All Sciences Proceedings

5th International Conference on Applied Engineering and Natural Sciences

July 10-12, 2023 : Konya, Turkey

All Sciences Proceedings
http://as-proceeding.com/

https://www.icaens.com/

Framework for Localization of Forgery Regions in Image

Canberk Şahin[1], Mustafa Özden[2]

[1]Department of Electrical and Electronics Engineering, Bursa Technical University, Bursa, 16310, Turkey
[2]Department of Electrical and Electronics Engineering, Bursa Technical University, Bursa, 16310, Turkey

[1]20278034074@ogrenci.btu.edu.tr
[2]mustafa.ozden@btu.edu.tr

Abstract **–** With the development of computer technologies, manipulations are made on digital images without leaving a clear trace thanks to image processing software. There is a great need for applications to detect forgered images made with malicious intent in many fields such as politics, law, medicine and military. Many studies have been carried out and various algorithms have been developed to detect forgered regions by detecting forgered images. Today, superior methods are developed by combining traditional image forgery techniques with deep learning techniques. In this study, the Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT) were used together with Convolutional Neural Networks (CNN) to locate the forgered regions of the forgered images. Three different methods were made to locate the forgered region. In the first method, DWT and CNN were used together. In the second method, DCT and CNN were used together. In the last method, DCT and DWT were combined in parallel and used together with CNN.

Keywords – Image forgery, Convolutional Neural Network (CNN), Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), Long-Short Term Memory (LSTM), Intersection Over Union (IOU).

I. INTRODUCTION

This With the rapid development of technology, digital images are used in many areas such as social media, health, law and politics, depending on the increase in the use of computers and smart phones. Modern image processing tools and various computer graphics software facilitate manipulation of digital images. Image manipulation can be done for good entertainment purposes, such as making images look better in order to gain more likes on social media or attract viewers who gain popularity, as well as malicious intentions such as political gain or operate malicious attacks, spreading incorrect information. If digital images are digitally changed using some image editing software and used for unethical purposes, this process is called image forgery. With the development of image processing software, images are manipulated without leaving a trace, so it is difficult to detect forgered images. The use of forgered images in areas such as politics and law negatively affect people's trust. In this context, the main purpose of image forensics is to detect the forgered images that are difficult to detect. Therefore, there is still a need to develop effective systems to detect image forgery.

Image forgery detection aims to verify the originality of a digital image. Image forgery detection is divided into two methods as active and passive [1]. Digital Watermarking and Digital Signature are the two basic methods used for active image forgery detection. Digital images manipulated with digital watermarks or digital signatures use a known authentication code embedded in the image content before being sent through an untrusted public channel. In order to understand that an image is authentic with active methods, it is determined by verifying whether the extracted watermark or digital signature matches the original one. The major disadvantage of digital watermarks or signatures is that they require specially equipped devices to place the watermark-signature during digital image creation. Passive forgery detection methods are based on detecting forgery images by considering the pixel distortions

of the images and statistical ratios in image consistency, and there is no need for any signature or watermark of the original image. The most popular passive methods are copy-move forgery and image splicing. Copy-move forgery is created by copying a specific region from an original image and pasting it to a location in the same image. In the image splicing method, a forgered image is created by clipping a region from one image and adding it to another image.

Passive methods generally extract features from the images as the first step, then select a classifier and train the classifier using the extracted features from the images, and finally classify the images as forgered or original [2]. In order to detect forgered images, the pixels in the manipulated part of the image should be evaluated statistically by extracting the features from the images. For this reason, researchers examined the pixel correlations of images using DCT technique [3] – [4]. In passive image forgery detection methods, the very high dimensions of feature maps are a disadvantage in detecting forgered images. By using DCT and DWT, the researchers reduced the size of the feature maps and increased the success of detecting forgered images [5].

