IMAGE FORGERY DETECTION USING INTEGRATED ERROR LEVEL ANALYSIS

[1] SAFIA NAVEED .S

[1]Department of Computer Science & Engineering
Jerusalem College of Engineering
Chennai, India

Abstract : Capturing photos and images is been popularly flourishing in recent times, owing to the widespread availability of cameras. Images are indispensable in our daily lives because they contain a lot of useful and important information, and it is often required to improve images to obtain additional information if one is required. A variety range of tools are available to improve a image quality; nevertheless, the tools are also frequently used to fake images, resulting in the spread of misinformation, threats, emotional distress, influence public opinions and actions. This increases and develops the severity and frequency of image forgeries, which is now a major source of concern. This action of tampering with images which are either for fun or to give false evidence has also been resulted in a disaster in some of the cases. Forgery is done in such a way that it cannot be determined by the naked human eye. To overcome these consequences there are various types of machine learning algorithms that are to be implemented for a better result. So, these machine learning techniques are used to extract the digital signature i.e. to extract the faked area of the image differentiating whether an image has been manipulated or not and also finds the maximum rate of forgery.

Index Terms **-** Convolution Neural Network, Digital Signature, Error Level Analysis, Fake, Forgery, Images, Machine Learning Techniques.

I. INTRODUCTION

The project concentrates on detecting the fake image by building a Machine learning model using Convolution Neural Networks which both are the important parts of Artificial Intelligence methods, which is emerging technology in this technological world. This system is to detect the presence of copy or move and Image splicing forgeries using Convolution Neural Networks and also to test and display the authenticity of the image [1].

They may also be beneficial for a variety of additional purposes in the future; including image forensics especially for women's who undergo threats by trending technique called morphing [2]. This morphing is nothing simple editing an image with another and then merged which gives as a single real image. Not only for women's it can used in social places to identify weather is the shown image fake or not. Thus various Machine Learning techniques associated with Convolution Neural Networks have been implemented for either fine or coarse image splicing, where as a technique dealing with both needs to be devised [3].

It can provide a better coverage (recall) for spoof and generated fake sites than looking up systems. Contemporary advancements in photo editing tools and enhanced new software have made much impact in security issues with respect to digital and social media domain[4]. In detail, the forged images has been uploaded to the internet of social platforms to create a panic situation for the users. Those synthesized images with fake and illegal content can be used in social media without their knowledge, which may in future lead to cause several problems of telecasting or posting [5]. Hence, it is important for the image forensics to detect the images weather it is a forged image or a manipulated image. With the high motive on effectively detecting the fake images and to provide security, the proposed work focused on developing an Advanced Fake Image-Feature Network (AFIFN) based on machine learning methods[6].

The tampered images are detected using neural networks which also recognize the regions of the image that have been manipulated and reveal the segments of the original image. It can be implemented on Android platform and hence made available to common users [7]. Another feature used along with compression ratio is imaging Meta data. Although it is possible to alter metadata content making content it unreliable on its own, here it is used as a supporting parameter for error level analysis decision [8].

Nowadays everywhere a lot of images has been clicked and many edit for good looking images and for fun in that list, forgery with images have become more popular even apart from threatening one individual to robbery these have been played a vital role [9]. One biggest issue to be solved using the proposed system is image forgery in money transfer, in recent times editing of one's name and amount in online money transfer platforms like Google pay, phone pay etc. So, to identify whether the show photograph is morphed or not this technique helps in identification.[10]

II. LITERATURE SURVEY

According to "Image Forgery Detection Using Machine learning technique, J. Malathi, B. Narasima Swamy, Ramgopal Musunuri," Been utilized with achievement in a couple of employments, as classification of surfaces, steganalysis, and bowing zone. We build up a new image counterfeit marker creating unequivocal descriptors recently proposed in the steganalysis field reasonably joining some of such descriptors, and redesigning a SVM classifier on the available training set. The issue with the present making is that majority of them see certain highlights in pictures changed by a particular tampering method, (for example, duplicate move, joining, and so forth). This proposes the structure does network always transversely over different evolving frameworks. Mix of no fewer than two pictures to make a completely phony picture is known as Image structure [11]

According to "Image Forgery Detection Using Deep Learning By Recompressing Images, Syed Sadaf Ali, Iyyakutti Iyappan Ganapathi, Neetesh," Capturing images has been increasingly popular in recent years, owing to the widespread availability of cameras A variety of tools are available to improve image quality; nevertheless, they are also frequently used to falsify images, resulting in the spread of misinformation [12].

