IJNRD.ORG ISSN : 2456-4184

INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG

An International Open Access, Peer-reviewed, Refereed Journal

Analysis and Implementation of Fake Image Detection Using Machine Learning

Vaishali Gedam[1], Sahil Meshram[2], Shrinivas Chinchanikar[3], Ganesh Pandey[4], Yash Lad[5], Punam Bhandarkar [6]

[1]Associate Professor, Department of Computer Science and Engineering, NIT, Nagpur
[2]B.E Graduate (iv year), Department of Computer Science and Engineering, NIT, Nagpur
[3]B.E Graduate (iv year), Department of Computer Science and Engineering, NIT, Nagpur
[4]B.E Graduate (iv year), Department of Computer Science and Engineering, NIT, Nagpur
[5]B.E Graduate (iv year), Department of Computer Science and Engineering, NIT, Nagpur
[6]B.E Graduate (iv year), Department of Computer Science and Engineering, NIT, Nagpur

Abstract : These days, the availability of image processing software, such as Adobe Photoshop or GIMP have made image manipulation so common. Detecting such fake images is unavoidable for unveiling of the image-based cybercrimes. An image taken by digital camera or smartphone is usually saved in the JPEG format due to its popularity. JPEG algorithm works on image grids, compressed independently, with a size of 8x8 pixels. While unmodified images, have a similar error level. For resaving operation, each block should degrade at around same rate due to similar amount of errors across the whole image. The compression ratio of this fake image is different from that of the original image and is detected using Error Level Analysis. The objective of our paper is to develop a photo forensics algorithm which can detect any photo manipulation. The error level analysis was then enhanced using vertical and horizontal histograms of error level analysis image to pinpoint the location of modification. Results show that the proposed algorithm could identify the modified image while showing the exact location of modifications.

Index Terms **-** Biometry, Identity, Recognition, Detection, Fake face.

INTRODUCTION

The use of technology in today’s world has been increased massively, one of the most common sources of communication is using images in these days, images has become pretty common these days they are used in newspapers, magazines, websites and advertisements and provide several information. The trust in images is increasing day by day due to their increase in everyday usage. Tampering or manipulating an image by altering some information within it is known as image forgery and to check whether the image is real or not is termed as Image Forgery Detection. Enormous number of people have become victims of image forgery in our modern society. A lot of people use image manipulating software’s to manipulate images and use it as evidence to mislead the court or several other people on social media sites or applications. This is why every image that is shared on the social media should be evaluated and generalized as either real or fake. Social media is one of the best platforms to socialize, share and spread knowledge but if no precautions are taken, it can mislead people resulting to cause havoc due to unintentional false propaganda. While it takes some practice to photoshop images and can clearly be observed due to pixelization and shady jobs by novices but some of them when manipulated by a professional can indeed appear genuine. Especially in the political aspect’s images can be manipulated to make or break a politician’s credibility. Forensic techniques practiced these days to manipulate images require an expert to analyze the credibility of an image. This approach may be practical for a small number of images however it is not recommended to be used for evaluating a large number of images such as on a social media website. Therefore, we need to implement a system that can determine whether an image either real or fake with the help of current machine learning algorithms available to us and thereby make it available for use to the common public. This paper will further unfold into three proposed methodologies that can be followed respectively to evaluate the originality of an image whereby first we will focus on the metadata analysis, secondly we will focus on the Error Level Analysis of the images and in the last part we will focus on generating a machine learning algorithm to evaluate the image.

