Pixel Based Image Forgery Detection Using Deep Learning

Prof. Ashwini Gavali**, Mr. Prajwal P Gowdra*, Mr. Rohan Kodag*,

Mr. Shubham Kadam*, Mr. Sidramaraddy*.

** Asst. Prof. CSE Department
* Student
KLECET, Chikodi, Tq: Chikodi, Dist: Belagavi, Karnataka, India.
2021-22

Abstract— **Society is becoming increasingly dependent on the internet and so does it become more and more vulnerable to harmful threats. These threats are becoming vigorous and continuously evolving. These threats distort the authenticity of data transmitted through the internet. As we all completely or partially rely upon this transmitted data, hence, its authenticity needs to be preserved. Images have the potential of conveying much more information as compared to the textual content. We pretty much believe everything that we see. In order to preserve/check the authenticity of images, image forgery detection techniques are expanding its domain. Detection of forgeries in digital images is in great need in order to recover the peoples trust in visual media. This proposed system is going to discuss different types of image forgery and blind methods for image forgery detection. It provides the comparative tables of various types of techniques to detect image forgery. It also gives an overview of different datasets used in various approaches of forgery detection. Therefore, it is crucial to detect such image forgery operations in the images. The image forgery detection can be done based on object removal, object addition, unusual size modifications in the image. Images are one of the powerful media for communication. In this proposed system, a survey of different types of forgery and digital image forgery detection has been focused.**

Keywords—Image Forgery, Authentication, Image Tampering.

Introduction

Image Forgery is not a modern concept as it comes along with the invention of photography. But it comes in the limelight nowadays, with the invent of easily accessible digital cameras supported with image editing software tools. Image Forgery begins with the first known fake image that was of Hippolyta Bayard, who released a fake picture of him committing suicide as an act of annoyance for the sake of losing the tag of inventor of photography to Louis Daguerre in 1840.

Digital visual media, nowadays, represent one of the prominent techniques of exchanging information, because of increase in easy to use and inexpensive devices. Moreover, visual media has greater expressive potential than any of the existing media. It describes convoluted scenes in an uncomplicated manner, whichever in a different way can be quite tough to transcribe. Malicious modification of digital images with intent to deceive for the sake of altering the public perception is termed as Digital Image Forgery. The modification is done in such a way that it hardly leaves any visually detectable traces. Manipulation of Digital images isn't any longer defined to experts with all the arrival and dispersal of handy image editing tools and software's. Some of the well-known images editing tools available online are Sum paint, Paint shop Pro, Photoshop CC, HitFilm Express.

Manipulation of visual media with such easily available tools is no longer a herculean task. It is not concerned whether an image is fake or not, until or unless it causes some harm. These images are accepted as certification of truthfulness almost by everyone and everywhere. So, confirmation of an Image authenticity is needed. Such confirmation is done with the help of image forgery detection techniques. These methods aim at validating the authenticity of images. There are several types of image forgery exposed to date and correspondingly the forgery detection techniques. This system aims to review the existing types of forgeries and their detection techniques.

I. Literature Survey

Before starting with this proposed system, it is necessary to carry out deep study of currently available systems in the industry and work carried out by different research scholars to understand the advantages and limitations of their research work. Below is the literature review of the work carried out by few research scholars which we have documented by referring to various research papers and international journals.

1) Review on Copy-Move Image Forgery Detection Techniques by Zaid Nidhal Khudhair (2021).

The system by Zaid made certain explanation to image forgery where the main focus was on the copy-move image forgery detection.

He compared all the recent algorithms (published in 2017-2020).

2) Detection of Image Forgery by Shubham Sharma, Sudeeksha Verma, Swapnil Srivastava (2020)

This system for detection of image forgery uses feature point extraction and morphological operation

It can divide the forged region by indicating the affected pixel

3) **A Robust Forgery Detection Method for Copy–Move and Splicing Attacks in Images by Manzur Murshed (2020)**

This system introduced an efficient and robust model for detecting splicing and copy–move attacks in both grayscale and colour images using traditional machine learning technique (SVM) and hand-crafted features.

SVM with RBF kernel was used for classification and the kernel parameters were tuned using Bayesian optimization.

4) **Digital Image Forgery Detection by S.Prayla Shyry, Saranya Meka, Mahitha Moganti (2019)**

This system reviewed various traditional techniques which are been used and also the

study of different image forgery techniques has been done elaborately with pros and cons.

Various types of passive approaches were being used like Image splicing, Image Retouching and Copy-Move attack.