According to "Image Forgery Detection By Using Machine Learning, J.Malathi," Dense local descriptors and AI have been utilized with achievement in a couple of employments, as classification of surfaces, steganalysis, and bowing zone. We build up a new image counterfeit marker creating unequivocal descriptors recently proposed in the steganalysis field reasonably joining some of such descriptors, and redesigning a SVM classifier on the available training set. The issue with the present making is that majority of them see certain highlights in pictures changed by a particular tampering method, (for example, duplicate move, joining, and so forth). This proposes the structure does not work always transversely over different evolving frameworks. Mix of no under two pictures to make a completely phony picture is known as Image structure.

A few optimization techniques such as Particle Swarm Optimization (PSO), Discrete Particle Swarm Optimization (DPSO), and Fractional OrderDiscrete Particle Swarm Optimization (FODPSO) Techniques based on which optimum features are selected. The region of interest can be captured and the same region can be inspected thoroughly in terms of pixel values, evaluating the degree of the noise present, analyzing the position of boundaries to study the region of interest, analyze the fluctuating intensities across the tongue region and study the texture to determine the abnormality. According to Nowadays, image manipulation is common due to the availability of image processing software, such as Adobe Photoshop or GIMP. The original image captured by digital camera or smartphone normally is saved in the JPEG format due to its popularity. JPEG algorithm works on image grids, compressed independently, having size of 8x8 pixels. For unmodified image, all 8x8 grids should have a similar error level. For resaving operation, each block shoulddegrade at approximately the same rate due to the introduction of similar amount of errors across the entire image. For modified image, the altered blocks should have higher error potential compred to the remaining part of the image. The objective of this paper is to develop a photo forensics algorithm which can detect any photo manipulation. The error level analysis (ELA) was further enhanced using vertical and horizontal histograms of ELA image to pinpoint the exact location of modification. Results showed that our proposed algorithm could identify successfully the modified image as well as showing the exact location of modifications [13].

From the interference of the review we have come to a conclusion that, they all have undergone each of the experiments utilizes various and numerous methods, strategies and algorithms. In spite they revolve around only one basis that is to detect the fake image and their results give a considerable amount of accuracy percentage of tampered images. But the major disadvantage in that, these methods cannot detect the real-time period in current decade due to empowering of new technologies..

III. EXISTING SYSTEM

Existing studies in image forgery detection have also worked and came up with many approaches on the comparison of image forgery; these approaches often limited in scope as well as only compare variants of the same algorithm on images that are specifically created for that type of methodology. It is conspicuous from the research of the existing systems that there is a need for advanced and most feasible version for image forgery detection than the traditional method There are also some forged images which cannot be detected by the existing project application due to, they fall under same algorithm limitations like following up only few traditional algorithms of deep learning. Added to it their accuracy of identifying the plotted area weather it is forgery or not varies with one another and does not exceed 90 percentages. Apart from this there are many drawbacks such as not using image splicing which results in false output of the detection

IV. PROPOSED SYSTEM

To overcome these drawbacks this proposed project is approached with various algorithms and techniques which undergoes several phases for example pre-processing unit which has two sub stages likewise the system trains with supervised Machine Learning Algorithms in collaboration with CNN , by using these algorithms the proposed system gives an accurate detection of image around 93% accuracy which is not accomplished by the existing systems The solutions which are currently available may provide a wide area of knowledge and satisfy the urgent needs of people.

Machine learning techniques now which holds the major portion of computer technology branches into three segments such as Supervised Learning, Unsupervised Learning and Reinforcement Learning. This implementation and whole system fall under supervised Learning.

This system with Machine Learning technique in collab with Convolution Neural Network techniques and algorithms together form the entire system, because this is not a traditional method so algorithms from different technique is been infused to derive the output. In order to verify the effectiveness of the proposed method, we tested and evaluated the algorithm.