Recent years have shown that CNN's can reveal complex statistical dependencies from high-dimensional sensor inputs such as cameras and efficiently learn their hierarchical representations, allowing it to generalize well to a wide variety of computer vision tasks, including image classification. Rao et al. proposed a new image forgery detection method that can automatically learn feature representations of images based on deep learning techniques [6].

Copy-move forgery is one of the most active research areas in forensics, as it is one of the most widely used methods. Researchers have turned to deep learning methods using databases such as ImageNet instead of using traditional block-based passive methods used to detect image forgery. Ouyang et al. proposed deep convolutional neural network-based copy-move image forgery detection method [7]. According to the results, the proposed method shows good results in computer generated forgered images, but it does not achieve successful result in detecting a forgered image generated based on the real scenario.

Cropping a region from one image and pasting it into another image is another method frequently used in image forgery. Due to the problems such as deep learning method-based image splicing detection, the proposed systems use high-dimensional feature vectors and it takes a long time for computers to train these feature vectors, researchers have used methods such as DCT and DWT to reduce the size of the feature vector. El-Latif et al. proposed an algorithm to detect image splicing using DWT-based CNN structure [8]. With the good results obtained in deep learning-based image forgery algorithms, studies have been carried out to detect the forgered region in the manipulated image. In order to detect forgered regions, the features of the manipulated regions were extracted and trained with various CNN structures. Manipulated regions were determined by giving images to the trained models [9] – [10].

In this study, a method that uses DCT and DWT together with CNN structure for feature extraction from images is proposed to detect manipulated regions of forgered images. For this purpose, three different models have been developed. In the first model, manipulated sections were determined using DCT and CNN structure. In the second model, manipulated sections were determined using DWT and CNN structure. In the last model, manipulated regions were determined with the proposed method consisting of CNN structure in which DCT and DWT are combined. Performance of the models were compared using precision, recall, F-measure and IOU metrics.

The remainder of this paper is organized as follows. The methodology of the proposed method is presented in detail in Chapter II. Experimental results with forgered images reserved for testing are presented in Chapter III. In Chapter IV, the success of the proposed method is compared with the success of different methods and the performance results are shown.

II. Proposed Method for Image Forgery Localization

A. Data Set in the Study

In the training phase of the proposed method in this study, the dataset named "Splited Copy" was used. In the dataset, there are 26000 forgered images made by image splicing as shown in Fig. 1. and copy-move forgery methods as shown in Fig. 2. In addition, there are mask images in black and white format consisting of the manipulated region prepared for each of the forgery images.

Fig. 1 An image created with image splicing forgery method (left) and the mask of the manipulated region of the image (right)

Fig. 2 An image created with copy move forgery method (left) and the mask of the manipulated region of the image (right).

After the model was trained, it was tested with the forgery images in the CASIA V.2.0 dataset. In addition, forgery images were obtained by manipulating random images downloaded from the internet that were not included in the data sets. The trained system was tested with forgery images not included in any dataset.

B. *Image Preprocessing*

Due to limited computation power and limited memory resources, computers accept images of certain sizes as inputs to existing CNN structures. Therefore, in the study, the images in the data set were resized to be 256 x 256 pixels. Color images were converted to gray level images in order to calculate the coefficients with DCT and DWT methods.

Since the aim of the study is to detect the manipulated location in the forgered images, each image is divided into patches of 32 x 32 pixels before being input to the CNN (Fig. 3). The reason for this is to improve feature extraction from manipulated regions when training with masks of manipulated regions. Thus, when a forgery image that is not used in the data set is given to the trained network, the location of the manipulated region can be estimated more precisely.

