ISSN 2347 - 3983
Volume 10. No.4, April 2022
International Journal of Emerging Trends in Engineering Research
Available Online at http://www.warse.org/IJETER/static/pdf/file/ijeter091042022.pdf
https://doi.org/10.30534/ijeter/2022/091042022

Digital Image Forgery Detection Using Deep Learning Models

S.Velmurugan[1], Dr.T.S.Subashini[2]
[1]ResearchScholar, Department of Computer and Information Science, Annamalai University, Tamil Nadu, India
Emial: ssvelmurugan@yahoo.com
[2]Professor Department Computer Science and Engineering, Annamalai University, Tamil Nadu, India
Email: rtramsuba@gmail.com

Received Date : March 4 , 2022 Accepted Date : March 24, 2022 Published Date : April 07, 2022

ABSTRACT

One of the challenges to image trust in digital and online apps, as well as on social media, is the current situation. Image forgery detection is a technique for detecting and locating fabricated components in a modified image. A sufficient amount of features is necessary for good image forgery detection, which can be achieved using a deep learning model that does not require human feature engineering other handcraft feature techniques. In this paper we used the GoogleNet deep learning model to extract picture features and the Random Forest machine learning technique to determine whether or not the image was fabricated. The proposed approach is implemented on the publicly available benchmark dataset MICC-F220 with k-fold cross validation approach to split the data set in to training and testing dataset and also compared with the state-of-the-art approaches.

Key words : Digital image forensics, Tampered Image Identification, Random Forest, GoogleNet.

1. INTRODUCTION

Digital images are being used in various spheres of real-time applications like media, military, science, law, education, politics, medical imaging and diagnosis, art piece, digital forensics, intelligence, sports, photography, social media, scientific publications, journalism, and business [1]. Digital images become a significant resource of information in the digital world as they are the fastest means of information and medium of communication. In recent years, forged images have affected the above-mentioned application areas [1]. Digital image acts a significant part of different technologies and fields. The use of digital cameras, personal computers, and sophisticated image processing software are available for modification and for manipulation of images. These tools are scalable and provide user interface features. Manipulating and tampering the images today can be effectively accomplished not only by specialists but also by novice users. These tampered images are not recognizable and so real in perception in a way that authenticity is lost [2].Therefore, integrity and authenticity verification of images has gained researchers attention in the image processing field. The approaches to detect any type of tampering are categorized into active and passive approaches [3] [4] as shown in the figure 1.

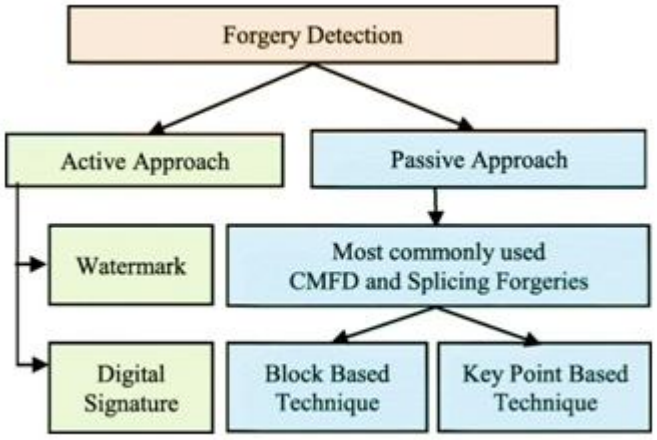

Figure 1: Approaches for Image Forgery Detection

Inactive approaches images need to be protected through digital signature or through watermarking techniques where as passive approaches do not require any kind of pre-embed operation of digital signature or watermarking. The drawback of active approaches is that it needs to pre pre-embedded either with digital signature or with watermarking, whereas a large number of images present today on web, social media and other applications are not active in nature[1].Thus we have focused on the detection of forgery with a passive approach which is described further in given sections. The contribution of this paper is to apply the GoogleNet [5] deep learning model for automatic feature extraction and to implement the Random Forest machine learning algorithm to

detect whether the image is forged or not. The organization of the paper is as follows: Section II highlights the recent and related approaches for Image forgery detection using deep learning and machine learning. Section III explains the proposed approach and section IV evaluates the performance of the proposed approach and section V ends with the conclusion and future scope.

