International Journal of Scientific Research & Engineering Trends
Volume 6, Issue 4, July-Aug-2020, ISSN (Online): 2395-566X

Image Forgery Detection Using Convolutional Neural Network

Alvina Aslam, Ankita Saxena, Sonali Saxena, Vaishnavi Raman Dwivedi
Assistant Professor Manish Gupta, Assistant Professor. Priyanka Goel
Dept. of Computer Science and Engineering,
Moradabad Institute of Technology Moradabad,UP

***Abstract*-with the advancement of high-resolution digital cameras and photo editing software featuring new and advanced features, the chances of image forgery have increased. the images can now be altered and manipulated easily. image trustworthiness is now more in demand. images in courtrooms for evidence, images in newspapers and magazines, and digital images used by doctors are few cases that demands for images with no manipulation. in this paper we discuss some of the types of image forgery and techniques to fight against these forgeries. he revolution in the digital world is changing the way in which we share and manipulate data, but this revolution has introduced many critical security issues that hampers the integrity of digital media. many sophisticated digital technologies and photo-editing software like adobe photoshop have made the manipulation of images a fair practice. as a result, digital images are becoming prone to forgeries and hence trust in digital images has been eroded. digital forgery is now a nightmare to individuals (e.g. fake images of celebrities and public figures), societies (fake images targeting religion or race), journalism, scientific publication etc. the existing image forgery detection techniques are widely divided into two categories-active approach and passive (blind) approach. active approaches rely on pre-registration or pre-embedded information. a shield is produced for the images to protect them from being manipulated. it is mainly based on digital watermarking. the main disadvantage of this approach is that the protection against manipulation must precede any attempt of forgery which means the pre-existing digital images and data cannot gain any profit using the approach. passive approach overcomes this disadvantage; the pre-existing images can also be catered using this approach. in this project, we propose the use of discrete cosine transform and a deep learning approach to learn features in order to detect tampered images. among different types of image forgery, copy-move forgery is the most popular to forge the digital imaged where a part of the original image is copied and pasted at another position in the same image. different methods have been developed to detect the image forgery in digital images. to address this issue, we present a pixel-based copy-move forgery detection method to check the genuineness of digital images.in this project we detect region duplication forgery by applying discrete cosine transform. we divide the image into overlapping blocks and then search for the duplicated blocks in the image. proposed method includes the following steps: (1) convert the color image into gray-scale image, (2) divide the gray-scale image into overlapping blocks of size 8×8, (3) feature extraction using dct on the basis of different feature sets,(4) implementation of convolutional neural network.**

***Key words*-Image forgery, Convolutional Neural Network, Support Vector Machine, Discrete Cosine Transform, zigzag scan, lexographic scanning.**

I.INTRODUCTION

In today's world it is easy to manipulate the image by adding or removing some elements from the image which result in a high number of image forgeries. Using the manipulation tools that are available on internet it is easy to tamper the digital images without any trace. Therefore, verification of originality of images has become a challenging task. An image can be manipulated with a wide variety of manipulation techniques such as scaling, rotation, blurring, resampling, filtering, cropping, etc. We need image forgery detection technique in many fields for protecting copyright and preventing forgery [2]. The verification of originality of images is required in variety of applications such as military, forensic, media, scientific, glamour, etc. Image tampering is a digital art which needs understanding of image properties and good visual creativity. Detection of image tampering deals with investigation on tampered images for possible correlations embedded due to tampering operations. Detecting forgery

in digital images is a rising research field with important implications for ensuring the credibility of digital images. Three of the most common types of image forgeries are:

1. Copy-move- a specific region from the image is copy pasted within the same image.

2. Splicing- a region from an authentic image is copied into a different image.

3. Image Resampling - To make an astounding forged image, some selected regions must undergo geometric transformations like rotation, scaling, stretching, skewing, flipping and so forth.

The basic aim of the project is to develop an efficient system which can detect copy-move image forgery. There are various ways to detect image forgery. They are classified into active and passive approaches. In the active approach, the digital image requires pre-processing of image such as watermark embedding or signature generation, which limits their application in practice. Unlike the watermark and signature-based methods, the passive techniques do not need any digital signature to be generated or to embed any watermark.