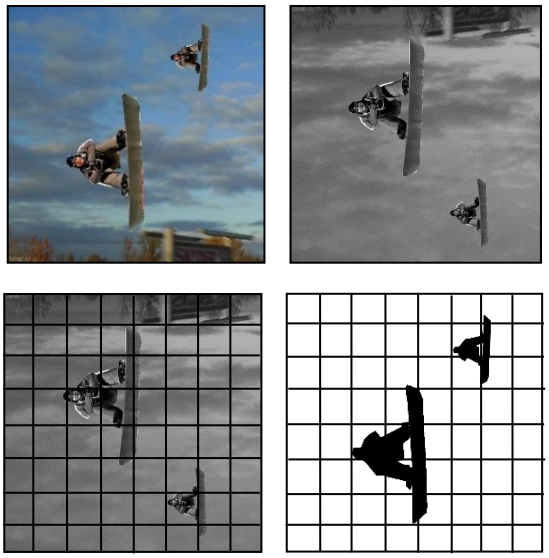

Fig. 3 Converting color images to gray level and divided into 64 patches

A total of 64 patches were obtained from an image by dividing the 256 × 256 pixel images into 32×32 pixel patches. In the same way, the masks of the images were also divided into patches and used in the training of the model.

C. *Discrete Cosine Transform*

Discrete cosine transform (DCT) is a conversion technique widely used in signal processing and image compression. It is closely related to the Fourier transform and is designed to analyze and represent images in terms of frequency components [11]. DCT converts a signal or image from spatial domain to frequency domain by decomposing it into the sum of cosine functions of different frequencies. The resulting DCT coefficients represent the frequency content of the image (Fig. 4) lower frequencies are concentrated at lower frequency coefficients and higher frequencies are concentrated at higher frequency coefficients.

Fig. 4 Obtaining coefficients (right image) by applying DCT to a forgered image (left image).

The combination of CNNs and DCTs when detecting forgery images takes advantage of the capability of both techniques to increase the accuracy and success rate of the forgered image detection methods. CNN structures are powerful

deep learning models that are good at learning complex image features and patterns and are effective at detecting general inconsistencies or manipulations in an image. However, CNNs are not good enough in precisely localization the manipulated regions in the image. On the other hand, DCT is a frequency-based transform that captures low-level details and structures in the frequency domain. By incorporating the DCT method, which analyzes the local frequency characteristics, into the CNN structures, better localization capabilities are gained in the forgered detection method. DCT helps identify specific frequency patterns or artifacts that are indicative of manipulated regions, thus improving the system's ability to localize regions where forgery is occurring.

D. Discrete Wavelet Transform

Discrete wavelet transform (DWT) is a mathematical method used to analyze signals and images by separating them into different frequency bands. DWT is particularly effective at capturing both frequency and spatial information simultaneously, making it useful in a variety of applications such as image processing, image compression and image noise reduction [12]. DWT decomposes a signal or image into a series of wavelet coefficients representing different frequency components at different scales or resolutions. This decomposition is achieved by applying a series of high-pass and low-pass filters to the signal, and then downsampling. The process is iteratively applied to approximation coefficients that reveal filters and subsampling, resulting in a multi-resolution analysis. The DWT method is applied to images by decomposing the images into a series of wavelet coefficients representing different frequency components at different scales or resolutions.

The results of DWT on an image are sets of approximation and detail coefficients at different scales. The approximation coefficients represent the low-frequency components of the image, capturing the overall structure and smooth variations. Detail coefficients represent high-frequency components that capture sensitive details and variations. DWT provides a multi-scale representation of the image where higher scales correspond to sensitive details and lower scales capture coarser details. This representation enables images to be analyzed and manipulated at different scales, facilitating tasks such as feature extraction, noise removal, compression, and forgery detection (Fig.5).

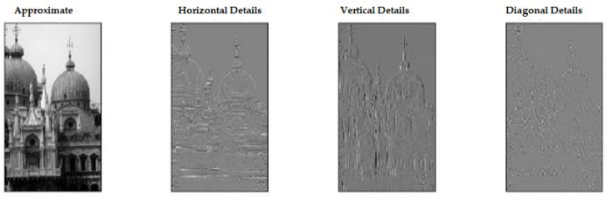

Fig. 5 Applying the discrete wavelet transform to a forgered image and obtaining horizontal, vertical and diagonal details.