V. METHODOLOGY

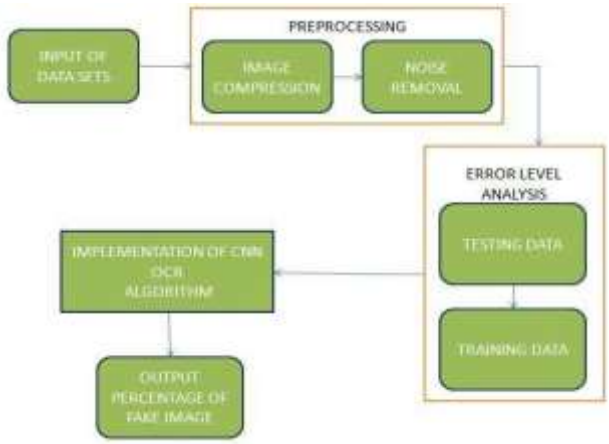

FIG. 1: BLOCK DIAGRAM OF SYSTEM MODULES

Image Acquisition: Collection of JPEG and PNG images is been used for the detection purposes normal camera photo copies are also collected from Nikon and canon. Input images can be given as much as a user want to identify its authenticity and forged areas.

Pre-processing stage: the extracted dataset (metadata) will be sent to the preprocessing area which is the heart of the proposed system where the data will be processed by separating the noise of the image with each pixel. By this way we can keep plotting the areas where is been tampered. The two main performances are BAG and Noise removal technique.

MLT Techniques: Applying of Machine Learning Techniques for the process is the next important stage of the project where various techniques like Convolution Neural Network and Optical Character Recognition is been used to identify the tampered area. Added to it the traditional technique called Error Level Analysis is also been used for exact derivation of forged images.

A. COLLECTION OF DATASETS:

The initial step is collection of data sets; we will be collecting the data's as Joint
Photographic Experts Group and Portable Network Graphics which is simply known as JPEG and PNG. The data's are collected through camera of Nikon and Canon and also internet
where we find huge availability of images and also from images taken from mobile phones
here since we provide 3000 input images instead of downloading each picture, the images are been downloaded from GitHub test images .
This is a phase were datasets are mixed up with real and fake images so that when we send these samples for testing the required output is deserved. The metadata which we collected should be only in JPEG or PNG format so that feature extraction takes place with proper executing format using this dataset.

B. PRE - PROSESSING OF DATASETS:

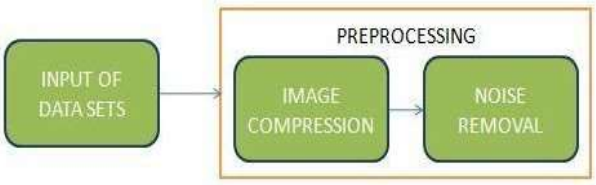

Fig 2: Input and Pre-Processing stage

Further, the extracted dataset (metadata) will be sent to the pre-processing area which is the heart of the proposed system where the data will be processed by separating the noise of the image with each pixel. By this way we can keep plotting the areas where is been tampered. Once the data is been collected the image goes through the following pre-processing units before final extraction of image.

BAG: The input data is been sent to the next sub stage for compression , so this JPEG compression is done because of it will identify some vertical or horizontal visual breaks in the image and these blocks are called Block Artificial Grid (BAG) this will appear in the border of each 8 X 8 pixel. Through this sub stage we can identify by 25.8% whether the image is altered or not.

Noise Removal: Now the image is sent for Noise Extraction, even if the image is compressed through BAG Algorithm in some circumstance it is not highly compressed and stored in high quality for some images. To increase the performance of the Noise removal system we use noise features to extract the forgery area.

C. ERRORR LEVEL ANALYSIS:

ELA is a method used in forensic which is designed toidentify the compressed area of an image that is when a image is been modified it definitely contains various levels of high definition impact when the image is been b=compressed it compresses each in different ratio according to the HD quality through this it is easily identified whether the image is tampered or not. This gives another 20% of conformation thisELA also contains certain algorithm that to be followed up for the process.