II. TYPES OF IMAGE TAMPERING TECHNIQUES

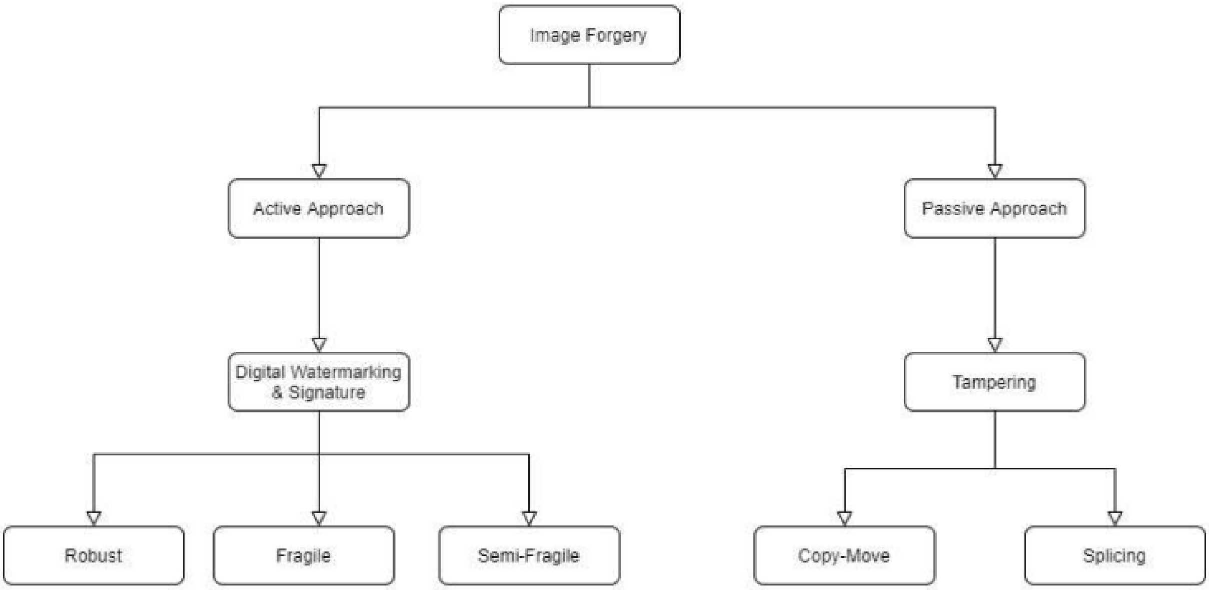
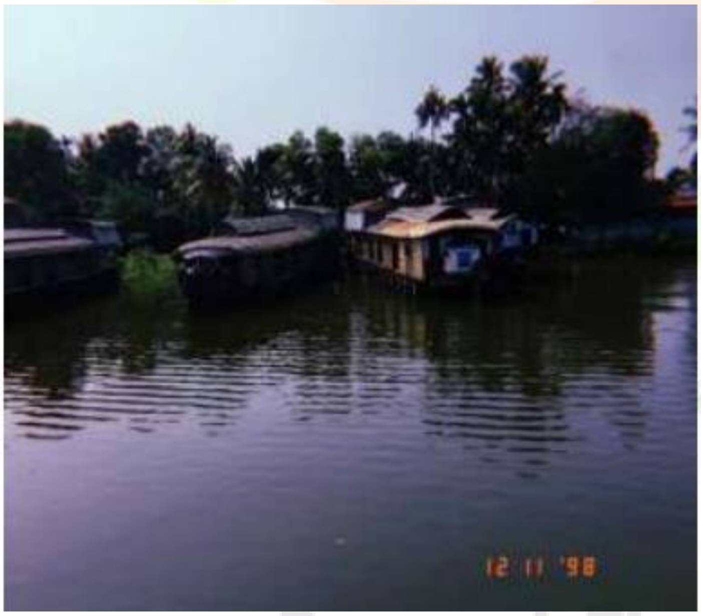
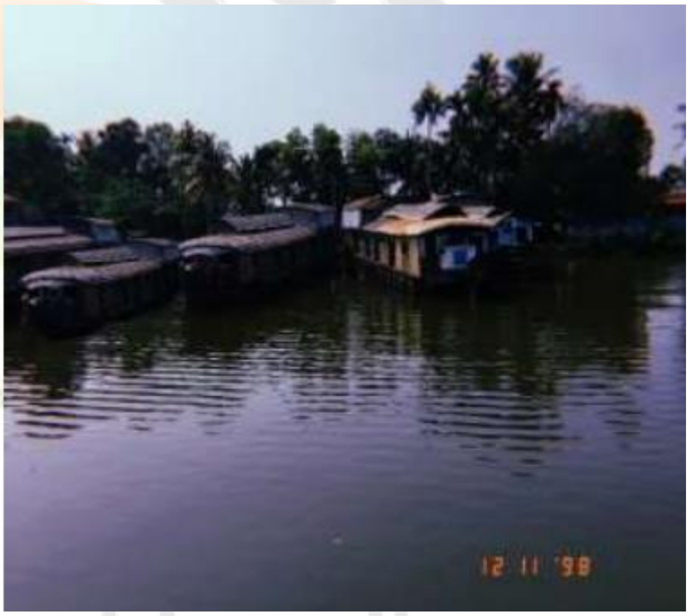


