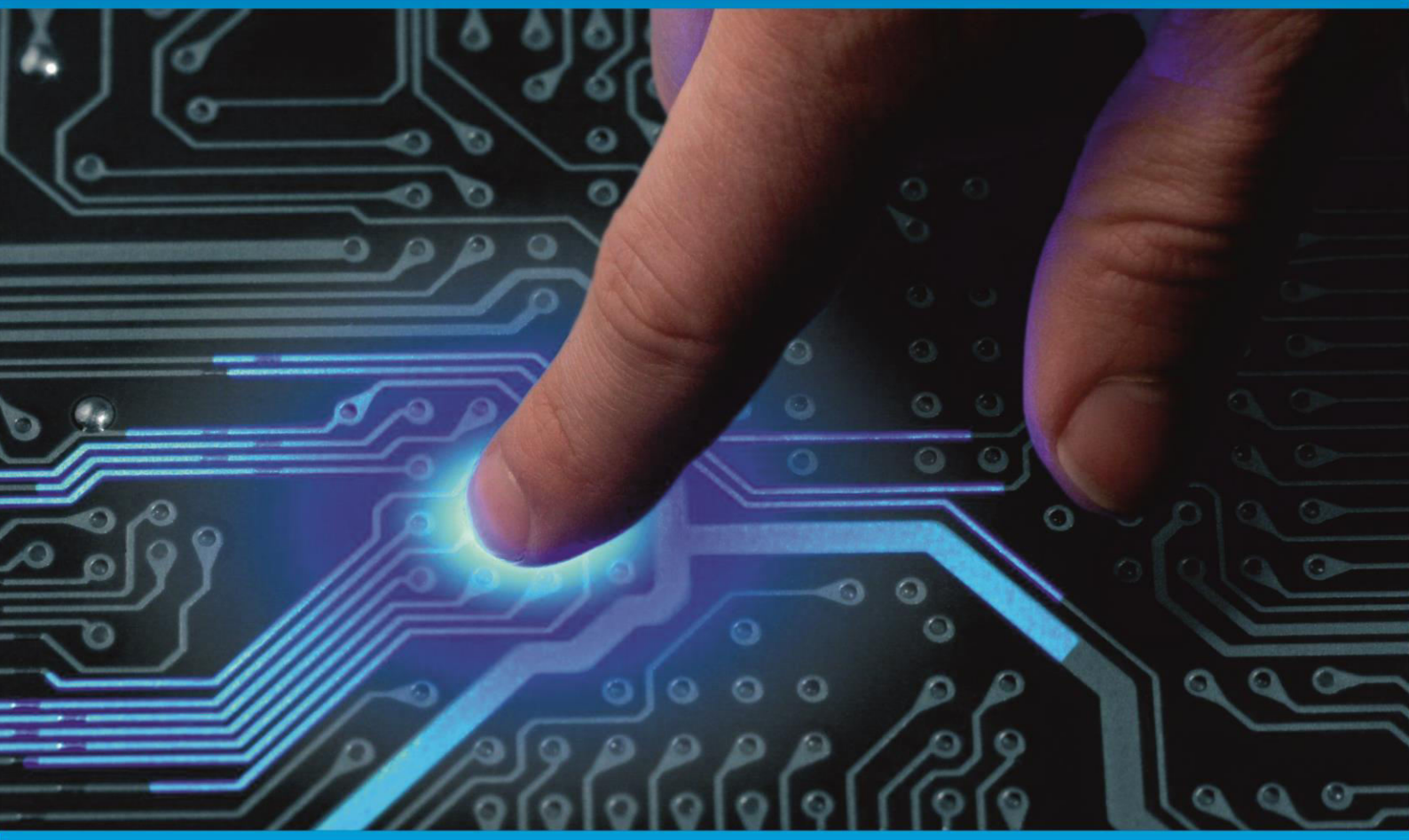
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 2, February 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.488

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 2, February 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0902029 |

Fake Image Detection Using Machine Learning

Prof. A.D.Gotmare [1] , Shital Hiwarkar [2] , Ruksar Shah [3] , Vaishnavi Dafare [4], Shehjad Sheikh[5], Ramesh Gugale [6]

Assistant Professor, Department of Computer Engineering, Bapurao Deshmukh College of Engineering, Sevagram, Wardha, India[1]

UG Students, Department of Computer Engineering, Bapurao Deshmukh College of Engineering, Sevagram, Wardha, India[2],[3],[4],[5],[6]

ABSTRACT: Nowadays, many false images are spreading in digital media. The detection of these false images is inevitable for the unveiling of image-based cybercrimes. Forging images and identifying such images are promising research areas in this digital era. Altered images are detected using a neural network that also recognizes the regions of the image that have been manipulated and reveals the segments of the image.
This original can be implemented on the Android platform and therefore made available to ordinary users. The compression ratio of foreign content in a false image is different from that of the original image and is detected using an error level analysis. The another function used with compression ratio is image metadata. Although it is possible to modify the metadata content, which makes it unreliable, here it is used as a support parameter for an analysis of the decision error level.

