International journal of imaging science and engineering

ISSN: 1934--9955 www.ijise.net

Vol-16 Issue-01 Mar 2021

A Review Paper on Digital Image False Positive Analysis

Kailash K,

India

Abstract: In today's digital world, many false pictures are circulating. The exposure of image-based cybercrimes will inevitably include the detection of such false pictures. In this digital age, studies into picture forgeries and methods for detecting them hold great promise. By training a neural network to identify altered areas, it is possible to uncover previously hidden parts of the original picture and identify instances of manipulation. Since it is compatible with the Android platform, it may be made accessible to the general public. Using Error Level Analysis, it is possible to identify phony images with foreign content by comparing their compression ratios to the originals. Metadata for images is another element utilized in conjunction with compression ratio. Despite the fact that metadata may be manipulated to make it unreliable when used alone, it is used here to enhance the decision-making process for error level analysis.

Keywords: Fake pictures, ELA, social media, digital fakes, fake photographs, Image Forensics

1. Introduction

Over the past few years there has been increase in the usage of Online Social Media (OSM) services as a medium for people to share, coordinate and spread information about events while they are going on. Though a large volume of content is posted on OSM, not all of the information is of good quality with respect to the event, like it may be fake, in-correct or noisy. Extracting good quality information is one of the biggest challenges in utilizing information from OSM. Over last few years, people have highlighted how OSM can be used to help in extracting useful information about real life events. But, on the other hand, there have been many instances which have highlighted the negative effects on con-tent on online social media on real life events. While digital photos are conveniently used, their credibility has been severely challenged due to numerous fraudulent cases involving image forgeries, e.g. the fake results on human stem-cell research. With the availability of powerful image editing tools, numerous image retouching techniques have become practical, which can be used to create great artistic works. However, malicious modification of image content forms a serious threat to the secure and legal usage of digital images. By skillful manipulation, forgery may be very difficult to recognize by the naked eye. Therefore, automatic detection of image forgery has attracted much research interest. In recent years, many image forgery detection techniques have been proposed, especially passive approaches which do not require any additional information besides the image itself. Some published methods make use of lighting abnormality; blur moment invariants, and similarity/dissimilarity of color and structural characteristics. Digital tools have enabled easy image creation, modification and distribution, which make fraudulent image forgeries easier than ever. Fakes are created either by merging two or more photos or altering an existing image. Because image manipulation happens at the pixel level, detection is not as easy as it was before the digital era. Tricky fakes can be exposed by algorithms that detect discrepancies or statistical irregularities at the bit level. An image is authentic if it represents a witness to an actual event, place, or time. A definition of image authenticity should enable us to distinguish an authenticimage from the fake images, such as the 2D composite images and the 3D computer graphics images. It is still a problem how to detect whether digital images are fake or real. Generally, there is an obvious boundary between the fake area and the real area, with the improvement of desktop photograph manipulation software, which cannot be used to distinguish fake images and real images. There are a many studies related to detect the fake images. In this paper an ELA system is discussed which will scan the image uploaded and create a database where the number of errors will detect the fake image. The system uses MATLAB software for the computation process.

In this technological era a huge number of people have become victims of image forgery. A lot of people use technology to manipulate images and use it as evidences to mislead the court. So to put an end to this, all the images that are shared through social media should be categorized as real or fake accurately. Social media is a great platform to socialize, share and spread knowledge but if caution is not exercised, it can mislead people and even cause havoc due to unintentional false propaganda. While manipulation of most of the photoshoped images is clearly evident due to pixelization & shoddy jobs by novices, some of them indeed appear genuine. Especially in the political arena, manipulated images can make or break a politician"s credibility. Current forensic techniques require an expert to analyze the credibility of an image. We implemented a system that can determine whether an image is fake or not with the help of error level analysis and thereby making it available for the common public. In this paper we have discussed about the methodology used for detecting the fake image and the conclusion of the project.

2. Methodology

Error Level Analysis

JPEG is a lossy format, but the amount of error introduced by each resave is not line. Any modification to the picture will alter the image such that stable areas (no additional error) become unstable. Fig.1 shows a modified image using Photoshop. The modified picture was based on the first 75% resave. Books on the shelf were duplicated and a

toy dinosaur was added to the shelf. The 95% ELA identifies the changes since they are areas that are no longer at their minimal error level. Additional areas of the picture show slightly more volatility because Photoshop merged information from multiple layers, effectively modifying many of the pixels. A 90% image resaved at 90% is equivalent to a one-time save of 81%. Similarly, saving an image at 75% and then resaving it at 90% (75% 90%) will generate virtually the same image as 90% 75%, or saved once at 67.5%. 19 The amount of error is limited to the 8x8 cells used by the JPEG algorithm; after roughly 64 resaves, there is virtually no change. However, when an image is modified, the 8x8 cells containing the modifications are no longer at the same error level as the rest of the unmodified image. 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".