Fig. 1.1 Types of image tampering techniques

A. COPY-MOVE:





.

Fig. 1.2 original image

Fig. 1.3 forged image

B. IMAGE-SPLICING:

This type of image tampering technique is used when a person uses a part from one image and then pastes it on some another image without further postprocessing such as smoothing of boundaries between different fragments.

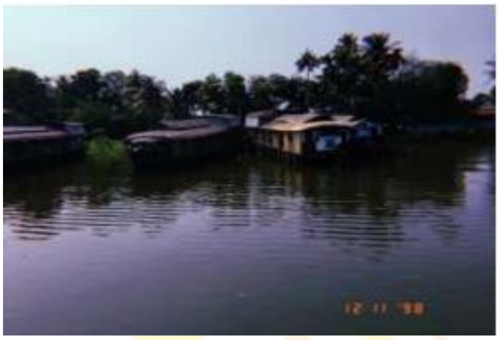


Fig. 1.4 first image

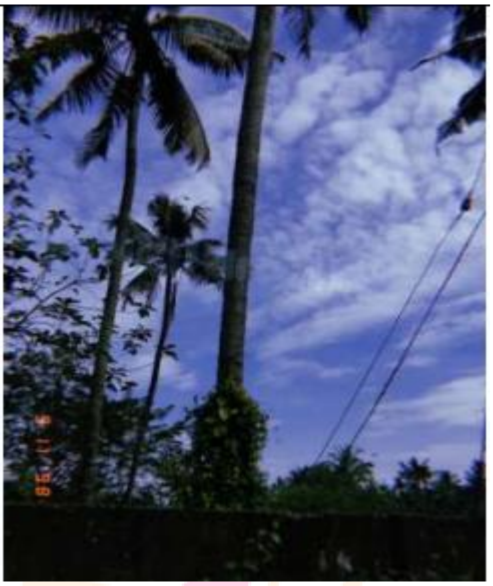


Fig. 1.5 second image

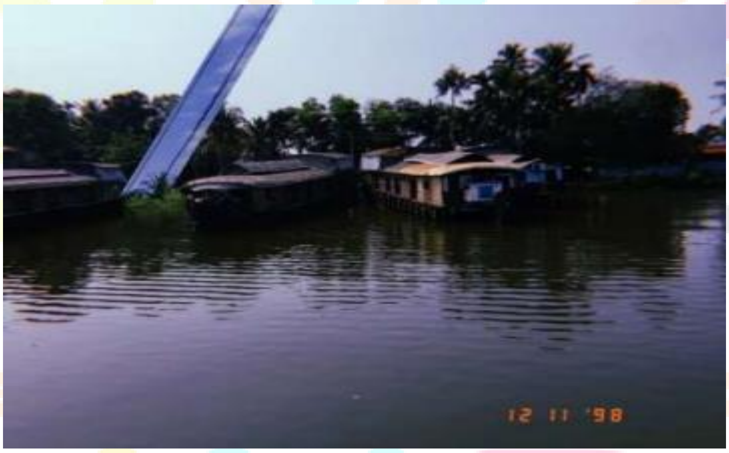


Fig. 1.6 tempered image

C. IMAGE- RETOUCHING

This type of image tampering technique is used when a person needs to alters an image to improve its appearance, retouching usually involves small localized adjustments to an image. Retouching an image can also be explained as polishing of an image by doing basic operations on it such as white balancing, cropping and adjusting other elements of an image