The compression levels are in the form of 8x8 gridwhich are obtained through the JPEG compression technique.JPEG is a lossy compression technique based on the combination between spatial domain and frequency domain inan image. In JPEG compression the raw data image is dividedinto 8x8 blocks and it is passed through a sampling and rgb transformation to reduce the amount of data for processing. Itis then transformed into the frequency domain using DiscreteCosine Transformation (DCT)

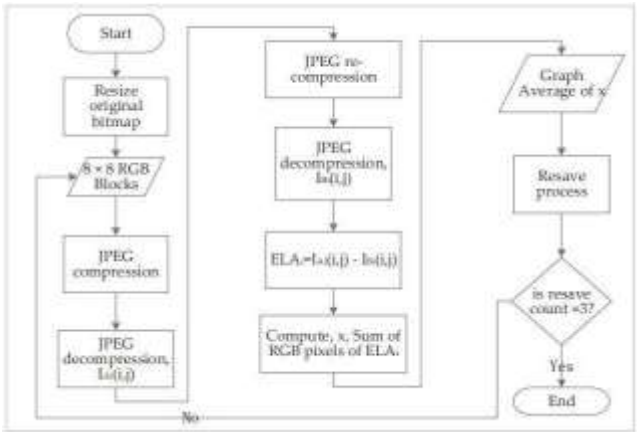

Fig 3: Error Level Analysis Algorithm Flow

The DCT coefficient is carried out in a zigzag manner and quantized using quantization table. The lossless entropy coding compress the quantization the coefficients to form a new compressed JPEG file.

The inverse process is occurred during JPEG decompression.The size of the JPEG compression image is depends on the size, content and quality of the image. The homogeneous image with few objects requires fewer bits of pixels as compared to an complex image with many bits. However, if the image is edited then the areas where the image have been manipulated should have a higher error potential than the other parts of the image . ELA works by intentionally resaving the image at a known rate of error, then measuring the difference between the images. When an image with JPEG extension is first saved, it compresses the photo for the first time. Many image editing software like Adobe Photoshop or GIMP supports the operation of JPEG compression. Therefore, if the image is opened in Photoshop, edited and saved again as a JPEG, the entire image will be compressed again. Because ofthis compression the original parts of the image have been compressed twice once by the camera that took the photo and again by Photoshop. Whereas, the "edited" part of the photographic image, were compressed once, by Photoshop. To the human eye, it is not possible to notice the difference bylooking at the image.

D. CNN ALGORITHM

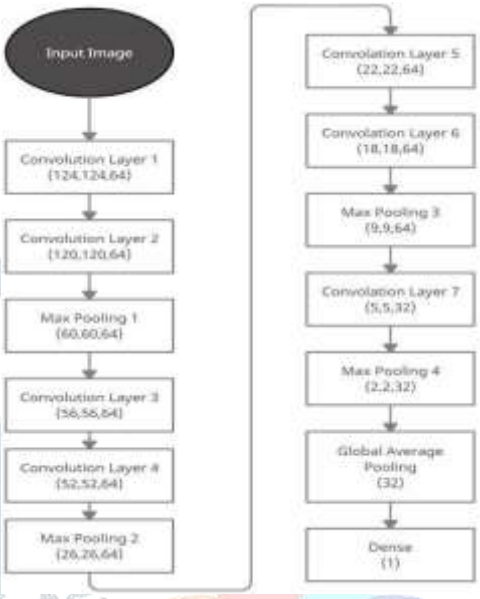

Fig 4 : CNN Algorithm Flow

The image after being recompressed it is been compared withthe original input image and the next most important part of the project is implementation of the revised algorithms of Convolution Neural Network and Optical Character Recognition through Error Level Analysis infused in the image as ELA image very after the noise removal process. The CNN algorithm consists of seven layers that are

- Padding Layer
- Convolution Layer
- Max – pooling Layer
- Convolution Layer
- Max –Pooling Layer
- Fully Connected Layer
- Fully Connected Layer

These are processed and then forwarded to the dense layer which is simple layer of the neuron were every neuron revivesinput from all the neurons from previous layer, so called as dense layer. Dense layer is used to classify image based on output from the convolution layers.

