ISSN: 2278-4632
Vol-13, Issue-04, No.06, April : 2023

FRAUD RESEMBLANCE IDENTIFICATION USING MACHINE LEARNING

[1]**Chitla.Vinay Santhosh**, Assistant Professor, Dept. of Computer Science and Engineering, DVR & Dr. HS MIC College of Technology, Kanchikacherla, Andhra Pradesh, India.
[2]**Dr. V.Sri Lakshmi**, Associate Professor, Dept. of Computer Science and Engineering, DVR & Dr. HS MIC College of Technology, Kanchikacherla, Andhra Pradesh, India.
[3]**Vaishnavi Nalajala**, UG Student, Dept. of Computer Science and Engineering, DVR & Dr. HS MIC College of Technology, Kanchikacherla, Andhra Pradesh, India.
[4]**Keerthi Bandi**, UG Student, Dept. of Computer Science and Engineering, DVR & Dr. HS MIC College of Technology, Kanchikacherla, Andhra Pradesh, India.
[5]**Raj Kumar Modugu**, UG Student, Dept. of Computer Science and Engineering, DVR & Dr. HS MIC College of Technology, Kanchikacherla, Andhra Pradesh, India.
[6]**Jessi Manda**, UG Student, Dept. of Computer Science and Engineering, DVR & Dr. HS MIC College of Technology, Kanchikacherla, Andhra Pradesh, India.

Abstract:
Image alteration has become so widespread in modern times thanks to the availability of image editing programmes like Adobe Photoshop or GIMP. For image-based cybercrimes to be exposed, it is necessary to find such phoney images. The JPEG format is frequently used to save images captured with a digital camera or smartphone. Using grids of 8x8 pixels that have been separately compressed, the JPEG algorithm operates. Unaltered photos, however, have a comparable error level. Since there are roughly the same number of mistakes throughout the entire image, each block should degrade at about the same rate during the resaving process. Using error level analysis, it is discovered that the compression ratio of this fake image differs from the actual images.
In this project, LBPNET, a machine learning convolution neural network, is being developed to identify false face photos. Here, we'll first extract LBP from the images before training a convolution neural network on LBP descriptor images to create a training model. Every time we add a new test image, the training model will use that image to determine if the image is false or not.

1. Introduction:

As a result of the huge expansion of technology in today's society, images are now one of the most widely utilized forms of communication. Images are used in publications like newspapers, magazines, websites, and advertising, and they may convey a variety of information. Due to their increased utilization in daily life, the trust in photographs is growing day by day. Image forgery is the act of tampering with or modifying an image by changing some of the information it contains, and image forgery detection is the process of determining whether the image is authentic or not.
Huge number of people have been victims of picture forgery in our current world. In order to deceive the court or numerous other people, many people modify photographs using image- editing software and present them as evidence. This is why it's important to assess and categorise every image published on social media as either real or fraudulent.
A significant media phenomenon is social networking sites these days and have drawn a lot of attention. The number of users [1] has surpassed three billion globally. The increase in the number of active users has surpassed 66% in the Gulf area [2]. More over 75% of Saudi Arabia's estimated 25 million residents [3] are active users of social media, placing the country ninth in the world in terms of social media usage.
Social media is one of the best venues for connecting with people, exchanging ideas, and disseminating information, but if safety measures aren't implemented, it can deceive users and cause chaos as a result of inadvertent false propaganda. While it takes some experience to photoshop images, some of them when edited by a professional might actually appear genuine. This is due to pixelization and dodgy jobs by novices. Images can be changed, especially in the political sphere, to increase or decrease a politician's credibility.
The believability of an image must be evaluated by a specialist in order to use the forensic techniques

DOI: 10.36893.JK.2023.V13I04N16.00107-00111

used today to change images. This method might work for a few photographs, but it is not advised to apply it while analysing many images, such as those on a social media platform. So, using the present machine learning methods at our disposal, we need to develop a system that can tell whether an image is real or fake and then make it available for usage by the general public.
This paper will then further break down into three suggested methodologies that can each be used to assess the originality of an image. First, we will concentrate on the metadata analysis, then on the LBPNET of the images, and finally, we will concentrate on developing a machine learning algorithm to assess the image.

2. Types of Tampering Techniques:

A. Copy Move:

This type of image tampering technique is used when a person needs to cover one part of the image in order to add or remove information, and is done with the help of textured regions from the same image because they have similar colour values, dynamic range, and noise variation properties in the image.