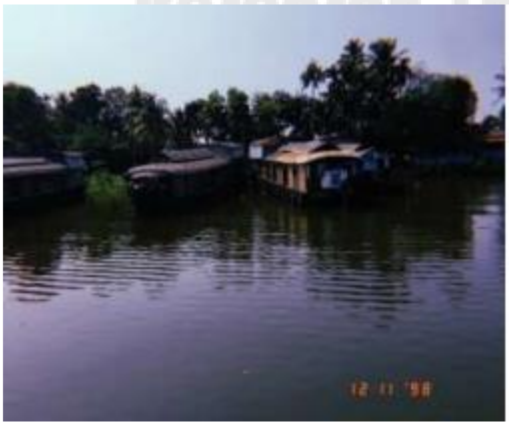


Fig. 2.1 original image

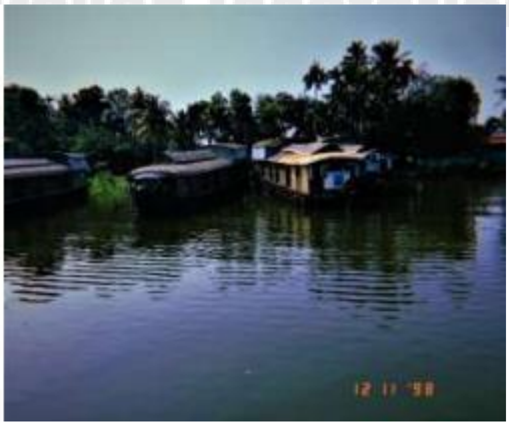


Fig. 2.2 retouched image

III. LITERATURE SURVEY

- S. Beram, et al. proposed a way for identify Doctoring in digital images. Doctoring typically involves several steps, usually in the

sequence of initial image-processing operations such as scaling, rotation, contrast shift, smoothing, and more. These methods used are
dedicated to three categories of statistical features including binary similarity, image quality and wavelet Statistics. The three categories of
forensic facilities are as follows:
1. Image quality measurements: They focus on the difference between the doctored image and its source. If the original image is unavailable, this test is simulated by a vague version of the image.
2. Higher order wavelet statistics: They are derived from multi-level decomposition of the image.
3. Binary similarity measurements: These measures capture the correlation and texture characteristics between and within the Bit planes of lesser importance, which are more susceptible to manipulation.
To influence the detection of doctoral effects, first, single tools are developed to identify the required image-processing functions. Then, these individual "weak" detectors assembled together to determine the presence of a doctorate in an expert fusion scheme.

- Swaminathan et al. [2] proposed a method for assessing camera and post-camera operation fingerprints for verifying the integrity

of the photos. A new methodology is operations. It primarily focuses on forensic analysis of digital camera images and observations based
on both internal and external acquisition devices.
The Internal fingerprints of several in-camera processing functions are made to assess by detailed imaging model and analysis of its
components. This is the primary function of Internal fingerprints. When further processing is enforced upon the camera, the image captured
is set to design as a manipulation filter so as to achieve linear time lapse estimation and to evaluate the internal fingerprints associated to
the operations performed later by the camera to blind deconvolution technique.
In cases when the test image is not a camera output but created by any other image processes that is due to the lack of fingerprints placed by the camera. Any diversity or change between new types of fingerprints and old projected camera fingerprints show that the image
underwent through some kind of process like stenographic embedding or image tampering after initial capture.

- H. Cao et al. [3] designed a new ensemble manipulation detector to detect a wide range of manipulation types on native image

patches.

- Fan et al. [4] proposed integrating statistical noise features with interchangeable image file format header features for manipulation

detection.

- M. C. Stamm and K. J. R. Liu, [5] proposed different methods not only for the detection of global and local contrast enhancement

but also for identifying the use of histogram equalization and for the identification of global addition of noise to a previously JPEGcompressed image.