KEYWORDS: Image forensics, Metadata analysis, Error level analysis, Multilayer perception network, Deep neural networks

I. INTRODUCTION

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 machine learning and thereby making it available for the common public. This paper will unfold into three sections whereby first will focus on the second will focus on the Implementation details while the last part showcase the experimental result.

II. THEORY

A. Metadata Analysis

Most image files not only contain an image, they also contain details (about metadata) per image.Metadata provides details about the lineage of an image, including the camera type used, color space details, and application notes. Some models, such as BMP, PPM, and PBM contain very small details beyond the size of the image and space. PNG files usually contain very small details, unless the image is converted to JPEG or edited by Photoshop.Conformed PNG files can include metadata from source file format. This information can be used to determine whether the metadata appears to be from a digital camera, processed by a graphical system, or modified to transmit misleading information. Some common things to check are:

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 2, February 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0902029 |

1) Make, Model, and Software
These identify the device or application that made the photo. Most digital cameras include the make and model in a block of EXIF metadata. (However, the original iPhone doesn't!) Software can describe the camera's firmware version or application information.

2) Image size
Metadata often records the size of the picture. Does the rendered image size (listed at the bottom of the metadata) match the other sizes in the metadata. Many applications resize or crop pictures without updating other metadata fields.

3) Timestamps
Look for a field that details the timestamp. These usually identify when the photo was taken or changed. Does the timestamp match the expected time frame?

4) Types of metadata
There are many different types of metadata; some are only done with cameras, while others are only done with apps

5) Descriptions
Many photos have embedded annotations that describe the photo, identify the photographer, and itemize the changes.

6) Missing metadata
Missing metadata field? If the photo is from a digital camera, it must contain camera-specific information. For some applications and online services, the metadata is deleted. In the absence of specific metadata, it usually indicates a re-saved photo rather than the original photo.

7) Altered Metadata
Metadata is similar to chain of custody for evidence processing, it can determine how an image was last created, processed, and saved. However, some people intentionally change metadata.

B. Error Level Analysis
JPEG is a lost format, but the size of the error that appears each time it is scanned is not linear. Any change to the image changes the image so that solid areas (no additional defect) become unstable. Fig. 1 shows a modified image with Photoshop. The image was adjusted based on the first 75% rendering. Books have been duplicated on the shelf and a toy dinosaur has been added to the shelf. The 95% ELA identifies the changes as these are areas that are no longer at their minimum error level. Additional areas The image shows a little more volatility as Photoshop merged information from multiple layers, effectively adjusting many of the pixels. A 90% rescanned image with 90% corresponds to a one-time saving of 81%. Likewise, saving the image at 75% and then resetting it to 90% (75%) will produce almost the same image as 90u%, or saving once at 67.5% .19 The error rate is limited to the 8x8 cells used by JPEG algorithm; after almost 64 saves, no change Same error rate as the rest of the unconfirmed image Error rate analysis (ELA) works by deliberately resizing an image with a known error rate such as 95% and calculating the difference between images. If there is no change, is a cell At this level, however, in the case of a large number of changes, the pixels are not at the local minima and are successfully "prime". JPEG is the missing format, but the number of errors generated by each rewrite is not a linear image change that will alter the image so that stable positions (without additional error) are unstable [1]. 95% of the ELA indicates a shift in sin with areas that are no longer in their level of minor error Almost all pixels in the original image are at their local minima e-opening (75%) shows large areas where pixels have reached their local minimum. The second risave introduces more areas that have reached their local minima by mistake.