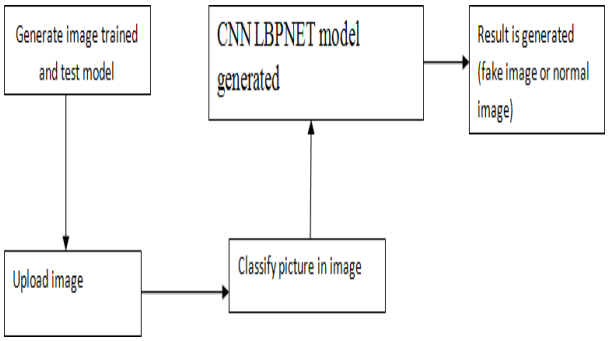
B. Image Splicing:

This kind of image manipulation technique is performed when someone copies a section of one image, pastes it on another image, and leaves the edges between the various parts unaltered.

C. Image Retouching:

Retouching is a type of image manipulation technique used when a person needs to make small localised adjustments to an image in order to improve its appearance. Retouching an image is also known as refining an image by doing fundamental activities such as cropping, white balancing, and image element adjustments on it.

Proposed System:



A) Meta Data Analysis:

A tag selection and search algorithm are essentially what a metadata analyzer does. The likelihood of being tampered with is raised if words like Photoshop, Gimp, Adobe, etc. are included in the text. The fakeness and realness variables, which are two different variables, are retained.
The significance of being real or phony is represented by each variable. After a tag has been collected, it is analyzed, and the corresponding variable is increased by a certain weight. These characteristics, which the cameras and any photo-editing software that may have been used have been added to the images, are not trustworthy and should only be used for initial research.

B) CNN:

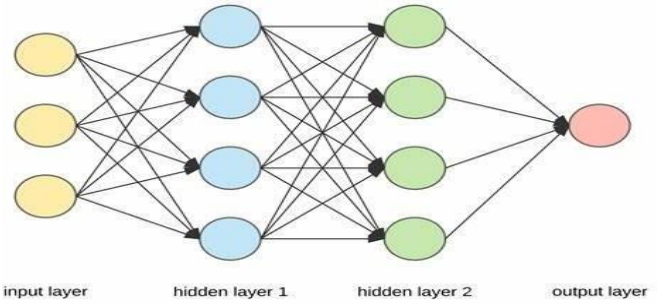
A neural network with many hidden layers that has an input layer, an output layer, and multiple hidden layers in between. When an image is selected the image is initially changed from the stage of compression and error level analysis for ELA representation for evaluation. The following step incorporates figuring out the ELA because 90% of the photos are needed to create the ELA image. The following stage involves per-processing and 100x100 pixels in width and height were transformed. An array with about 30,000 integer values is created from the image using serialization.
These pixels have elements of red, green, and blue, therefore 10,000 of them will have 30,000 values. The array will be provided to the neural network as input during training, and output neurons will also be set. Real and false images are represented by the two output neurons. The neuron is set to one if the image is fraudulent
if the neuron is real, it is set to zero. The image array will be fed into the input neurons during testing,

 DOI: 10.36893.JK.2023.V13I04N16.00107-00111

ISSN: 2278-4632
Vol-13, Issue-04, No.06, April : 2023

and the values of the output neurons willbe taken to show the analysis's results.

The Various Layers of Convolutional Neural Network is Convolution Layer, Pooling Layer, ReLU Layer, Fully Connected Layer.

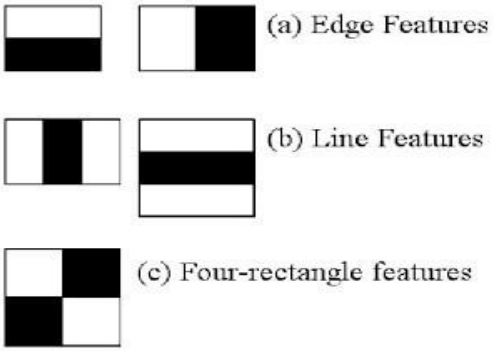


C) Haar Cascade:

A Haar classifier, or a Haar cascade classifier, is a machine learning object detection program that identifies objects in an image

1) Haar Features:

The first step is to collect the Haar features.A Haar feature is essentially calculations that areperformed on adjacent rectangular regions at a specific location in a detection window. The calculation involves summing the pixel intensitiesin each region and calculating the differences between the sums. Here are some examples of Haar features below:



Identifying these elements in a huge photographcan be challenging. Integral images come into play in this situation since they allow for a reduction in the number of operations.

2) Creating Integral Images:

Integral pictures, to put it briefly (check out thepaper if you're interested in the mathematics underlying it), essentially accelerate the calculation of these Haar features. It constructs sub-rectangles and array references for each of them rather than computing at each individual pixel.

3) Adaboost Training:

In essence, Adaboost selects the top features andtrains the classifiers to use them. To produce a "strong classifier," it combines several "weak classifiers," which the algorithm can then use to find things.

4) Implementing Cascading Classifiers:

The cascade classifier is composed of a number of stages, each of which contains a group of weaklearners. Boosting is used to train weak learners, resulting in a highly accurate classifier from the average prediction of all weak learners.

