgery detection techniques," in Journal of Physics: Conference Series, IOP Publishing, 2021, p. 012010. Accessed: Nov. 13, 2024. [Online]. Available: https://iopscience.iop.org/article/10.1088/1742-6596/1892/1/012010/meta

[4]S. S. Ali, I. I. Ganapathi, N.-S. Vu, S. D. Ali, N. Saxena, and N. Werghi, "Image forgery detection using deep learning by recompressing images," Electronics, vol. 11, no. 3, p. 403, 2022.

[5]H. Wu, J. Zhou, J. Tian, and J. Liu, "Robust image forgery detection over online social network shared images," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 13440–13449. Accessed: Nov. 13, 2024. [Online]. Available: http://openaccess.thecvf.com/content/CVPR2022/html/Wu\_Robust\_Image\_Forgery\_Detection\_Over\_Online\_Social\_N etwork Shared Images CVPR 2022 paper.html

[6]R. Agarwal, D. Khudaniya, A. Gupta, and K. Grover, "Image forgery detection and deep learning techniques: A review," in 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), IEEE, 2020, pp. 1096–1100. Accessed: Nov. 13, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9121083/

[7]E. U. H. Qazi, T. Zia, and A. Almorjan, "Deep learning-based digital image forgery detection system," Applied Sciences, vol. 12, no. 6, p. 2851, 2022.

[8]M. Zanardelli, F. Guerrini, R. Leonardi, and N. Adami, "Image forgery detection: a survey of recent deep-learning approaches," Multimed Tools Appl, vol. 82, no. 12, pp. 17521–17566, May 2023, doi: 10.1007/s11042-022-13797-w.

[9]N. K. Rathore, N. K. Jain, P. K. Shukla, U. Rawat, and R. Dubey, "Image Forgery Detection Using Singular Value Decomposition with Some Attacks," Natl. Acad. Sci. Lett., vol. 44, no. 4, pp. 331–338, Aug. 2021, doi: 10.1007/s40009-020-00998-w.

[10]A. Sahani and K. Srilatha, "Image Forgery Detection Using Svm Classifier," vol. 3, no. 3.

[11]M. Ahmad and F. Khurshid, "Digital Image Forgery Detection Approaches: A Review," 2021, pp. 863–882. doi: 10.1007/978-981-33-4604-8 70.

[12]J. Khan, "Random Forest based Copy-Move Forgery Detection," 2020, Accessed: Nov. 13, 2024. [Online]. Available: https://osf.io/preprints/vc42d/

[13]S. Uma and P. D. Sathya, "Forgery Detection of Digital Images Using Teaching–Learning Based Optimization and Principal Component Analysis," Sens Imaging, vol. 24, no. 1, p. 19, Jun. 2023, doi: 10.1007/s11220-023-00424-7.

[14]N. T. Pham, J.-W. Lee, and C.-S. Park, "Structural Correlation Based Method for Image Forgery Classification and Localization," Applied Sciences, vol. 10, no. 13, Art. no. 13, Jan. 2020, doi: 10.3390/app10134458.

[15]N. K. Rathore, N. K. Jain, P. K. Shukla, U. Rawat, and R. Dubey, "Image Forgery Detection Using Singular Value Decomposition with Some Attacks," Natl. Acad. Sci. Lett., vol. 44, no. 4, pp. 331–338, Aug. 2021, doi: 10.1007/s40009-020-00998-w.

[16]P. Bevinamarad and P. H. Unki, "Robust Image Tampering Detection Technique Using K-Nearest Neighbors (KNN) Classifier," in Innovations in Computational Intelligence and Computer Vision, vol. 1424, S. Roy, D. Sinwar, T. Perumal, A. Slowik, and J. M. R. S. Tavares, Eds., in Advances in Intelligent Systems and Computing, vol. 1424. , Singapore: Springer Nature Singapore, 2022, pp. 211–220. doi: 10.1007/978-981-19-0475-2\_19.



[17]B. Diallo, T. Urruty, P. Bourdon, and C. Fernandez-Maloigne, "Robust forgery detection for compressed images using CNN supervision," Forensic Science International: Reports, vol. 2, p. 100112, 2020.

[18]"CASIA 2.0 Image Tampering Detection Dataset." Accessed: Nov. 13, 2024. [Online]. Available: https://www.kaggle.com/datasets/divg07/casia-20-image-tampering-detection-dataset?resource=download

[19]A. Singh and J. Singh, "Image forgery detection using deep neural network," in 2021 8th International conference on signal processing and integrated networks (SPIN), IEEE, 2021, pp. 504–509. Accessed: Nov. 13, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9565953/

[20]E. U. H. Qazi, T. Zia, and A. Almorjan, "Deep learning-based digital image forgery detection system," Applied Sciences, vol. 12, no. 6, p. 2851, 2022.



INTERNATIONAL STANDARD SERIAL NUMBER INDIA







## **INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH**

IN COMPUTER & COMMUNICATION ENGINEERING

🚺 9940 572 462 应 6381 907 438 🖂 ijircce@gmail.com



www.ijircce.com