Passive image forgery detection techniques roughly can be divided into five categories. Pixel-based techniques detect statistical anomalies introduced at the pixel level; format-based techniques leverage the statistical correlations introduced by a specific lossy compression scheme; camera-based techniques exploit artifacts introduced by the camera lens, sensor, or on-chip post-processing; physical environment-based techniques explicitly model and detect anomalies in the three-dimensional interaction between physical objects, light, and the camera; and geometry-based techniques make measurements of objects in the world and their positions relative to the camera. However, we have developed a copy- move image forgery detection algorithm that includes use of DCT followed by implementation of Convolutional Neural Network. accuracy. The proposed model consists of following steps such as: taking multiple authentic and tampered images from MICC-F220 Dataset, pre-processing, edge detection and morphological processing, followed by training and testing of neural network.

1. Image forgery detection

Image Forgery is not new. History has recorded that it happens as early as the 1840s. Hippolyta Bayard, the first person to create a fake image as recorded by history, is famous for a picture of him committing suicide (see picture on the left). It all started as an act of frustration because he had lost the chance of becoming 'the inventor of photography' to Louis Daguerre[3]. Daguerre patented a photography process earlier than him and owns all the glory. whether the picture is unique or manipulated. There is fast increment in digitally controlled falsifications in standard media and on the Internet. This pattern shows genuine vulnerabilities and abatements the credibility of digital images.

2. Methods of image forgery detection

- Pixel Based Forgery Detection
- Format-based image forgery detection
- Camera based image forgery detection
- Physical environment-based image forgery detection
- Physical environment-based image forgery detection

II. PIXEL BASED FORGERY DETECTION

The pixel-based image forgery detection aims to verify the authenticity of digital images without any prior knowledge of the original image. There are many ways for tampering an image such as splicing or copy-move, resampling an image (resize, rotate, stretch), addition and removal of any object from the image.
Technique- There are many approaches that have been proposed by various authors for detecting pixel-based image forgery. Figure below shows the general process of detecting copy-move image forgery

III. METHODOLOGY USED TO DETECT IMAGE FORGERY

- Module1: Taking image as input and converting into grayscale.
- Module 2: Implementation of DCT.
- Module 3: Implementation of Zigzag Scan.
- Module 4: Implementation of Quantization.
- Module 5: Implementation of lexicographic ordering.
- Module 6: Euclidean is applied to determine the similarity of vectors with the neighboring vectors.
- Module 7: Calculation of direction of vectors passing the distance threshold.
- Module 8: Determination of accuracy.
- Module 9(final Module): Implementation of CNN.

IV. DETECTING COPY MOVE FORGERY USING DCT

One approach to detect copy move forgery detection, proposed by Fridrich et al., basically performs a rigorous search by comparing the image to every cyclic-shifted version of it. But the complexity of this approach is very high, it requires (mn)2 steps to execute for an image of size M×N so it is difficult to implement it practically.

One of the distinguish property of copy move forgery detection is the feature extraction process. Some methods are based on dimensionality reduction, moments , colour properties, frequency domain transform . Other techniqueto detect copy move forgery is by using

Discrete Cosine Transform (DCT). Junfeng He et.al. proposed the method that can detect forged jpeg image and locate the doctored part by applying the DCT transform on images. This method has many other advantages like fast speed etc. There is an approach that can detect doctored JPEG images, by examining the double quantization effect hidden among the DCT coefficients. Our method detects forgery by dividing the image into overlapping blocks and then we search for the matching region in the image. We show the effectiveness of this technique on credible forgeries and compute its robustness also.

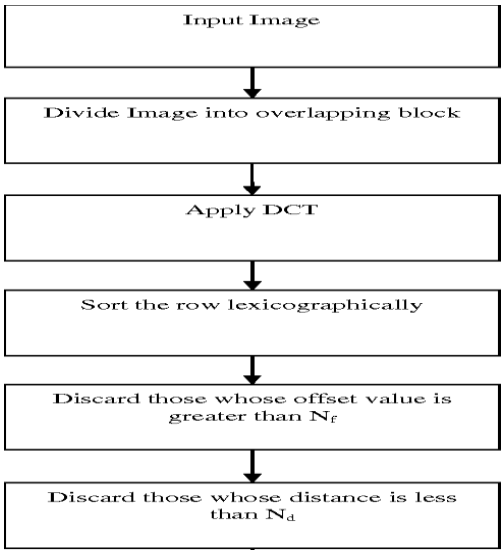

Fig 1 Duplication Detection Algorithm.

1. Experiment Result:

Fig 2 Input Image.

Fig 3 Tempered Image.

