FAKE IMAGE VERIFICATION USING MACHINE LEARNING

K. Manikandan

Assistant Professor, Department of Information Technology, G. Venkataswamy Naidu College (Autonomous), Kovilpatti

ABSTRACT

Presently, numerous fake images are being disseminated through digital media. The disclosure of image-based cybercrimes necessitates the detection of such counterfeit images. In this digital age, promising research areas include identifying forged images. It tended to be executed in the Android stage and made accessible to regular clients. Error Level Analysis is used to identify a fake image's foreign content's compression ratio, which differs from the original images. Image metadata is another feature used in conjunction with compression ratio.

Keywords: Image forensics, Error level analysis, Machine Learning, Counterfeit images and Metadata analysis.

Introduction

Image forgery has affected a large number of people in this technological age. Many people manipulate images with technology and use them as evidence to deceive the court. To that aim, every photo uploaded on social media should be correctly identified as real or fraudulent. Although social media is a fantastic tool for connecting with people, sharing knowledge, and educating others, its abuse may lead to confusion and even pandemonium as a consequence of inadvertent false propaganda. Although most photoshopped images clearly show manipulation due to pixelization and amateur work, some appear genuine. Manipulated images can make or break a politician's credibility, especially in politics.

Metadata Analysis

The bulk of image files contain more than just a picture. Image metadata is also included. They are confined within them. The metadata that records the history of a photograph includes application comments, colour space information, and the type of camera used. Various image styles contain various types of information.

Certain formats, such as BMP, PPM, and PBM, only retain a tiny amount of information in addition to the image's size and colour space. Unless the picture was edited with Photoshop or converted from a

JPEG, PNG files generally carry relatively little information. Metadata from the source file format may be present in the converted PNG file. The metadata describes how the file was generated and managed. This data can be used to determine.

Unless the picture has been edited with Photoshop or converted from a JPEG, PNG files normally carry relatively little information. The converted PNG file may contain metadata from the original file format. The metadata describes how the file was produced and how it was handled. This data can be used to evaluate whether the metadata was captured with a digital camera, processed by a graphical software, or looks to have been manipulated to present misleading information.

Model and Software

They are used to identify the programme or device that created the image. In the EXIF metadata block, most digital cameras include a Make and Model number. The first iPhone, on the other hand, does not. The Product may display the camera's firmware or application info.

Image sizes

The image's dimensions are frequently recorded in the metadata. Does the rendered image's size match the other measures in the metadata listed at the bottom? Many applications crop or resize images without updating any additional metadata.

Metadata Types

Metadata comes in a variety of forms. Only applications can create some, whereas cameras only have a few.

Error Level Analysis

Even though JPEG is a lossy compression format, each save does not introduce a line of error. The image will be altered so that stable areas will become unstable upon modification (no additional error). The figure depicts a Photoshop-modified embodiment. 1. The foundation for the altered image was the initial 75% resave. A toy dinosaur and duplicated books were added to the shelf. The 95% ELA identifies the changes because Photoshop effectively altered many pixels by combining information from multiple layers and no longer being at their minimal error level. The image's volatility is slightly higher in other areas.

The 80% picture saved at 80% equates to an 81% one-time save. In essence, holding a photograph once at 67.5% or twice at 75% yields the same result as holding it once at 75% or 90%. 19 The usage of 8x8 cells in the JPEG method decreases inaccuracy; after around 64 resolves, there is almost no difference. When a picture is altered, the 7x7 cells that contain the altered data no longer have the same error level as the rest of the image. ELA works by storing the picture at a certain error rate, such as 96%, and then calculates the difference between the two. The cell has arrived at its goal if there is practically no change.

Even if there is a lot of change, the pixels are still "original" since they aren't at their original minima. The quantum of error introduced by each save isn't direct, despite the fact that JPEG is a lossy format (1). rather, if the commodity is changed in the image, stable areas where there's no other error — will become unstable. A toy dinosaur and duplicated books were added to the shelf. The 95 ELA finds the changes because these areas are no longer at their minimum error position.

Other parts of the image exhibit a little bit more volatility as a result of Photoshop effectively altering many of the pixels by combining information from multiple layers. Nearly every pixel in the original image is not at its local minimum. In the first resave, large areas with pixels that have reached their local minima are visible (75%). Additional regions with reached local error minima are added in the second resave.

From the pattern on the left side of Figure 1, we can determine which parts of the ELA-applied image are created by the human eye and can see small changes on the fly.