In forgery detection, the DWT coefficients are analyzed to identify inconsistencies, defects or changes that occur during image manipulation. Deviations or anomalies in the detail coefficients at different scales can indicate manipulated regions or forgery, while approximation coefficients provide contextual information about the overall structure of the image. Forgery detection algorithms can identify suspicious areas by comparing the DWT coefficients of an image with an original or unmodified reference image. It can also detect various types of manipulation, such as copy-paste resizing and blending.

The combination of DWT and CNNs in forgery image detection combines multi-scale frequency-based analysis, spatial and semantic analysis, enabling the accurate detection of manipulated regions in images. The frequency-based analysis provided by DWT helps capture a wide variety of manipulations, including subtle changes or tampering, that might be difficult to detect with CNNs alone.

E. Designing and Training the CNN Structure

Convolutional Neural Networks are a class of deep learning models specifically designed for processing data such as images. CNNs consist of multiple layers, including convolutional layers, pooling layers, and fully connected layers. The basic operation in CNNs is convolution, which involves applying a series of filters to the input data, detecting patterns and spatial relationships, and effectively extracting features. Pooling layers help reduce the spatial dimensions of data by reducing computational complexity. Finally, fully connected layers perform high-level querying and make predictions based on extracted features [13].

In this study, a new structure that is different from traditional CNN structures, which consists of

convolution layers, DCT and DWT layers, as well as long short-term memory (LSTM) layers, is designed. In the designed structure, convolution layers, DCT and DWT layers work in parallel. The Fig.6 shows the designed sutructure.

In the proposed method developed using this sutrucure, color images are given as 32 × 32 pixels input to CNN after image preprocessing. There are 6 convolutional layers in the CNN structure. After 3 convolution layers in the CNN structure, the output data is resized and given as input to three 32 × 32 size long-term memory layers in total. Convolution layers with 32 filters are applied to the data that is the output of the LSTM layers. DCT and DWT layers working in parallel with the data coming out of 3 convolution layers were applied. DCT and DWT layers were used to extract the features of the images in the frequency domain as described in the previous sections. The mask of the manipulated region of a forgered image is estimated from the features extracted from all layers.

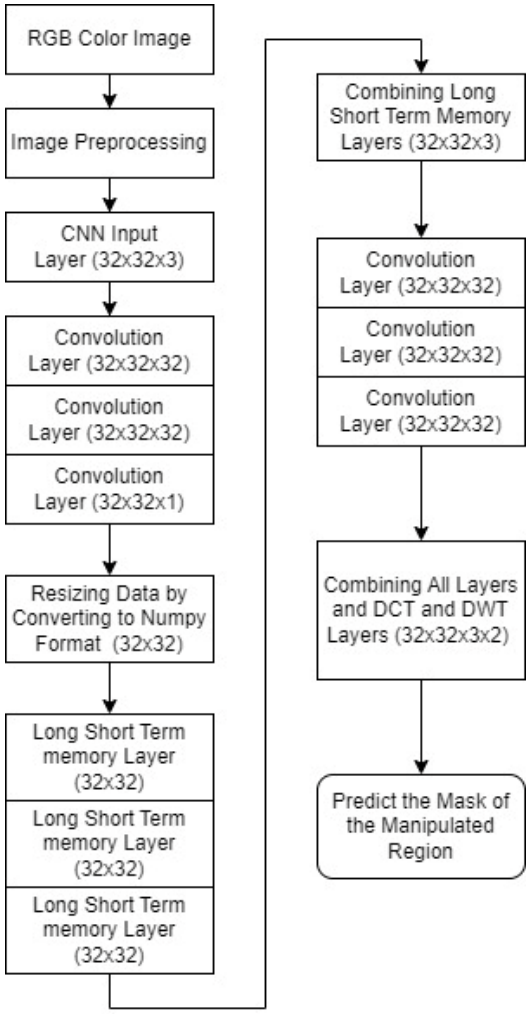

Fig. 6 The structure of the proposed method with the designed CNN layers.