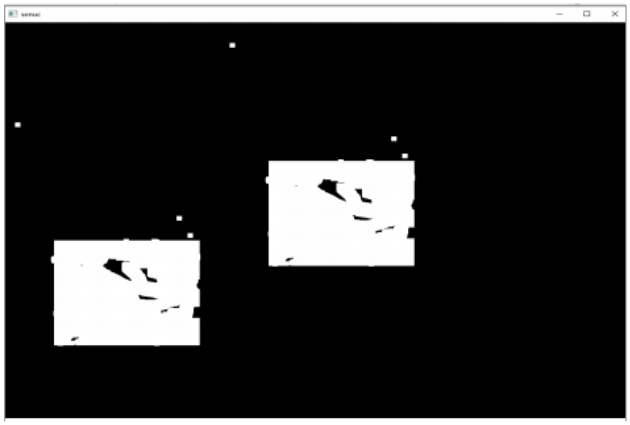

Fig 4 Output Image.

2. Result analysis

Comparison of execution time when block size = 8×8 (Using DCT)

Size of Image	Execution time
1024 X 768	240 seconds (approx.)
174 X 132	201 seconds(approx.)
256 X 256	7 seconds(approx.)
128 X 128	27 seconds(approx..)

3. Efficiency of Algorithm

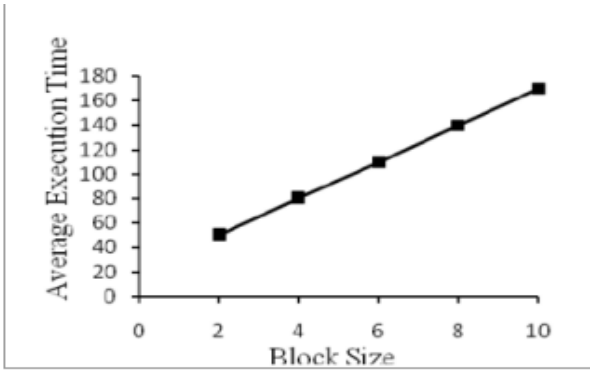

Fig 5 Graph between Avg. Execution Time and Block size.

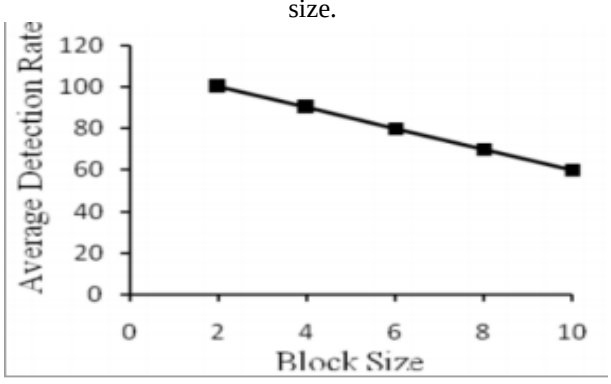

Fig 6 Graph between Avg. Detection Rate and Block size.

V. ZIGZAG SCAN

Zigzag scanning is a transform-based coding .it is employed for non-uniform quantization of N x N DCT coefficients. Lower coefficients have most of the energy and it is distributed circularly symmetric about the origin. The net result is that it results in 1D sequence, after certain number of non-zero coefficients most of the remaining become 0.

VI. QUANTIZATION

Quantization, involved in image processing, is a lossy compression technique achieved by compressing a range of values to a single quantum value. When the number of discrete symbols in a given stream is reduced, the stream becomes more compressible. For example, reducing the number of colors required to represent a digital image makes it possible to reduce its file size. Quantization makes the range of a signal discrete, so that the quantized signal takes on only a discrete, usually finite, set of values. Relation of Quantization with gray level resolution: Now 256, or 5 or whatever level you choose is called gray level. Remember the formula that we discussed in the previous tutorial of gray level resolution which is,

$$L=2k$$

The quantization is operated using the 8 × 8 quantization matrix. Each DCT coefficient is quantized, that is, divided by the corresponding value given in the quantization matrix. In this way, a smaller number of bits can be used for encoding the DCT coefficients[4].

This is an example of DCT coefficient matrix:

$$\begin{bmatrix} -415 & -33 & -58 & 35 & 58 & -51 & -15 & -12 \\ 5 & -34 & 49 & 18 & 27 & 1 & -5 & 3 \\ -46 & 14 & 80 & -35 & -50 & 19 & 7 & -18 \\ -53 & 21 & 34 & -20 & 2 & 34 & 36 & 12 \\ 9 & -2 & 9 & -5 & -32 & -15 & 45 & 37 \\ -8 & 15 & -16 & 7 & -8 & 11 & 4 & 7 \\ 19 & -28 & -2 & -26 & -2 & 7 & -44 & -21 \\ 18 & 25 & -12 & -44 & 35 & 48 & -37 & -3 \end{bmatrix}$$