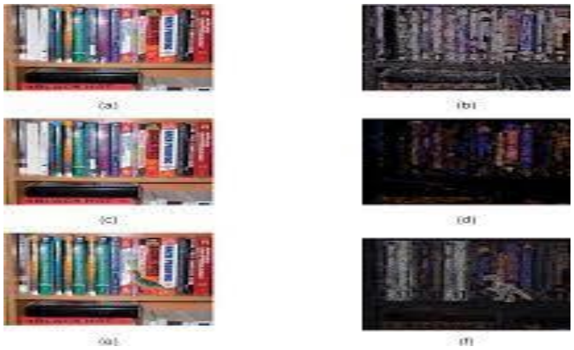


Figure 1 shows an error-level assessed image on the left and a false image on the right.

Machine Learning

Data mining and machine learning share certain similarities. Both systems can recognize patterns in data. In any case, rather than removing information for human interpretation, as in information mining applications, AI uses that information to detect patterns in data and change programmed actions. Machine learning algorithms are often classed as either supervised or unsupervised. New data may be applied to what supervised algorithms have already learned. Unsupervised algorithms can draw inferences from datasets.

Facebook's News Feed personalizes each user's feed using machine learning. If a member regularly pauses scrolling to read or "like" the postings of a certain buddy, the News Feed will show more of that person's activities earlier in the meal. The software uses statistical analysis and predictive analytics to find trends in the user's data, which is subsequently utilized behind the scenes to populate the News Feed. They track how users like, comment on, share, and interact with different types of material. These behaviours will permanently affect the composition of the news stream.

System Design

Metadata Analysis

The Java programming language is used to develop the entire system. The metadata-extractor library is used to extract image metadata. The metadata information for many different image types can be removed using the metadata extractor [5]. When a picture is chosen for handling, it is burrowed into two separate stages. Metadata analysis is the initial step. The metadata analyzer is essentially an algorithm for searching for tags.

If terms like Adobe, Gimp, Photoshop, and so on are found in the text, the likelihood of being altered increases. Fakeness and realness are the two distinct variables that are kept track of. The weight of an image, is genuine or fake, is represented by each variable. When the tag is removed, the comparison variable is increased by a predetermined weight. The preceding chart compares weight gains and addresses catchphrases. When all tags have been processed, the final values of the fakeness and realness variables are sent into the output stage, as illustrated in Fig. 2.

Table 1 shows a list of keywords.

List Of Keyword	Realness / Fakeness	Inc. Value
Photoshop	Fakeness	5
Gimp	Fakeness	5
Corel	Fakeness	5
Adobe	Fakeness	3
Exif Info	Realness	2
Camera Tags	Realness	2

Error Level Analysis:

Imagery is utilized for error-level analysis [2]

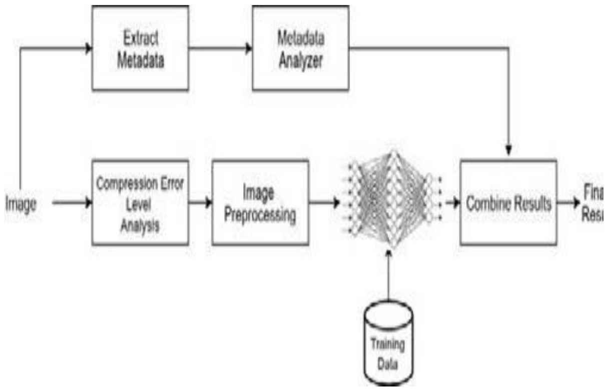


Image library lets you save images with a certain percentage of compression in the

JPEG format. A picture is first saved at 100% quality by the system. The same concept is then turned into an 80 %--quality image with ImageJ, and the difference between the two is calculated using the difference technique. The picture generated is the ELA image necessary for the input image.

Machine Learning

A Neuroph [4] Java library is used to implement machine learning. Neuroph is chosen for its ease of use and simplicity in putting neural networks into action. We have implemented a momentum back propagation learning rule multilayer perceptron network. Table 2 depicts the structure of a neural network.

One input sub-caste, three retired layers, and one affair sub-caste comprise a multilayer perceptron neural network. At the contraction and error position analysis stage, the picture chosen for assessment is transformed into an ELA representation. Following the computation of ELA, the concept is preprocessed to be turned into a 100x100px range and height. This is completed with 90 of the photos. The picture is republished into an array after preprocessing. Each integer number in the collection represents 2,000 pixels.