II. PROPOSED SYSTEM

In this proposed system we deal with the concept of Pixel based Image forgery detection. The objective of the project is to develop an algorithm to determine image tampering and show the corresponding results to the user. Few of the objectives are:

1. To provide a GUI (Graphical User Interface) for the user to interact with the application in order to check the authenticity of the images.
2. To provide a percentage of authenticity as output to the user which in turn will help in determining whether the image is forged or not in the fake image detection application.
3. To provide various functionalities to further determine the authenticity of the image.

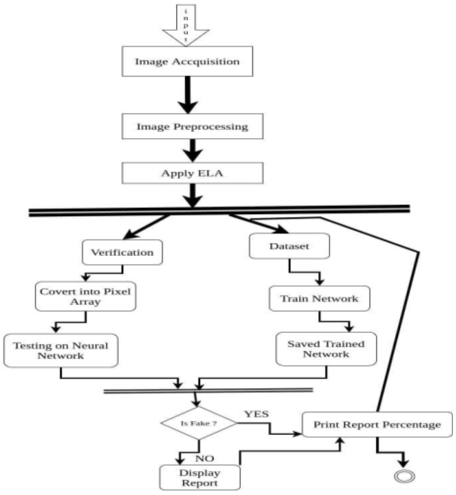

Fig 1: Activity diagram of Fake Image detection using CNN.

The above (Fig 1) process works as follows:

Initially the user inputs an image to the system. The same image passes through Image acquisition module where it will retrieve an image from the source. After the acquisition the pre-processing module comes into action, where the steps taken to format an image before they are used by model training and inference like resizing, orienting and colour corrections. After this module the actual technique comes into action where the technique which has major impact starts to process the image The technique is called ELA(Error Level Analysis).

ELA works by intentionally resaving the image at a known error rate, such as 95%, and then computing the difference between the images. If there is virtually no change, then the cell has reached its local minima for error at that quality level

However, if there is a large amount of change, then the pixels are not at their local minima and are effectively original. The ELA then splits the section into two categories i.e., Verification and Dataset. Dataset corresponds to all the images that are required for carrying out the verification operation. In the Train network section, the process of identifying each image is carried out and further the same trained data is stored in source which is the next section i.e., Saved Trained Network.

The trained network data is then further used in the verification section. In the next section the image is converted into a pixel array based on the height, width, channel format. All the converted images are tested on the neural network, here the fake area of the image will be detected using the compression level of the image. If the image has a higher compression score than the other areas then the image is tampered. Lastly, we check the corresponding output by a condition that if it is fake then it will print report percentage after which user can either exit or redo the step again. If it not fake then it will display report and print report percentage. Based on output percentage, we can determine whether the image is fake or not.

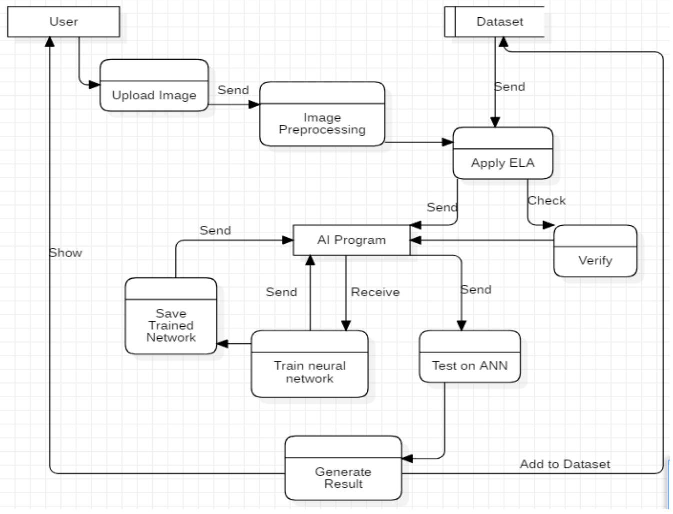

Fig 2: Data-flow diagram of the proposed system.

In the above (Fig2) data-flow diagram we have complete set of different modules for the processing of image. Initially, the user will upload a jpg image and forward it to the Image Pre-processing module where here there are steps taken to format images before they are used for model training and inference.

Then we proceed to apply ELA where it has two choices one is to send and other is to check for verification.