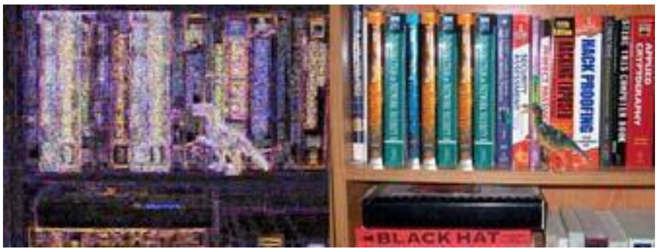

Figure 1

Books on the shelf were duplicated and a toy dinosaur was added to the shelf. The 95% ELA identifies the changes since they are areas that are no longer at their minimal error level. Additional areas of the picture show slightly more volatility because Photoshop merged information from multiple layers, effectively modifying many of the pixels. Nearly all pixels in the original image are not at their local minima. The first resave (75%) shows large areas where the pixels have reached their local minima. The second resave introduces more areas that have reached their local minima for error. By analyzing the pattern in the ELA applied image (Fig 1 left part), we can determine which part of the image is possibly faked. It is hard for the human eye to detect small scale changes to image so that we have decided to use machine learning to detect the anomalies in the error level analyzed images.

Error level analysis is done with the help of ImageJ library. ImageJ provides option to save image in JPEG format with certain percentage of compression. The system first saves an image at 100% quality. Then the same image is converted into 90% quality image using ImageJ. The difference between these two is found out though difference method. The resulting image is the required ELA image of the input image. This image is saved as a buffered image and sent to the neural network for further processing.

Machine learning is implemented using Neuroph library for java. Neuroph is selected because of the simplicity andeasiness to implement neural networks. We have implemented a multilayer perceptron network with momentum back propagation learning rule. A multilayer perceptron neural network is used having one input layer, 3 hidden layers and 1 output layer. Once the image is selected for evaluation, it is converted to ELA representation from Compression and Error Level Analysis stage. 100%, 90% images are used for the construction of ELA image. Once ELA is calculated, the image is preprocessed to convert into 100x100px width and height. After preprocessing, the image is serialized in to an array. The array contains 30,000 integer values representing 10,000 pixels. Since each pixel has red, green and blue components, 10,000 pixels will have 30,000 values.

Most image files do not just contain a picture. They also contain information (metadata) about the picture. Metadata provides information about a picture's pedigree; including the type of camera used, color space information, and application notes. Different picture formats include different types of metadata. Some formats, like BMP, PPM, and PBM contain very little information beyond the image dimensions and color space. In contrast, a JPEG from a camera usually contains a wide variety of information, including the camera's make and model, focal and aperture information, and timestamps. PNG files typically contain very little information, unless the image was converted from a JPEG or edited with Photoshop. Converted PNG files may include metadata from the source file format. Metadata provides information related to how the file was generated and handled. This information can be used to identify if the metadata appears to be from a digital camera, processed by a graphical program, or altered to convey misleading information.

3. Result

Error level analysis has shown promising result in non- shared images. It is able to detect anomaly in all „photo shopped‟ or „gimped‟ images under a very small processing. It failed on images shared through WhatsApp, Google+ etc. Moreover, it became completely erroneous when images with manipulated metadata given. Neural network has been successfully trained using the error level analysis with 4000 fake and 4000 real images. The trained neural network was able to recognize the image as fake or real at a maximum success rate of 83%. The use of this application in mobile platforms will greatly reduce the spreading of fake images through social media. This project can also be used as a false proof technique in digital authentication, court evidence evaluation etc. By combining the results of metadata analysis (40%) and neural network output (60%) a reliable fake image detection program is developed and tested. Error level analysis helps to find the changes in the image uploaded by the number of pixels in the image.

4. Discussion

Digital material may now be edited and transformed in ways that were unimaginable even a decade ago, all thanks to modern technology. We will most likely be able to control digital media in ways that are unfathomable now, thanks to tomorrow's technologies. It is crucial for the field of digital forensics to strive to stay up with the ever-changing technologies. I really hope that the tools my lab is developing will assist the media, the courts, and our society in navigating this fascinating and sometimes perplexing digital era. Raising awareness and enacting reasonable policies and laws are also important goals.

5. Conclusion

A system for detecting false images was developed in this study using the ELA approach. Some additional information is added to the photographs before they are made public. Concerning the confidential data, the study presented the mathematical ELA operation in a fabricated picture. This article successfully modifies earlier research. The outcome of the image's authenticity was determined using the MATLAB-math work program. The outcome is contingent upon the picture pixel that can be identified by the Error Level Analysis, which in turn determines whether any alterations have been performed or not.

Acknowledgement

Previous research publications on Identifying Fake Images using image processing were written by Hany Farid of Dartmoth College and Sangwook Lee of Texas Tech University. I am thankful to them. My college, ITM Universe, has been very helpful to me throughout this research process, and for that I am eternally grateful.

References

The value of an image, Forensic and digital image analysis, Dr. N. Krawetz of Hacker Factor Solutions earned a Ph.D. in 2007.

[2]Hello there! Scientific multidimensional pictures are the focus of ImageJ, an open-source image processing application. In his third edition of A Treatise on Electricity and Magnetism, Clerk Maxwell authored volume 2. Vol. 1, pages 68–73, Oxford: Clarendon, 1892.

the thirdVisit http://forensics.idealtest.org/ / CASIA V2.0. The new version is bigger, more realistic, and uses post-processing of tampered sections to challenge fraudulent photos. It has 7,499 original and 5,123 altered color photos.

[4]Here is the Neuroph Framework: http://neuroph.sourceforge.net/ Develop common neural network topologies using Neuroph, a lightweight Java neural network framework. The open-source Java library it comes with is well-designed and offers a modest number of classes that cover the fundamentals of NN.

[5]This URL: https://github.com/drewnoakes/metadata-extractor, One simple Java library that can read picture metadata is metadata-extractor.