2. RELATED WORK

The extraction of handcrafted features is used in most picture fraud detection approaches in the literature, including geometrical based, wavelet based, statistical based [6], key-point based, block based, transformations based, texture based, and so on. The majority of the features produce good results, however they are not resistant to various geometrical and post-processing processes for various sorts of picture forgeries. To improve the accuracy [7] of image forgery detection, some researchers utilized [8] machine learning, deep learning and convolution neural network [9] (CNN) based approaches [10]-[11]. In [12] authors proposed an approach for image forgery detection using Scale Invariant Features Transform (SIFT) features for the dataset MICC-F220 and MICC-F2000 and able to deal with affine geometric transformations. The False Positive Rate (FPR) and True Positive Rate (TPR) achieved is 8% and 100% respectively. In [13] authors proposed an image forgery detection approach using speeded up robust features (SURF) [17] and hierarchical agglomerative clustering (HAC) for the dataset MICC-F220. In [14] the approach is based on discrete cosine transform (DCT) [17] features for each block and through lexicographical sorting of block-wise DCT coefficients forgery of the image is detected. This approach is only able to identify forgery with small variations in scaling and rotation. In [15] authors applied PCA on image blocks to reduce the dimension space and performed lexicographical sorting and robust to minor variations in the image due to lossy compression or additive noise. In [6] authors proposed the modified version of CNN to detect cut and paste forgery. A filter layer was added before the first convolutional layer to take an image as its input and output the Median Filtering Residual (MFR) of the image.

The proposed method learned hierarchical features representation automatically with low false rate and high detection rate. In [7] authors stated automated hierarchical feature representations learning model to detect splicing and copy-move forgeries. They proposed the CNN model with 8 convolutional layers and a fully connected layer with a 2-way classifier.

In [16] presented the two-stage deep learning approach using the Stacked Auto encoder (SAE) model for the detection of forged images. In [11] authors presented the CNN model with a blocking strategy for image forgery detection. Firstly, the image was divided into blocks using tight blocking and marginal blocking. Then, the blocks were inputted in to the rich model Convolutional Neural Network (rCNN). At last, the pooling was performed, followed by the classification of the input image based on the feature vectors using the SVM classifier.

3. PROPOSEDAPPROACH

This proposed approach is using the hardware as Intel (i5) CPU with 2.9 GHZ, 8GB RAM, GPU and software as Windows Pro OS with Matlab release R2016R.

3.1 Dataset

In this section, MICC-F220 [12] publicly available benchmark dataset is used for the experimental result. This dataset consists of 110 non-forged and 110 forged with 3 channels i.e. color images of size 722 × 480 to 800 × 600 pixels with 5 various geometrical and transformational attacks are used. Figure 2 shows the original image, while Figure 3 shows a range of geometrical and transformational images. This dataset is used for the detection of forged images where cloned or copy-move forgery is carried out.

Figure 2: Original Image

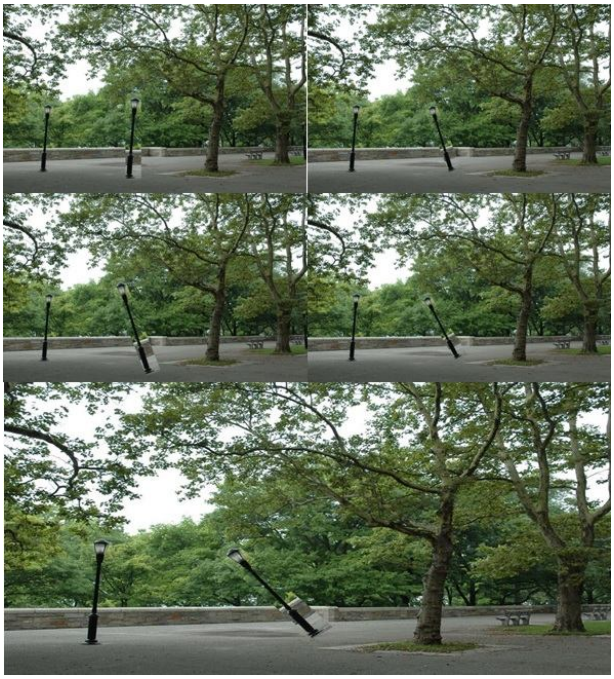

Figure 3: Various geometrical and transformational attacks

3.2 Machine Learning Algorithm

Random Forest is one of the widely used and popular algorithms in machine learning. This can be used as both regression and classification techniques. Random Forest is the forest of decision trees. A dataset is divided into uniform subsets repeatedly for calculating the class membership through DT classifier. In every intermediary state, the acceptations and rejection of class labels are achieved through the hierarchical classifier. The node partitioning, identification of term in all nodes and allocating the class label to leaf nodes are the three major parts of the decision tree. While taking decisions or prediction, the majority of votes by decision trees are taken into consideration.

3.3 Approach

In this approach k-fold cross validation approach is used with the k value as 5 for dividing the dataset into training and testing. Google Net issued to extract the features to train the Random Forest machine learning algorithm as shown in Figure 4.