The methods used are as follows.
I). Identification of globally applied contrast enhancement: contrast enhancement operations appear as non-linear pixel mappings that
display artifacts in the image histogram. Non-linear mapping is divided into locally mapped regions. Contract mapping maps multiple
unique input pixel values to a single output pixel value. Results in spontaneous peak in the image histogram.
ii). Improve the locally applied contrast image:
Contrast enhancement operation can be implemented locally to hide the visual evidence of image tampering. Localization of these activities may find evidence of cut-and-paste type forgery. Forensic techniques can be extended in a manner to detect such cut-paste forgery.
iii). Detecting histogram equalization in the image:
Similar to other contrast enhancement operation, the Histogram Equalization Operation introduces spontaneous peaks and gaps in the
image histogram. Methods for the detection of histogram equalization in the image have been extended.
iv). Image noise detection:
Additive noise can be applied to a picture globally
in an attempt to eradicate visual evidence of forgery as well as legally necessary indicators of other tampering tasks. The noise detection technique can detect whether the image is in speckled noise, Gaussian sound etc.

- M. Stamm and K. Liu [6] focuses on identifying the image change, focus on the functions used to retrieve and edit the potential

information about the unaltered version of the image. Even though the detection of such activities is probably not related to malicious
tampering yet they doubt the image's authenticity and content. An iterative method based on the likelihood model is proposed to estimate
the contrast enhancement mapping used to unify the image. The likelihood model identifies the histogram entries that occur with the

V. PROPOSED METHOD

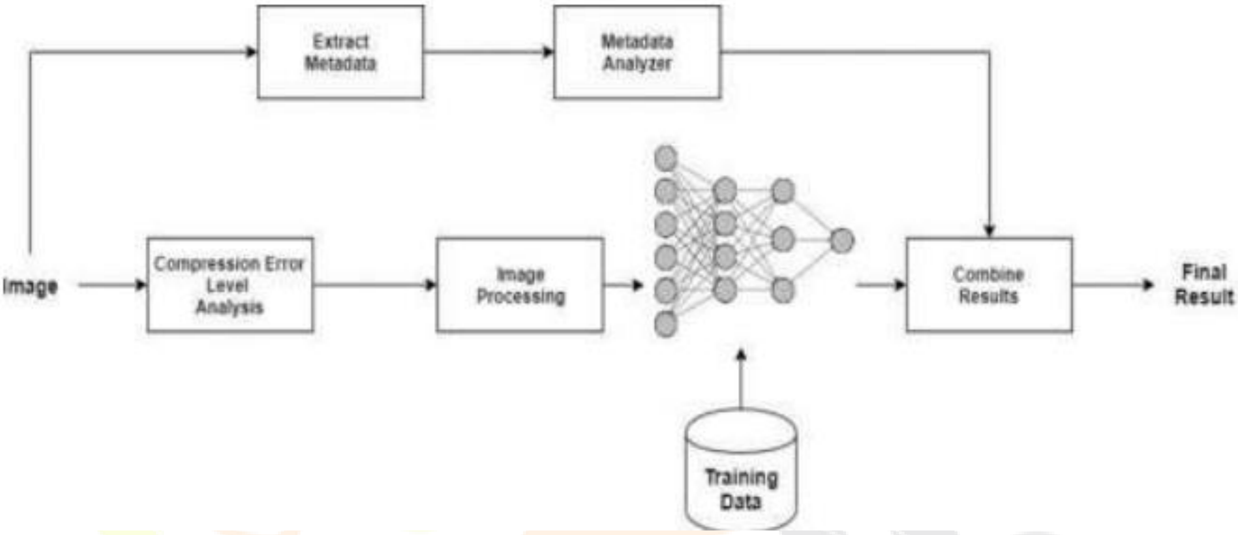
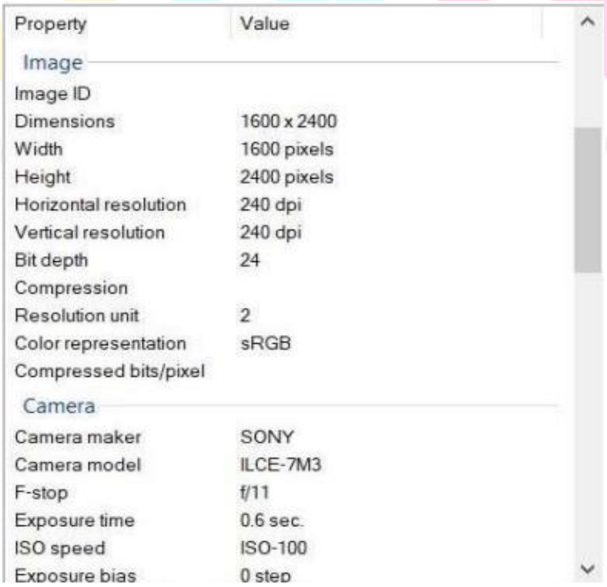