If it has been sent then the AI Program will further forward it to test it on ANN (Artificial Neural Network). It will also perform training on neural network by sending and receiving through the AI program. Further down the line after the neural network has been trained it will then forward it to save trained network module where the test results will be saved in a module and then reverted back to the AI Program After a complete test on ANN it will move on to generate result module where the results will be displayed.

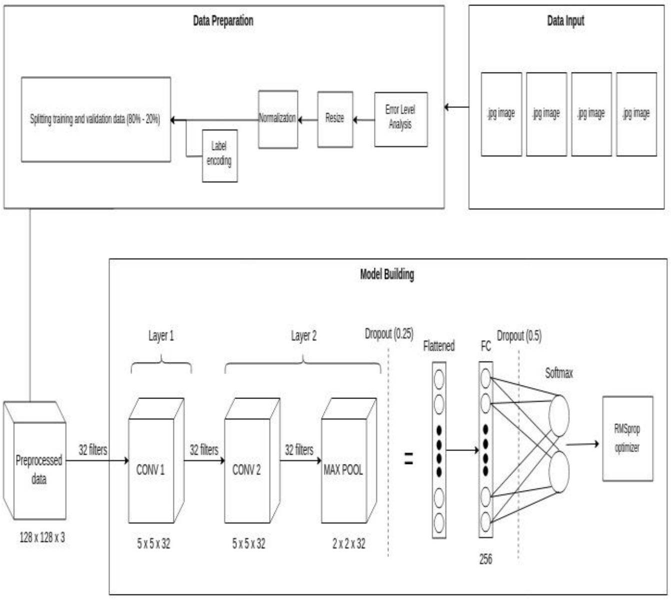

Fig 3: Model Architecture.

The convolutional neural network was originally proposed by LeCun et al. for handwritten recognition has been successful in image identification, detection, and segmentation of the image [13]. CNN has a high ability in large-scale image classification. CNN consists of three layers: convolutional layer, pooling layers, and fully connection layers. A Convolutional layer and pooling layer is the most important layer on CNN. Convolutional layer is used for extract feature by combining the image area with many filters. Pooling layer reduces the size of the output map of the convolution layer and prevents overfitting. Through these two layers the number of neurons, parameters, and connections is much less than there is a CNN model. This makes CNN more efficient compared with BP networks with similar layers. The final formula of the single output image channel of the convolution layer as:

$conv(I,K)xy=\sigma(b+\Sigma\Sigma\Sigma Kijk\ \cdot Ix+i-1,y+j-1,kdk=1wj=1hi=1)$ (1)

The layers are determined by specific kernels, K along with the bias value on each kernel. It then operates by calculating the output image of the previous layer with each of the kernels. Convolution is a mathematical term that means applying a function to the output of another function repeatedly. The kernel moves from the top left corner to the bottom right. So, the result of the convolution of the image can be seen in the picture on the right. The goal of convolution in image data is to extract features from the input image. The convolution will produce a linear transformation of the input data according to the spatial information in the data. A very popular approach to down sampling is to use pooling layers. Pooling layer usually deciphers the image (like 2x2) in the aggregation into a single unit. The most popular scheme for aggregation is the incorporation of the maximum value (max-pooling).

Subsampling is the process of reducing the size of the image data. In image processing, subsampling also aims to increase the position invariance of features. In most CNN, the subsampling method used is max pooling. Max pooling divides the output from the convolution layer into several small grids and then takes the maximum value of each grid to construct a reduced image matrix. The red, green, yellow and blue grids are the grid to be selected for maximum value. So the results of the process can be seen on the grid set on the right. The process ensures that the features obtained will be the same even if the image object is translating (shifting). The formula of max-pooling is as follows:

$Yil+1,jl+1,d\ =\ max0\leq i<H,0\leq j<W\ \ \mathcal{X}il+1\times H+i,jl+1\times W+j,dl$ (2)

Layer is a layer that is usually used in the application of MLP and aims to transform the dimensions of data so that data can be classified in a linear. Each neuron in the convolution layer needs to be transformed into one dimensional data first before it can be inserted into a fully connected layer. Because it causes data to lose spatial information and not reversible, fully connected layer can only be implemented at the end of the network. Applying CNN for fake image classification and original image converted into error level form on image. We know through the previous literature that CNN can achieve competitive performance and even better than humans in some visual problems, and we wanted to test CNN's ability to classify forgery image and original images via Error Level Analysis.

III. RESULTS

In this section, we will describe the experimental results from the recognition of the original image and the fake image. We analyse the accuracy percentage of the drawing training. Percentage of training varies from 60-90%. This shows that the method we use is able to study the data despite the small amount of data. From the training we have done, then we get results based on Figure 4.