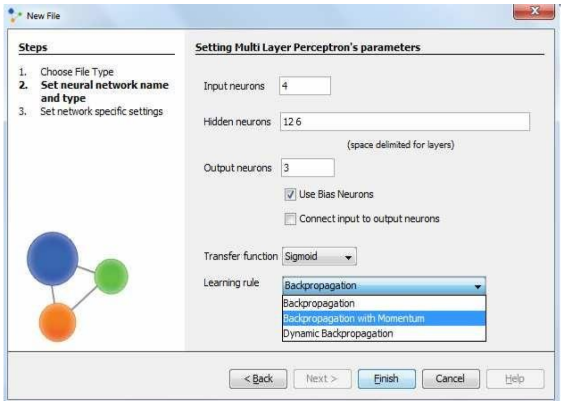
During training, more neurons are added, and the array is used as input to the multilayer perceptron network. The MLP is a neural network that is entirely linked. Two affair neurons exist. The alternate neuron represents the picture, whereas the first neuron represents the fake. If the picture is fraudulent, the genuine neuron is put to zero and the artificial neuron is set to one. Otherwise, absolute and unnatural are set to 0 respectively.

To change the neuronal connection weights, a momentum backpropagation learning algorithm was used. It is a supervised learning rule that attempts to reduce the error function. Table 3 displays the chosen learning rate, momentum, and efficiency. During testing, the image array is fed into the input neurons, and the output neural network's values are recorded. We utilized the sigmoid activation function.

Table 2. Structure of Neural Networks

Layer	Remarks
Input Layer	30,000 neurons
Hidden Layer 1	5000 neurons, Sigmoidactivation function
Hidden Layer 2	1000 neurons, Sigmoidactivation function
Hidden Layer 3	100 neurons, Sigmoid activation

	function
Output Layer	2neurons



Steps

1. Choose File Type
2. **Set neural network name and type**
3. Set network specific settings

Setting Multi Layer Perceptron's parameters

Input neurons: 4
Hidden neurons: 12 6
(space delimited for layers)
Output neurons: 3
☑ Use Bias Neurons
☐ Connect input to output neurons
Transfer function: Sigmoid
Learning rule: Backpropagation
Backpropagation
Backpropagation with Momentum
Dynamic Backpropagation

Fig 3. Neuroph framework

Experimental Result

In non-shared images, metadata analysis has produced promising results. Under minimal processing, it can identify anomalies in all "photoshopped" or "gimped" images. Photos shared via WhatsApp, Google+, and other platforms failed. In addition, when images with manipulated metadata were provided, it became utterly incorrect.

The CASIA dataset is used to train the neural network [3]. The dataset has 7491 authentic images and 5123 altered images of varying sizes. In the neural network with a 30,000-pixel value, each image has been preprocessed to 100x100 pixels. We used 4000 real and fake photos from the dataset for our training. The neural network was tested on the remaining images. The efficiency of various neural network configurations is depicted in Table 3. Best results are obtained with momentum set to 0.7 and a learning rate of 0.2.

Table 3 shows the outcomes of neural network training.

Rate of Learning	Momentum	Epoch	Efficiency
0.01	0.5	500	60%
0.05	0.5	500	62%
0.1	0.5	500	68%
0.2	0.5	500	66%
0.1	0.4	500	69%
0.1	0.3	500	68%
0.1	0.6	500	75%
0.1	0.7	500	76%`
0.2	0.7	500	82%
0.2	0.7	1000	83%

Conclusion

The neural network was successfully trained using an error-level analysis on 3000 false and 3000 actual photos. With an 83-percent success rate, the trained neural network was able to distinguish whether the image was actual or phony. Fake photos will be far less likely to propagate through virtual entertainment if this program is used on portable devices. This project may also be utilized in automatic validation, court-proof assessment, and other applications as a fake confirmation approach. A reliable false picture-detecting algorithm is constructed and tested by merging the findings of 40% of the metadata analysis and 60% of the neural network output.

References

Image J is a free and open-source software programme for improving multidimensional scientific pictures. A Treatise on Electricity and Magnetism, Third Edition, Vol. J. Clerk Maxwell 2. Clarendon Press, Oxford, 1892, pp. 68-73. http://imagej.net/Welcome

Welcome to ImageJ, an open-source program for processing scientific images with multiple dimensions. A Treatise on Electricity and Magnetism, 3rd ed., J Clerk Maxwell, vol. 2. Oxford: 68–73, Clarendon, 1892.

CASIA V2.0 uses post-processing of tampered regions to challenge fake images, making them larger and more realistic. It has 5123 altered and 7491 genuine color images.

The Neuroph Framework can be found at http://neuroph.sourceforge.net/. It is a lightweight Java neural network framework for creating standard neural network architectures. A well-designed open-source Java library with a few basic classes that correspond to fundamental NN concepts is included.