By classifying an object as a non-object, your object recognition method would be significantly hampered, hence it is crucial to maximize a low false negative rate. The Haar cascade is demonstrated in the video below. The red boxes represent "positives" from the underachievers.

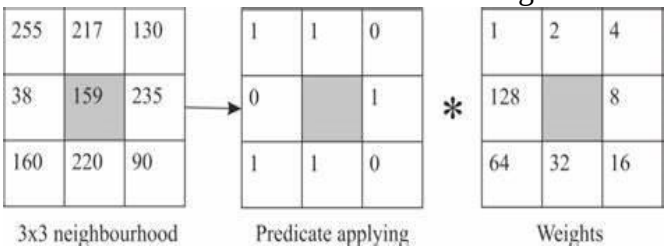
D) LBPNET:

Stands for Local Binary Points, where it uses center value to calculate the pixel value around the center value. If value is >= center value thevalue is given as 1 and if < then center value thevalue is given as 0.

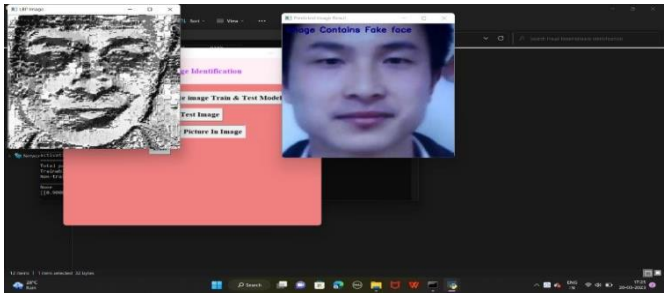
DOI: 10.36893.JK.2023.V13I04N16.00107-00111

ISSN: 2278-4632
Vol-13, Issue-04, No.06, April : 2023

LBPNET is used along with CNN to make the identification of image more accurate



3. Result Analysis:



We are seeing the results for the fake face in the image in the screen above. Similar to that, you can try different photos. If you wish to experiment with fresh photos, you must submit them to us so that we can familiarise the CNN model with them and enable it to recognize them as well

4. Conclusion:

Since the internet develops quickly in contemporary culture, there are several social networking platforms like Facebook, Instagram, and others that have been used both for positive and negative purposes. Images are being stolen under these conditions for nefarious reasons. Such unlawful intents must be found via digital forensics.

Using error level analysis, we suggested image modification detection methods in this work. After quickly reviewing the relevant publications, the proposed model was thoroughly explained. The proposed model was thoroughly tested, and analysis revealed that at least 95% accuracy was attained.

The proposed model can be used to identify whether or not the image is modified, and can be employed for detection of more manipulation techniques if a better model is produced in further investigations. In the future study, it will also be feasible to use it to different multimedia and films. Images are being stolen under these conditions for nefarious reasons. Digital forensics must therefore be able to identify these unlawful goals.

References:

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

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.

Reshma P.D and Arunvinodh C IMAGE FORGERY DETECTION USING SVM
CLASSIFIER Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), 2015.

Selling Stock. (2014). Selling Stock. [online] Available at: https://www. selling-stock.com/Article/18-billionimages-uploaded-to- the-web-everyd [Accessed 12 Feb 2019].

Li, W., Prasad, S., Fowler, J. E., & Bruce,
L. M. (2012). Locality preserving dimensionality reduction and classification for hyperspectral image analysis. IEEE Transactions on Geoscience and Remote Sensing, 50(4), 1185– 1198.

R. Raturi, (2018). Machine Learning Implementation for Identifying Fake Accounts in Social Network. International Journal of Pure and Applied Mathematics, 118(20), 4785-4797.

A. Krizhevsky, I. Sutskever, & G. E. Hinton, (2012). Image net classification with deep convolutional

 DOI: 10.36893.JK.2023.V13I04N16.00107-00111

neural networks. In Advances in Neural Information Processing Systems, 1097–1105.
K. Ravi, (2018). Detecting fake images withMachine Learning. Harkuch Journal
L. Zheng, Y. Yang, J. Zhang, Q. Cui, X. Zhang, Z. Li, et al. (2018). TICNN: Convolutional Neural Networks for Fake NewsDetection. United States
J. Bunk, J. Bappy, H. Mohammed, T. M. Nataraj, L., Flenner, A., Manjunath, B., et al. (2017). Detection and Localization of Image Forgeries using Resampling Features and Deep Learning. The University of California, Department of Electrical and Computer Engineering, USA.
M. Villan, A. Kuruvilla, K. J. Paul, & E. P. Elias, (2017). Fake Image Detection Using Machine Learning. IRACST—International Journal of Computer Science and Information Technology & Security (IJCSITS).
S. Shalev-Shwartz, & S. Ben-David, (2014). Understanding Machine Learning: From Theory to Algorithms. New York: Cambridge University Press.

 DOI: 10.36893.JK.2023.V13I04N16.00107-00111