For the training of the proposed method, preprocessing was done on the images. There are a total of 26000 forgered images in the data set used. A total of 6000 images were randomly selected from the data set. Each image is cropped and resized to 256 × 256 pixels and divided into patches of 32 × 32 pixels. The same operations were performed on the mask images. The resulting images were randomly divided into groups as testing, training and validation. The reason for this is to prevent the trained model from memorizing and overfitting. CNN was trained to have 200 epochs.

III. Tests and Analysis of the Proposed Method

Various tests have been carried out to measure the performance of the method that locates the manipulated region in the proposed image forgery. The proposed method has been implemented using the Python language in the Google Colab environment. Various additional applications have been made to evaluate the performance of the proposed method. First, the proposed method was trained and tests were carried out with images that were not used in the training selected from the data set. Then, tests were performed with forgered images that were outside the data set. The DCT layer used in the proposed system was removed and the method was retrained using only the DWT layer. Then, the DWT layer was removed and the method was retrained using only the DCT layer. A total of 3 methods were tested and success rates were compared with each other. Evaluation metrics used in training are presented in the next subsection.

A. Evaluation Metrics

Describe Various evaluation metrics [14] – [15] were used to measure the performance of the proposed method, including accuracy, recall, precision, F-score, and intersection over union metrics (IoU).

Accuracy measures the overall accuracy of the model's predictions by calculating the ratio of correctly predicted manipulated regions to the total number of images and is defined by equation 1 [16]:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FN + FP} \times 100 \qquad (1)$$

Describe In equation 1, TP occurs when the forgery image detection system correctly identifies a forged image as being forged. In other words, the system correctly detects the presence of forgery in an image that is actually forged. TN occurs when the forgery image detection system correctly identifies an non forged image as being non forged. The system correctly classifies an image as original when it is indeed original. FP occurs when the forgery image detection system incorrectly identifies an original image as being forged. The system mistakenly flags an original image as a forgery, even though it is original. FN occurs when the forgery image detection system incorrectly identifies a forged image as being original. The system fails to detect the presence of forgery in an image that is actually forged.

Recall is a performance metric that measures the system's ability to correctly identify all instances of forgery in a given dataset. The recall is calculated by dividing the number of true positives (TP) by the sum of true positives and false negatives (FN) and is calculated by equation 2 [16]:

$$\text{Recall} = \frac{TP}{TP + FN} \times 100 \qquad (2)$$

Precision is a performance measure that measures the system's ability to accurately identify forgered images among images that it classifies as forgered and is calculated by equation 3 [16]:

$$\text{Precision} = \frac{TP}{TP + FP} \times 100 \qquad (3)$$

F-measure (also known as F1 score) is a combined performance metric that takes into account both precision and recall. It provides a single value that represents the overall effectiveness of the system in detecting forged images. The F-measure is calculated using the harmonic mean of precision and recall and is calculated by equation 4 [16]:

$$\text{F} - \text{measure} = 2 \times \frac{Recall \times Precision}{Recall + Precision} \qquad (4)$$

Intersection over Union (IoU), also known as the Jaccard index, is a performance metric commonly used in forgery image detection systems to evaluate the accuracy of object detection or image segmentation tasks. It measures the overlap between the predicted bounding box or region and the ground truth bounding box or region. To calculate the IoU, the intersection area between the predicted region and the ground truth region is divided by the junction area of both regions and calculated by equation 5 [16]:

$$\text{IOU} = \frac{Intersection\ Area}{Union\ Area} \qquad (5)$$

B. Experimental Studies and Tests

This section describes the experiments and tests performed on the method that detects the manipulated region in the proposed forgered images. One of the tests aims to find out whether the manipulated location is detected when the trained network gives an image that is in the dataset but not used in the training of the model Fig. 7.