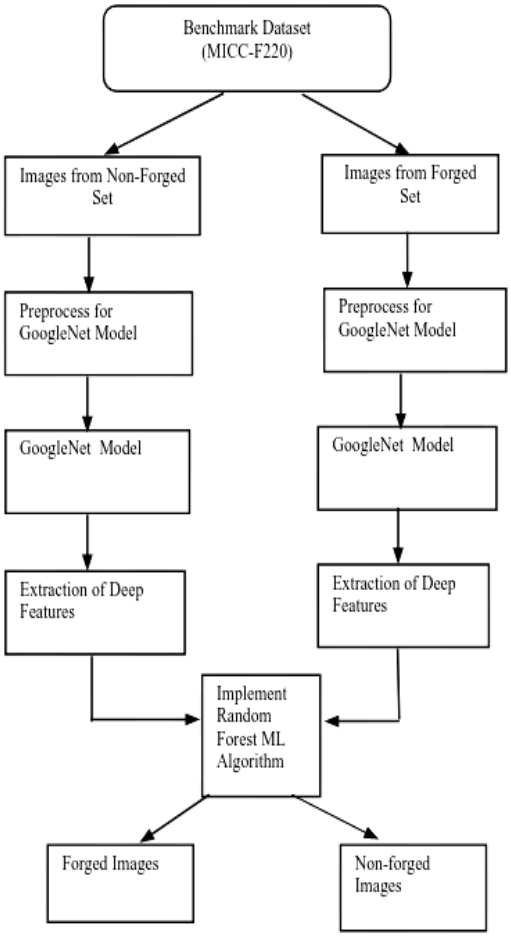

Figure 4: The proposed image forgery detection technique's process

4. RESULTS AND DISCUSSION

The following performance measures were calculated using the confusion matrix. Here TP (true positive) refers to forged images correctly identified. An image misidentified as forged when they are not refers to FP (False Positive). Original images that are correctly identified are termed as TN (True Negative) and those identified as forger are termed as FN (False Negative).

1. Recall (R) also known as Sensitivity or True Positive Rate (TPR) also) measures how well the copy moved regions are correctly identified and is given as

$$R = \frac{TP}{TP + FN}$$

2. Precision (P), also known as Confidence and Positive Predictive Value, is a measure of the accuracy of a prediction. The probability that a detected copy moved region is accurate is measured by PPV.

$$P = \frac{TP}{TP + FP}$$

3. Accuracy measures the proportion of the total number of predictions that are correct.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

4. F-1 score measures the harmonic-mean of Precision (P) and Recall (R) and is given by

$$F1 = 2\frac{P * R}{P + R}$$

.

Where TP-True Positive, FN-False Negative, FP-False Positive, and TN-True Negative values respectively. The Confusion Matrix of the predicted class and the actual class is computed for the evaluation of the proposed method as shown in Table 2. It is observed that the accuracy is 89.55%, Precision is 85.95%, TPR is 94.54%, FPR is 15.45%, F1score is 90.04% and execution time is 0.43 sec with Area under Curve (AUC) is 55.92%.Figure 5 shows the ROC Curve for Random Forest machine learning algorithm for theMICC-F220 Dataset.

Table 1: Confusion matrix for assessing the performance metrics

	Predicted Images		
Actual Images	Image Dataset	Forged	Non-Forged
	Forged	47.27%	2.73%
	Non-Forged	7.72%	42.28%

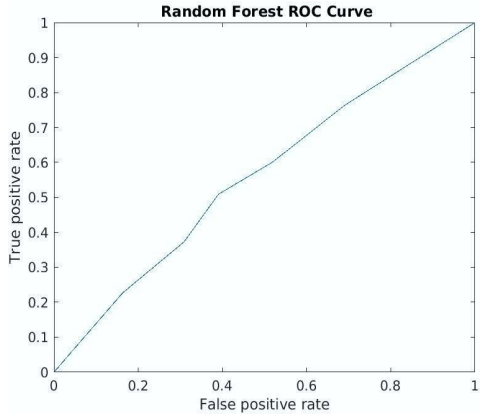

Figure 5: ROC Curve for Random Forest machine learning algorithm for theMICC-F220Dataset

Table 2: Comparison of performance metrics with other approaches

Approach	FPR, %	TPR, %	Time'(ms)
Amerinietal.[12]	8	100	4.94
Mishraetal.[13]	3.64	73.64	2.85
Fridrichet al.[14]	84	89	294.69
Popescu & Farid[15]	86	87	70.97
Proposed Approach	15.45	94.54	0.43

5. CONCLUSION

In this paper, Random Forest machine learning algorithm is implemented on the extracted features of digital images using GoogleNet deep learning model for image forgery detection. The proposed approach achieves better results as compared to the state-of-the-art approaches. As a future work, more machine learning algorithms and other emerging deep learning models can be explored and implemented for image forgery detection with other publicly available benchmark datasets.