A common quantization matrix is:

$$\begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

Dividing the DCT coefficient matrix elementwise with this quantization matrix, and rounding to integers results in:

$$\begin{bmatrix} -26 & -3 & -6 & 2 & 2 & -1 & 0 & 0 \\ 0 & -3 & 4 & 1 & 1 & 0 & 0 & 0 \\ -3 & 1 & 5 & -1 & -1 & 0 & 0 & 0 \\ -4 & 1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

For example, using −415 (the DC coefficient) and rounding to the nearest integer

$$\text{round}\left(\frac{-415}{16}\right) = \text{round}\left(-25.9375\right) = -26$$

VII. LEXICOGRAPHICAL ORDERING

Lexicographical ordering approach correlates the pixel values and reconstructs the correlation to a new color space that is essential for human perception. A vector represents a pixel in color space and vectoral ordering is required for comparison between the vectors.

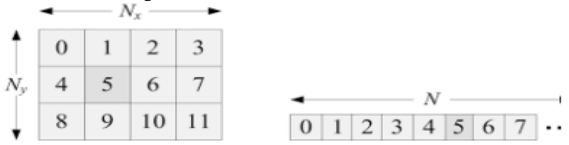

Fig 7 Lexicographical Ordering.

VIII. MEASURING EUCLIDEAN DISTANCE IN IMAGES

The Euclidean distance is the distance between two points in Euclidean space. The two points P and Q in two dimensional Euclidean spaces and P with the coordinates (p1, p2), Q with the coordinates (q1, q2).

$$d(\mathbf{p},\mathbf{q}) = d(\mathbf{q},\mathbf{p}) = \sqrt{(q_1-p_1)^2+(q_2-p_2)^2+\cdots+(q_n-p_n)^2}$$
$$= \sqrt{\sum_{i=1}^{n}(q_i-p_i)^2}.$$

Step1: load the column vector x;

Step2: load the code book;

Step3: minimum distance is initially set to the first element of cb.

Step4: i.e. set idx=1;

Step5: compute distance by normalized values of (x-cb) for all cb;

Step6: if d is less than distance set distance is equal to d;

Step7: set idx=index;

Step8: end

Figure 8 Euclidean Distance Algorithm.

IX. CALCULATION OF DIRECTION OF VECTOR PASSING THE DISTANCE THRESHOLD

Change-Vector Analysis (CVA) is a bi-temporal method that was originally designed for only two spectral dimensions (2-D CVA): Brightness (an indicator of overall reflectance) and Greenness (an indicator of vegetation), both from the Tasseled Cap transform. The CVA produced as output two change components: magnitude and direction[5].
The mathematical framework of the n-D CVA calculation is presented below.

$$X^{t_1} = \{x_1^{t_t}, x_2^{t_1}, \ldots, x_N^{t_1}\} \text{ and } X^{t_2} = \{x_1^{t_2}, x_2^{t_2}, \ldots, x_N^{t_2}\}$$

x1,x2……..xn are the sets of N spectral bands of images acquired at time t1 and t2, respectively, the angle between the vector and each axis (θ1, θ2, θ3,…θN) can be expressed by the arccosine function:

$$\theta_i = cos^{-1}\left(\frac{X_i^{t_1} - X_i^{t_2}}{ED}\right) \; i = 1 \ldots N$$

where ED is the Euclidian distance, expressed as the square root of the band-wise sum of the squares of the differences:

$$ED = \sqrt{\Sigma_{i=1}^{N}(X_i^{t_1} - X_i^{t_2})^2}$$

X. DETERMINE THE FORGED REGION BY MASK

The copy-move forgery detection baseline was first developed by making feature maps from input image extracts, followed by the construction of relevant feature statistics based on percentage pooling process from up-sampled feature maps.

XI. CONVOLUTIONAL NEURAL NETWORK

A convolutional neural network (CNN, or ConvNet) is a class of deep neural networks, most commonly applied to analyzing visual imagery.

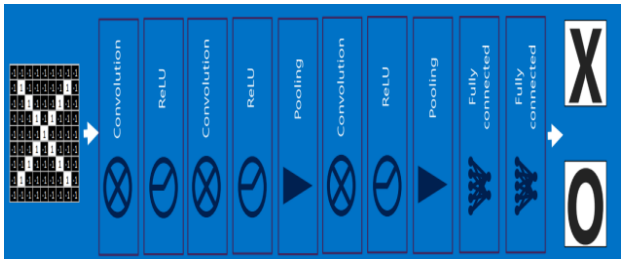

Fig 9 Fully connected Layers.