Fig. 3.1 Proposed method architecture

1. Metadata Analysis

Metadata analyzer is basically a tag selecting and searching algorithm. If keywords like Photoshop, Gimp, Adobe etc. are found in the text and then the possibility of being tampered is increased. Two separate variables are maintained which are called fakeness and realness. All of the variables represent the weight of being real or fake. Once a tag taken, is analyzed and corresponding variable is incremented by some predefined weight. These properties are already added in the photos by the cameras and the photo manipulation software's if they are used, but can easily be tampered or changed and cannot be trusted and should be only be used for preliminary analysis.



Property	Value
Image	
Image ID	
Dimensions	1600 x 2400
Width	1600 pixels
Height	2400 pixels
Horizontal resolution	240 dpi
Vertical resolution	240 dpi
Bit depth	24
Compression	
Resolution unit	2
Color representation	sRGB
Compressed bits/pixel	
Camera	
Camera maker	SONY
Camera model	ILCE-7M3
F-stop	f/11
Exposure time	0.6 sec.
ISO speed	ISO-100
Exposure bias	0 step

Fig 3.2 Metadata of an image

2. Error Level Analysis

Error Level Analysis resaves a particular image at a certain error rate, for example 96%, then it checks for a virtual change, if any detected it means that the cells have reached its local minima for error at that quality level. Nevertheless, if many changes are discovered, then the pixels tend to be original. These analyses differentiate the real pixels from the fake one. The system first saves an image at 100% quality and then it is resaved into the 90% quality image. The difference between these two is found out though difference method. The resultant image is the required error level analysis (ELA) image of the input image. Now, this image is saved as a buffered image and sent to the neural network for further processing. Example: - Error Level Analysis suggests a certain form of digital manipulation by showing variable error levels throughout the image. Areas like eyes, hair-color, nose and clothes are noted. All of these features are eventually at different error levels rather than their surroundings. This suggests that mostly several areas are brightened and the colors have been altered.