VI . PERFORMANCE METRICS

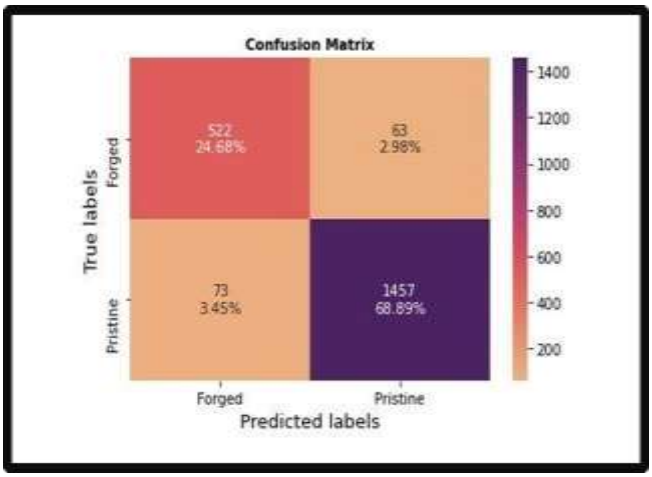

Fig 5: Confusion matrix

Classification Report

	precision	recall	f1-score	support
0	0.88	0.89	0.88	585
1	0.96	0.95	0.96	1530
accuracy			0.94	2115
macro avg	0.92	0.92	0.92	2115
weighted avg	0.94	0.94	0.94	2115