ACKNOWLEDGEMENT

The corresponding author S.Velmurugan here by acknowledge the funding support provided to him by DST-PURSE-Phase II Annamalai University, Tamilnadu, India in carrying out this study as part of this Ph.D Research work.

REFERENCES

[1] Doegar,A., Dutta,M., & Kumar,G .(2019).“A review of passive image cloning detection approaches”, Proceedings of2nd International Conference on Communication, Computing and Networking(pp.469-478). Springer, Singapore.

[2] Chauhan,D., Kasat,D., Jain,S. and Thakare,V., "Survey onkeypointbasedcopy-moveforgerydetectionmethodson image",ProcediaComputer Science85(2016)January1.

[3] K.Asghar, Z.Habib, and M.Hussain, “Copy-move and splicing image forgery detection and localization techniques :a review”, Australian Journal of Forensic Sciences.,vol. 49,no.3, (2017)pp. 281–307.

[4] Doegar, A., Dutta, M., & Gaurav, K.,"CNN based Image Forgery Detection using pre-trained AlexNet Model", International Journal of Computational Intelligence &IoT.,vol.2,no.1, (2019).

[5] S zegedy C, LiuW, JiaY, SermanetP, ReedS, Anguelov D,Erhan D, Vanhoucke V, Rabinovich A.,“ Going deeper with convolutions”, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,(2015).

[6] J.Chen, X.Kang, Y.Liu, and Z.J.Wang, “Median Filtering Forensics Based on Convolutional Neural Networks”, IEEE Signal Processing Letters., vol. 22, no. 11,(2015) pp.1849–1853.

[7] Rao, Yuan, and Jiangqun Ni. "A deep learning approach to detection of splicing and copy-move forgeries in images", Proceedings of (2016) IEEE International Workshop on Information Forensics and Security (WIFS),2016.

[8] Zhang J., Liao Y., Zhu X., Wang H., and Ding J., “A Deep Learning Approach in the Discrete Cosine Transform Domain to Median Filtering Forensics”, IEEE Signal Processing Letters.,vol.27,(2020), pp. 276–280.

[9] ZhangY., GohJ., WinL.L., and Thing V., “Image Region Forgery Detection: A Deep Learning Approach”, Cryptology and Information Security Series.,vol.14, (2016),pp. 1–11.

[10] J.Zhong and C.Pun, "An End-to-End Dense-Inception Net for Image Copy-Move Forgery Detection", IEEE Transactions on Information Forensics and Security, vol. 15,(2020)pp. 2134-2146.

[11] ZhouJ.,NiJ., RaoY., "Block-Based Convolutional Neural Network for Image Forgery Detection", EditedKraetzerC.,ShiYQ.,DittmannJ.,KimH.,DigitalForensicsandWatermarking.IWDW2017.LectureNotesin ComputerScience,Springer,vol10431,(2017),pp. 65-76.

[12] I.Amerini, L.Ballan, R.Caldelli, A.Del Bimbo, and G. Serra, “A SIFT-based forensic methodforcopy-moveattackdetectionandtransformation recovery”,IEEE Transactions on Information Forensics and Security., vol.6,no.3 (2011), pp.1099–1110.

[13] P.Mishra, N.Mishra, S.Sharma, and R.Patel, "Region duplication forgery detection technique based on SURF and HAC", The Scientific World Journal.,(2013).

[14] Fridrich,A.J.,Soukal,B.D.,&Lukáš,A.J.,"Detectionofco py-moveforgeryindigitalimages",ProceedingsofDigital ForensicResearchWorkshop(2003).

[15] Popescu, A.C. and Farid,H.,2004."Exposing digital forgeries by detecting duplicated image regions", Dept. Comput. Sci., Dart mouth College,Tech.Rep.TR2004-515, (2004), pp. 1-11.

[16] ZhangY., GohJ., WinL.L., and Thing V., "Image Region Forgery Detection:A Deep Learning Approach",Cryptology and Information SecuritySeries.,vol.14,(2016), pp. 1–11.

[17] S. Velmurugan, T.S. Subashini, M.S. Prashanth, "Dissecting the literature for studying various approaches to copy move forgery detection", International Journal of Advanced Science and Technology (IJAST) 29 (04), (Jun. 2020) 6416–6438.