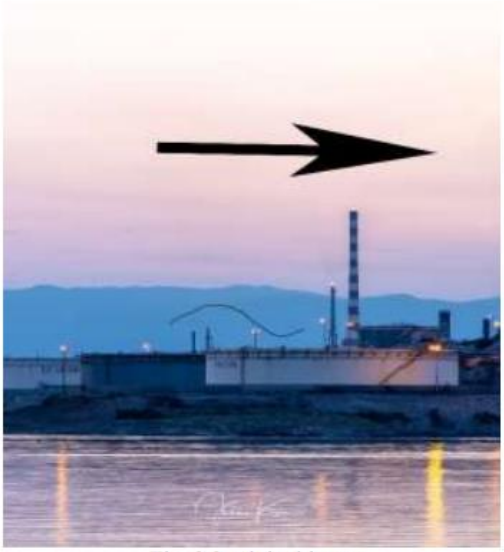


Fig. 3.3 original image

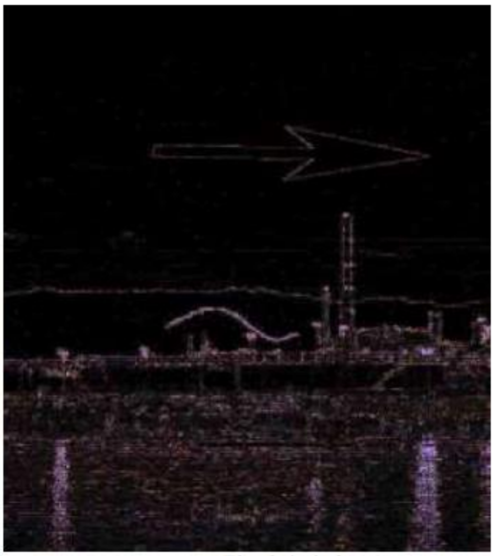


Fig. 3.4 ela image

Convolutional Neural Network

A multilayer perceptron neural network having an input layer and output layer both with a few hidden layers. When the image is opted for evaluation, the image is first converted for ELA representation from Compression and Error Level Analysis stage. The second step includes calculating the ELA since 90% of the images are used to construct ELA image. In the next step, the image is preprocessed and converted into 100x100px width and height. The image is serialized into an array containing approximate 30,000 integer values while representing 10,000 pixels. These pixels include red, green, and blue components; therefore, 10,000 pixels will have 30,000 values. When the data is being trained, the array will be given as input to the neural network and output neurons also set. The 2 output neurons represent fake and real image. If the image is fake, then the neuron is set to one while if the neuron is real then it is set to zero. During testing, the image array will be fed into the input neurons and values of output neurons will be taken to display the result of the analysis.

Transfer Learning

Transfer Learning transfers information from one domain to the required domain by improving the learner. It is a technique where a model is developed for one task but then reused as a starting point for another task. Let's take an example of two people and teach them drums. The first person has never learned or played any kind of musical instrument before while the other has an extensive knowledge of playing xylophone. Hence, the person with the experience of xylophone will learn drums way more efficiently by transferring previously learnt knowledge and apply it in the current task. Transfer Learning is required when there is a narrow amount of target training data. The major reason for that maybe data being pricy and rare. But we need to use it because it decreases the training time for a model and gives a lower error level.

VGG 16 Model

VGG16 is a convolutional neural network architecture which focuses on having Convolution layer with stride 1 and same padding and maxpool layer of stride 2. VGG Network uses 3x3 convolutional layers putting right above each other. While the depth increases with each stack. It makes advanced changes over AlexNet since it replaces kernel sized filters to multiple 3x3 kernel sized filters.

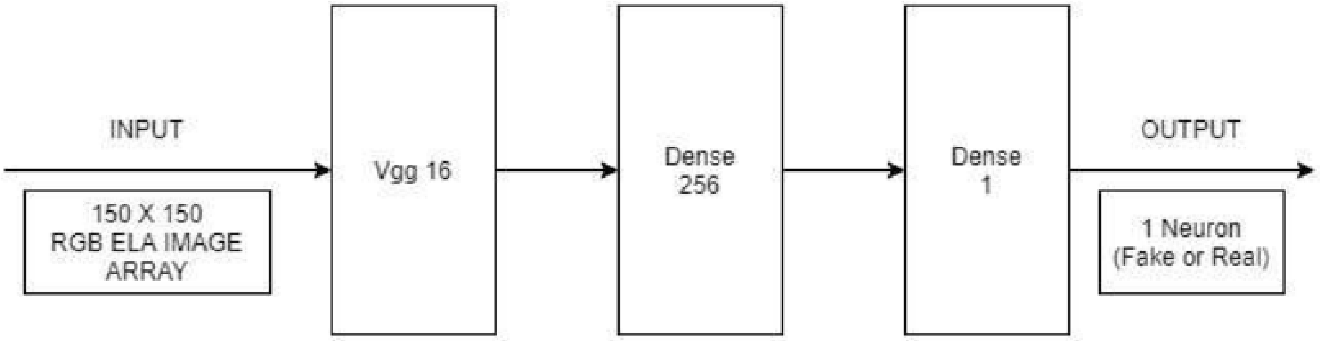
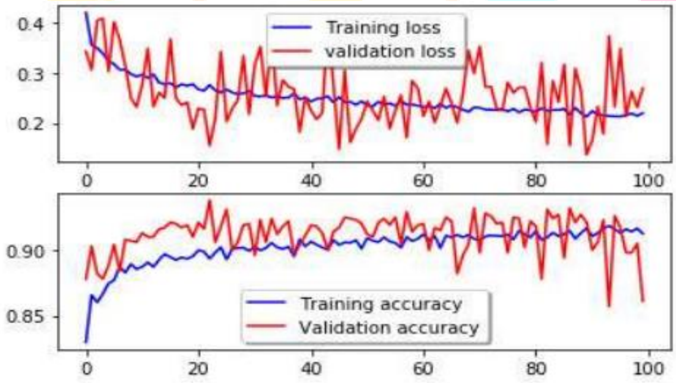