Fig. 7 The forgered image obtained by the image splicing method.

A forgered image was created by cutting and adding an eagle from an original image to an original image with a desert scene. It was observed that the proposed method was able to detect the location of the manipulated eagle when the forgered image was given to the trained model Fig. 8.

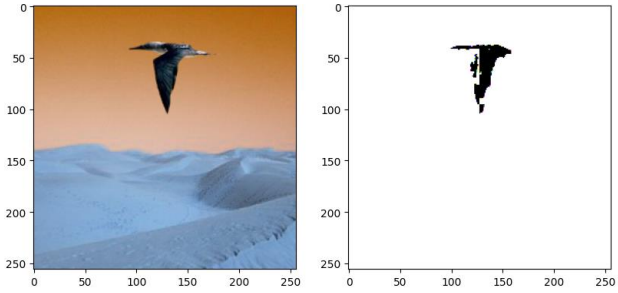

Fig. 8 Detecting the mask of the region of the manipulated eagle image.

Another test was carried out to detect the manipulated part of a forgered image that did not exist in the data set. A forgered image was produced by adding a round black object to an original nature-

themed image (shown in Fig.9) with an image processing software.

Fig. 9 Forgered image produced by adding a round black object to an original nature-themed image.

By giving the forgered image as input to the trained model, the manipulated region was successfully detected (Fig. 10).

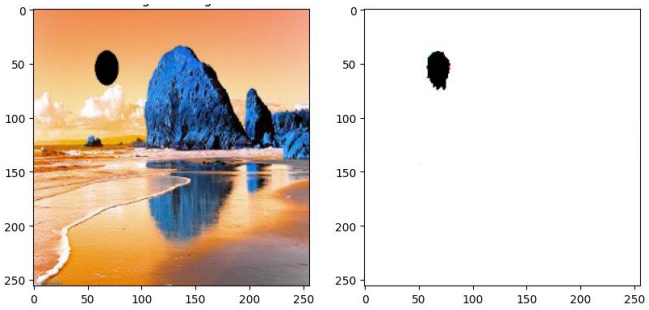

Fig. 10 Detection result the manipulated region in the forgered image obtained by adding a black object.

In order to determine the success of the proposed method, tests were carried out on two additional methods. In the first method, the DWT layer in the proposed system was removed and the manipulated region was determined using only the DCT layer. In the other method, the DCT layer in the proposed system was removed and the manipulated region was determined using only the DWT layer. While testing the methods, Casia V.2.0 dataset and Splited Copy dataset were used. Tests were carried out by selecting 1000 random images from each data set.

Table 1. Splited Copy Dataset Performance Metrics

Data Set	**Splited Copy**		
Performance Metric	**CNN+DWT+DCT (Proposed Method)**	**CNN+DCT**	**CNN+DWT**
Accuracy	94,37	89,25	91,48
Precision	96,43	88,42	93,24
Recall	94,67	89,17	92,29
F-Measure	97,08	92,48	93,59
IOU Score	94,63	94,15	95,64

Splited Copy is the data set used in the training of CNN. However, images not used in CNN training were used in the tests. The results of the performance metric are given in Table 1 by performing tests for three different methods on the data set.

Table 2. Casia V.2.0 Dataset Performance Metrics

Data Set	**Casia V.2.0**		
Performance Metric	**CNN+DWT+DCT (Proposed Method)**	**CNN+DCT**	**CNN+DWT**
Accuracy	94,29	88,34	92,17
Precision	95,08	89,62	92,88
Recall	95,86	90,44	89,78
F-Measure	96,77	91,68	92,28
IOU Score	95,38	95,43	95,14

In order to compare the success of the proposed systems, three different methods were also tested with an external dataset called Casia v.2.0, which was not used in CNN training. The performance metric results obtained are shown in Table 2.