XII. FINAL RESULT

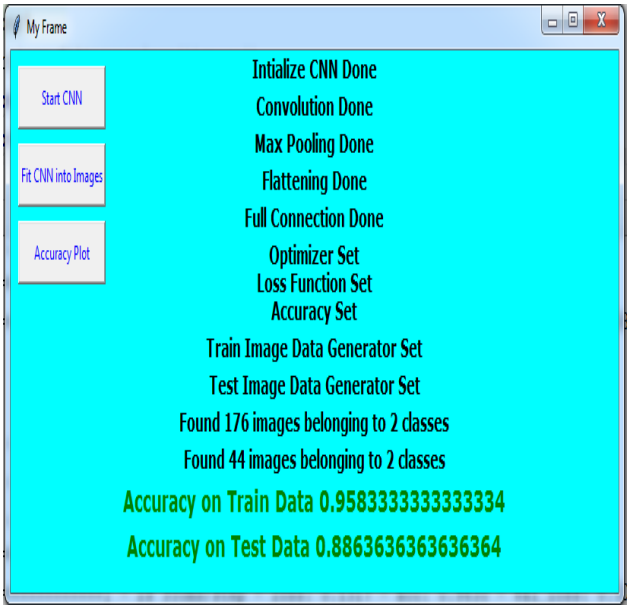

Fig 10 CNN Training and Testing Accuracy.

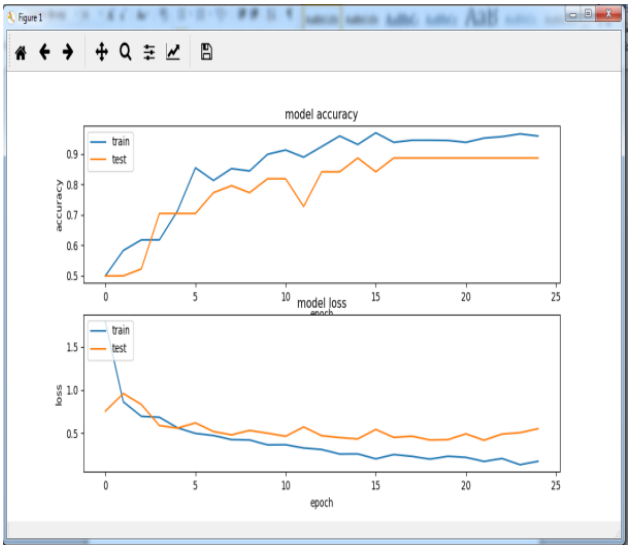

Fig 11 Accuracy and Loss Plots.

XIII. CONCLUSION

Copy-move forgery is one of the most frequently applied forgery techniques. In this we use a robust method to detect the duplicated region in the digital image. We have conducted some test on the algorithm against sample images from the internet. The result of the test is very encouraging since we got improvements in the detection rate and the detection time of the copy-move attack detection algorithm that we used. We are happy that the project is able to meet the outlined objectives proved that the use of DCT is better than using PCA for detecting copy-move attacks in highly textured images. We can improve the efficiency of forgery detection by applying wavelet transform. For future work, we plan

to further optimize the data structures to gain additional query performance and further improve accuracy. The process can be further extended to different formats and works for binary scale, gray scale and color images also[6].

REFERENCES

[1] R. B. W. a. E. J. Delp, "A Watermark for Digital Images," in International Conference on Image Processing, vol.3, 1996.
[2] C. S.-F. Ng T-T, "A data set of authentic and spliced image block," [Online]. Available: http://www.ee.columbia.edu/trustfoto .
[3] P.Pradyumna Deshpande, "Pixel Based Digital Image Forgery Detection Techniques," International Journal of Engineering Research and Applications (IJERA), pp. 539-543 , 3, May-Jun 2012.
[4] N. R. W. P. G. Gomase, "Advanced Digital Forgery Detection: A Review," IOSR Journal of Computer Science (IOSR-JCE) e-ISSN: 2278-0661, pp. 80-83, 2010.
[5] N. S. V. Ashima Gupta, "Detecting copy move forgery using DCT," International Journal of Scientific and Research Publications, Volume 3 ISSN 2250-3153 , 5 May 2013.
[6] G. B. C. P. H. S. a. B. S. A. S.Murali, "Comparison and Analysis Of Photo Image Forgery Detection Techniques," International Journal on Computational Sciences & Applications (IJCSA), pp. Vo2, No.6, 6, December 2012 .