Fig. 3.5 Neural network architecture

V. RESULTS

From using transfer learning on the VGG16 model we were able to attain the following results. We can see from the graph below that in both Training and Validation the approach we used was significantly optimal as during both the training and validation in the graph it can be seen that our neural network was neither over-fitting or under-fitting , due to limited resources on our end we were still able to achieve a validation accuracy of 86.12%.



VI. CONCLUSION

As the internet advances rapidly in modern society, there are many social network services such as Facebook, Instagram and so on which have been used not only for good reasons but also some take the misuse them for negative purposes. Under these circumstances,
crimes against images are appearing for illegal purposes. Digital forensics need to detect such illegal purposes.

In this paper, we proposed image manipulation detection techniques using error level analysis. After we briefly observed the related works, the proposed model was explained in detail. Through intensive experiment the proposed model was analyzed and showed that at
least 95 % accuracy was achieved.

The proposed model can be used to determine whether or not the image is manipulated, and can be applied for detection of more manipulation techniques if a better model is established in later studies. In addition, it will be possible to apply it to various multimedia as
well as videos in the further research. Under these circumstances, crimes against images are appearing for illegal purposes. Hence, digital
forensics need to detect these illegal purposes.

VII. FUTURE WORKS

Here are a few more approaches we are trying to get more accuracy:

1. more aggressive data augmentation
2. more aggressive dropout
3. use of L1 and L2 regularization (also known as "weight decay")
4. Fine-tuning one more convolutional block (alongside greater regularization)

VIII. REFERENCES

1. Luo, Weiqi, Jiwu Huang, and Guoping Qiu. "Robust detection of region-duplication forgery in digital image." Pattern Recognition, 2006. ICPR 2006. 18th International Conference on. Vol. 4. IEEE, 2006.

2. S. Gholap and P. K. Bora, Illuminant colour based image forensics, in Proc. IEEE Region 10 Conf. 2008

3. Leida Li, Shushang Li, Hancheng Z -Journal of Information Hiding and Multimedia Signal Processing, Vol. 4, No. 1, pp. 46-56, January 2013.

4. Tiago and Christian et al Exposing Digital Image Forgeries by Illumination Color Classification. IEEE Transactions on Information Forensics and Security (Page: 1182 1194)Year of Publication: 2013.

5. Reshma P.D and Arunvinodh C IMAGE FORGERY DETECTION USING SVM CLASSIFIER
Conference on Innovations in Information, Embedde and Communication Systems (ICIIECS), 2015.
6. S.Shaid."TypesofImageForgery."Internet:http://csc.fsksm.utm.my/syed/research/image-forensics/11-types-of-mageforgery.html,
Feb.08, 2010 12:17 [Dec. 4, 2012].
7. Z. He, W. Sun, W. Lu, and H. Lu. "Digital image splicing detection based on approximate run length," Pattern Recogn .Lett., vol.
32, pp. 1591-1597, 2011