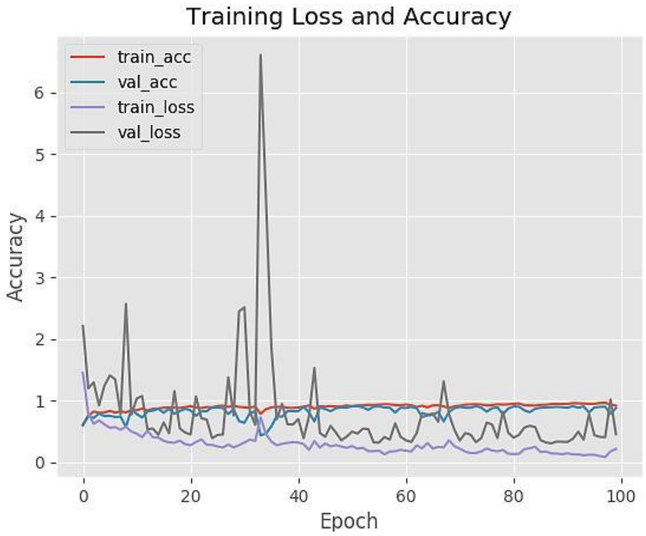

Fig 4: Model accuracy vs model loss image forgery detection

From the picture above (Fig4) that known with training accuracy of the model achieved up to 98.13% using 30 epoch. Thus, by using the deep learning architecture of proposed system in analysing error level image analysis for image forgery can be applied and get good results on recognition.

IV. Conclusion

Authenticity and integrity of the digital images are well thought-out to be important to overcome these issues because of the forging in fields such as forensic, medical imaging, e-commerce, industrial photography, etc. The authenticity verification check of the image is popularly used where the images are considered to supporting evidences, historical records, insurance claims, etc. Because of the drastic increase in the software availability for the advanced image manipulation and processing

In this paper, we have solved the problem of distinguishing real images and forgery images using deep learning. We propose a new system from combination Error Level Analysis and Convolutional Neural Network in machine learning and computer vision to solve the problems above. First, we divide the dataset into tampered images and original images, then we determine the architecture that will be used to train the recognition.

V. Future work

We can conduct a CNN architecture variant to get the best accuracy and do other approaches in processing image processing to recognize the original image and forgery image and conduct the test for more than 50 epoch.

References

1. Zaid Nidhal Khudhair, "A Review on Copy-Move Image Forgery Detection Techniques", et al 2021 J. Phys.: Conf. Ser.

2. Shubham Sharma, Sudeeksha Verma, Swapnil Srivastava, "Detection of Image Forgery", August 2020 Department of Computer Science and Engineering, PSIT College of Engineering Kanpur, India.

3. Hisyam Fahmi1, Wina Permana Sari "Effectiveness of Deep Learning Architecture for PixelBased Image Forgery Detection 2020", Faculty of Science and Technology, Universitas Islam Negeri Maulana Malik Ibrahim, Indonesia

4. Mohammad Manzurul Islam, Gour Karmakar, Joarder Kamruzzaman and Manzur Murshed, "A Robust Forgery Detection Method for Copy–Move and Splicing Attacks in Images", School of Engineering, IT and Physical Sciences, Federation University Australia.

5. S.Prayla Shyry, Saranya Meka, Mahitha Moganti, "Digital Image Forgery Detection", International Journal of Recent Technology and Engineering (IJRTE), July 2019.

6. Pradyumna Deshpande, Prashasti Kanikar, "Pixel Based Digital Image Forgery Detection Techniques" May-Jun 2019.

7. Bo Xu, Guangjie Liu, and Yuewei Dai , "Detecting Image Splicing Using Merged Features in Chroma Space‖, the Scientific WorldJournal",2015.

8. Arunvinodh C and M.F, Reshma P.D, "Image Forgery Detection Using Svm Classifier," IEEE Sponsored 2nd International Conference on Innovations in Information Embedded and Communication Systems", 2015.

9. Christian Riess and Tiago Jose de Carvalho, "Exposing Digital Image Forgeries by Illumination Color Classification", IEEE Transactions on Information Forensics And Security, vol. 8,2014.

10. B. L. Shiva Kumar and Lt. Dr.Santosh, "Detecting Copy Move Forgery in Digital Images: A Survey and analysis of current methods", GlobalJournal of Computer Science and Technology, vol.10, 2014.