Fig 6: Classification report

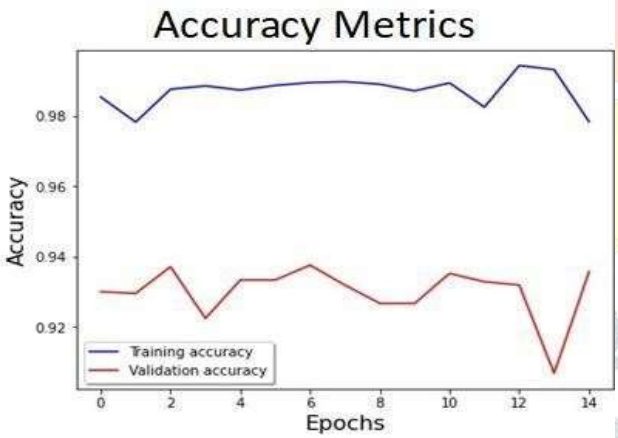

Fig 7: Accuracy Metrics

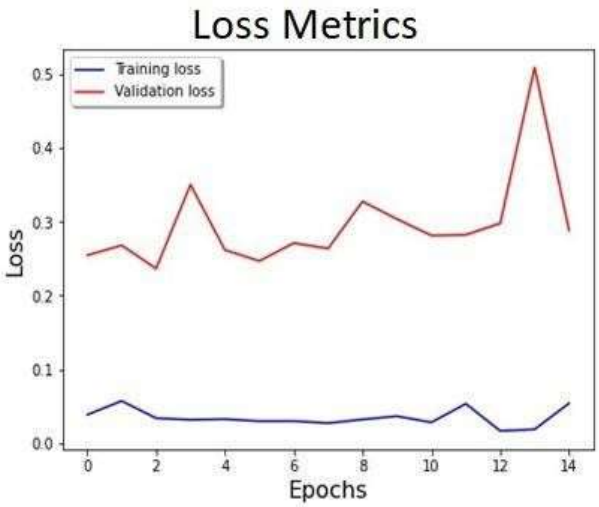

Fig 8: Loss Metrics

VII. ARCHITECTURE DIAGRAM

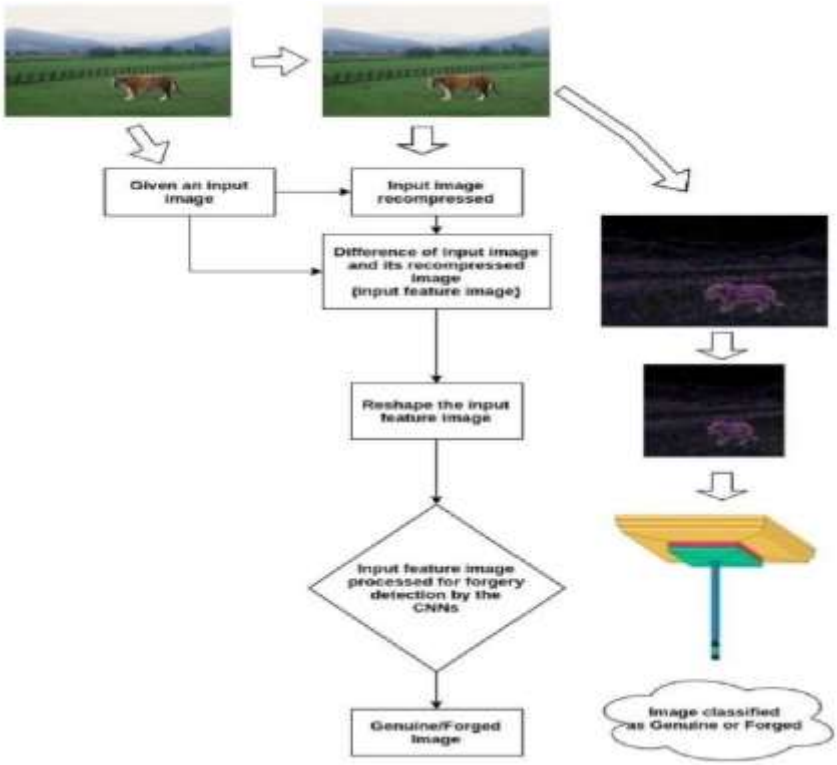

Fig 9: Architecture Diagram

VIII. RESULT

By proposing Integrated Algorithms to the existing algorithms we locate the forgery area by a single authentication process. The image is finally sent to the execution and then the desired output is derived weather the input image is been forgery or a real image with maximum accuracy

Fig 10 - Output of Forged Image

Fig 11 - Output Of Authentic Image

IX. CONCLUSION

A Machine learning model has been developed using Error level analysis and Convolution Neural Network and has been tested for detecting the authenticity of the images, we have achieved a testing accuracy of 93.89% and a validation accuracy of 93.57% for the model. The trained neuralnetwork was able to perceive the image as tampered or real ata maximum success rate of 94.8%. The usage of this proposed application in mobile platforms will greatly reduce the spread-out of fake images through the social media. This project system can also be used as a false proof technique in digital authentication, court evidences evaluation and especially for women's who undergo private issues etc

REFERENCES

[1] A picture's worth, Digital Image Analysis and Forensics, N Krawetz - 2007 Ph D, Hacker Factor Solutions

[2] http://imagej.net/WelcomeImageJ is an open source image processing program designed for scientific multidimensional images .J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73

[3] http://forensics.idealtest.org/CASIAv2.0CASIAV2.0 is with larger size and with more realistic and challenged fake images by using post-processing of tampered regions .It contains 7491 authentic and 5123 tampered color images.

[4] CNN based Image Forgery Detection using pre trained AlexNet Model(2018) – Amit Doegar, Maitreyee Dutta, Gaurav Kumar

[5] https://github.com/drewnoakes/metadata- extractor Metadata-extractor is a straightforward Java library for reading metadata from image files.

[6] https://www.github.com/afsalashyana/FakeImageDetection GitHub repositor for fake image detector desktop application written in javafx.

[7] Development of Photo Forensics Algorithm by Detecting Photoshop Manipulation using Error Level Analysis (2017)- Suriya Gunawan, Siti Amalina Mohammad Hanafiah, Mira Kartiwi, Anis Nurashikin Nordin.

[8] A deep learning Approach to Detection of splicing and copy - move forgeries in images (2016)- Yuan Rao, Jiangqun Ni http://neuroph.sourceforge.net/NeurophFrameworkNeuroph is lightweight Java neural.

[9] Network framework to develop common neural network architectures. Rong Fu, Dang – "University Classroom Attendance Based on Deep Learning", 2017.

[10] Boosting Image Forgery Detection using Resampling Features and Copy-move Analysis Tajuddin Manhar Mohammed

[11] Robust forgery detection for compressed images using CNN supervision BoubacarDiallo 2017

[12] Video Forgery Detection using Machine Learnin Gaikwad Kanchanl. https://www.irjet.net/archives/V6/i11/IRJET-V6I1180.pdf

[13]AHP validated literature review of forgery type dependent passive image forgery