By analyzing the pattern in the applied ELA image (Fig 1 left part), we can determine which part of the image is probably false. to detect small-scale changes to the image, so we decided to use machine learning to detect discrepancies in the error level of error.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 2, February 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0902029 |

C. Machine Learning

The process of machine learning is similar to the process of data mining. Both systems search through data to find patterns. However, instead of extracting data for humans to understand as is the case with data mining applications, machine learning uses that data to detect patterns in the data and tailor program operations to fit. Machine learning algorithms are often classified as supervised or unsupervised. Monitoring algorithms can apply what has been learned in the past to new data. Unsupervised algorithms can draw inference from data sets. If a member frequently stops scrolling to read or "like" a particular friend's post, the News Feed will start showing ear more of that friend's activity than in the source feed. Behind the scenes, the software is using statistical analysis and predictive analysis to identify patterns in the user data and use the templates for inclusion in the News Feed. User activities such as like, comment, share etc. on different posts and based on these activities, the content on news feed will be adjusted continuously.

III. SYSTEM DESIGN

A. Metadata Analysis The whole system is developed using java programming language. To extract the metadata of an image, the -extractor metadata library is used. Metadata decompressor can extract metadata information of many different types of images. After extracting the metadata, the metadata text will be put into the metadata analysis module. Metadata analyzer is basically a tag search algorithm. If keywords like Photoshop, Gimp, Adobe, etc. is found in the text and then the likelihood of being tampered with is increased. The two separate variables that are maintained are called false and true. Each variable represents the weight of being a real or fake image Once the marker has been taken, it is analyzed and the corresponding correction is increased by a certain predefined weight. The following table represents the keywords and the corresponding weight gain. After processing all the tags, the final values of the fake and actual variables are included in the Error rate analysis with the help of the ImageJ library starts saving the image at 100% quality. The same image is converted into 90% image quality using ImageJ.The difference between the two is found in the method differences.The resulting image is the required ELA image of the input image.This image is saved as a counter image and sent to the neural network for processing.

B. Machine Learning Machine learning is implemented using the Neuroph [4] library for java. Neuroph is chosen because it is easy and easy to implement a neural network. We implemented a multi-layer perceptron network using momentum backpropagation learning rules. The structure of the neural network is shown in the table. 2. Multilayer perceptron neural networks are used in one input layer, three hidden layers and one output layer. When an image is selected for evaluation, it is converted to an ELA representation from the compression and error level analysis stages. 100%, 90% Image ELA Used to build images Once the ELA is calculated, the image is preprocessed to be converted to a width and height of 100x100px. After preprocessing, the images are serialized into an array. The array contains 30,000 integer values that represent 10,000 pixels. It has red, green, and blue components, and 10,000 pixels has a value of 30,000. During training, the array is provided as input to the multilayer perceptron network and the output neurons are also set up. The MLP is a fully connected neural network. It has 2 output neurons. The first neuron represents the fake and the second neuron represents the real image. If the given image is a fake 1, the fake neuron is set to 1 and the real is set to 0, otherwise the fake is set to 0 and the real is set to 1. Adjust neuronal connection weights using the momentum backpropagation learning rule. Supervised learning rule to minimize the error function. The

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 2, February 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0902029 |

selected learning rates and momentum are shown in Table 3 along with the achieved efficiencies. During testing, an array of images is fed to the input neuron and the value of the output neuron is taken. We have used a sigmoid activation function.

IV. EXPERIMENTAL RESULT

Metadata analysis showed promising results in images unallocated. In addition, it became completely inaccurate when images were used metadata provided. all given The neural network will be 30,000. From the database we used 4000 real and non-training images.The remaining images were used to test the neural network.Table 3 shows the various neural network configurations and neural network efficiency. You can get the best results by setting the learning rate to 0.2 and the momentum to 0.

V. CONCLUSIONS

The neural network is successfully trained using error level analysis with 4000 fake and 4000 real images.The trained neural network was able to detect the image as unrealistic or realistic with a maximum success rate of 83% .Using this application on mobile platforms as a means of false evidence in digital verification, court evidence testing etc. By combining the results of metadata analysis (40%) and neural network output (60%) A reliable image detection system is developed and tested.

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/ CASIA v2.0 CASIA V2.0 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] http://neuroph.sourceforge.net/ Neuroph Framework Neuroph is lightweight Java neural network framework to develop common neural network architectures. It contains well designed, open source Java library with small number of basic classes which correspond to basic NN concepts.
[5]https://github.com/drewnoakes/metadata-extractorMetadata-extractor is a straightforward Java library for reading metadata from image files.
[6]https://www.github.com/afsalashyana/FakeImageDetectionGitHub repositor for fake image detector desktop application written in javafx.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor:
7.488

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details