IV. CONCLUSION

Today, with the development of image processing software, image forgery has increased. Image spliced and copy-paste techniques are the most used methods in image forgery. In this study, a method for the localization of manipulated regions of forgered images is proposed. In the proposed method, deep learning techniques were used together with DWT and DCT methods. The aim of the proposed method is to combine the advantages of traditional methods such as DWT and DCT with frequency analysis, durability against certain transformations and size reduction advantages of

CNNs that can learn by detecting inconsistencies in lighting texture or pixel level changes. Thus, the locations of the manipulated regions are determined more precisely. Various tests have been carried out to evaluate the performance of the proposed algorithm. In the proposed method, using the DWT layer and the DCT layer separately, a total of three different methods were tested on Splited Copy and Casia v.2.0 datasets. The proposed method showed the best performance in terms of accuracy, precision and F-Score. The method using CNN with DWT outperformed the method using CNN with DCT in terms of accuracy, precision and recall metric.

Future work under consideration includes methods for detecting image forgery, such as identifying the source camera and detecting the forgery technique. In addition, more accurate manipulated region estimation can be made by using Generative Adversarial Networks (GAN) together with CNN structures.

REFERENCES

[1]. Birajdar, G.K. and V.H.J.D.i. Mankar, Digital image forgery detection using passive techniques: A survey. 2013. 10(3): p. 226-245.

[2]. Parveen, A., Z.H. Khan, and S.N.J.I.J.o.C.S. Ahmad, Block-based copy–move image forgery detection using DCT. 2019. 2: p. 89-99.

[3]. Jaiprakash, S.P., et al., Low dimensional DCT and DWT feature-based model for detection of image splicing and copy-move forgery. 2020. 79: p. 29977-30005.

[4]. Dua, S., J. Singh, and H.J.P.C.S. Parthasarathy, Image forgery detection based on statistical features of block DCT coefficients. 2020. 171: p. 369-378.

[5]. Hayat, K., T.J.C. Qazi, and E. Engineering, Forgery detection in digital images via discrete wavelet and discrete cosine transforms. 2017. 62: p. 448-458.

[6]. Rao, Y. and J. Ni. A deep learning approach to detection of splicing and copy-move forgeries in images. in 2016 IEEE international workshop on information forensics and security (WIFS). 2016. IEEE.

[7]. Ouyang, J., Y. Liu, and M. Liao. Copy-move forgery detection based on deep learning. in 2017 10th international congress on image and signal processing, biomedical engineering and informatics (CISP-BMEI). 2017. IEEE.

[8]. Abd El-Latif, E.I., et al., A passive approach for detecting image splicing based on deep learning and wavelet transform. 2020. 45: p. 3379-3386.

[9]. Zhang, Z., et al. Boundary-based image forgery detection by fast shallow cnn. in 2018 24th International Conference on Pattern Recognition (ICPR). 2018. IEEE.

[10].Hebbar, N.K. and A.S. Kunte. Image Forgery Localization Using U-Net based Architecture and Error Level Analysis. in 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N). 2021. IEEE.

[11].Makhoul, J.J.I.T.o.A., Speech, and S. Processing, A fast cosine transform in one and two dimensions. 1980. 28(1): p. 27-34.

[12].Merry, R. and M.J.l.s. Steinbuch, Eindhoven university of technology, Department of mechanical engineering, Control systems technology group, Wavelet theory and applications. 2005.

[13].Aloysius, N. and M. Geetha. A review on deep convolutional neural networks. in 2017 international conference on communication and signal processing (ICCSP). 2017. IEEE.

[14].Zhang, Y., et al., Image region forgery detection: A deep learning approach. 2016. 2016: p. 1-11.

[15].Han, J.G., et al., Efficient Markov feature extraction method for image splicing detection using maximization and threshold expansion. 2016. 25(2): p. 023031-023031.

[16].Hussain, M., et al., Evaluation of image forgery detection using multi-scale weber local descriptors. 2015. 24(4): p. 1540016.