International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 5 Issue 1, November-December 2020 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

# Imagically - Image Forensic Tool

**Raj Saundatikar[1], Dr. Lakshmi J. V. N[2]**

[1]Master of Computer Application, [2]Assistant Professor,
[1,2]Department of MCA, Jain Deemed-to-be University, Bangalore, Karnataka, India

**ABSTRACT**

With the increasing popularity of social media such as Instagram, Facebook and twitter, fake news with fake pictures which look real has become a huge problem. Digital Photos are more integral to communication than ever before, but the wide availability of easy image editing and manipulation tools at the disposable of anyone with a computer, or a smartphone, it makes them a risky proposition when trust is important. The point of this undertaking is to give a thorough idea of the best in class in the territory of Image forensics investigation. In this era to effectively change the data of a picture without leaving any undeniable which achieves viably and precisely the Image tampering discovery task. This strategy has been intended to recognize to decide if the Image is original or altered, without the knowledge on any data about the picture under examination. The tool work by identifying the certain pattern and metadata and other function giving a better result in forensically extraction.

***Keywords:*** *Error Level Analysis, MIME (Multipurpose Internet Mail Extension), Steganography*

***How to cite this paper:*** Raj Saundatikar | Dr. Lakshmi J. V. N "Imagically - Image Forensic Tool" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-5 | Issue-1, December 2020, pp.385-387, URL: www.ijtsrd.com/papers/ijtsrd35808.pdf

IJTSRD35808

## INTRODUCTION

The reliability of photos has a fundamental job in numerous territories, including: measurable examination, criminal examination, reconnaissance frameworks, knowledge administrations, clinical imaging, and news coverage. The specialty of making picture fakery has a long history. In any case, in the present advanced age, it is conceivable to handily change the data spoke to by a picture without leaving any undeniable hints of altering. In spite of this, no framework yet exists which achieves successfully and precisely the picture altering recognition task.

The advanced data insurgency and issues worried about sight and sound security have additionally created a few ways to deal with computerized crime scene investigation and altering identification. By and large, these methodologies could be partitioned into dynamic and latent visually impaired methodologies. The region of dynamic strategies just can be separated into the information concealing methodology (e.g., watermarks) and the advanced mark approach. We center around dazzle strategies, as they are viewed as another heading and rather than dynamic techniques, they work without any ensuring procedures and without utilizing any earlier data about the picture. To identify the hints of altering, dazzle techniques utilize the picture work and the way that falsifications can bring into the picture explicit distinguishable changes (e.g., measurable changes).

In this project, we will discuss about how the image forensics can be used in computer forensics, mobile forensics in chapters 3 and 4 .The tool is built using python programming and tool lets you get the information on the image, information like image meta data, image time stamp, mime type and file size and can help find error level, image hash, pixel hash and lets you compare image hash pixel by pixel. we will also provide the demo/screenshots of the project.

### Description of Research work

**Problem Statement:** Most of the forensic software in the market are paid and the size of the software is also big.

**Proposed Methodology:** This project is going to solve the above mentioned problem by making the tool free, open source, fast and lightweight.

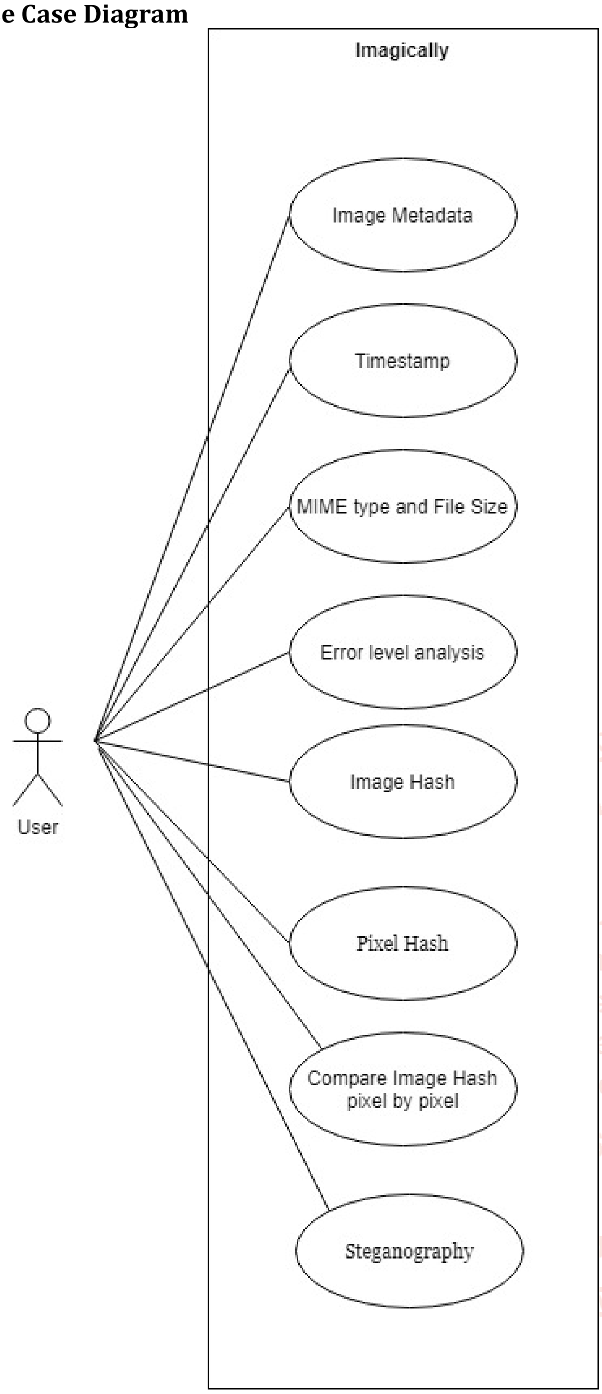
**Motivation:** The motivation of this project is to make an effective open source image forensic tool.

**Scope:** Anyone who want to find out whether the image is tampered or altered or need a source code for research work or add some additional new functionality to it. It can be also used for other functions mentioned in the introduction.

### Requirements

Hardware: 2GB RAM
Software: Windows OS, Linux, Python, Libraries such as NumPy, ImageHash, whratio, SciPy, Pillow.

**Use Case Diagram**



**Module-wise Description**

This tool is made up of following modules:

### 1. Image Metadata:

Image metadata is text information related to an image file that is embedded into the file or stored in a separate file that is associated with it. Image metadata includes details appropriate to the image itself as well as information about its production. Some metadata is generated automatically by the device capturing the image. Extra metadata may be added manually and edited with the help of software or general image editing software such as GIMP or Adobe Photoshop. Metadata can be added directly by some digital cameras. Image metadata can be very important for cataloging and contextualizing visual information. A large number of visual artists find the features useful in providing data about themselves and their images. With this module we can gather information related to the image such as image width, height, bits/pixel, compression rate, pixel format and aspect ratio.

### 2. Image Timestamp:

Timestamp is the data which is encoded that shows whether a certain event occurred, usually gives date and time of day, sometimes it can also be accurate to a small fraction of a second. The term Timestamp is derived from rubber stamps which were used in offices to stamp the current date, and sometimes time, in ink on paper documents, to record when the document was received. Files in computer contains timestamp that tells us when a certain file was last modified, and also digital cameras add timestamps to the pictures they take which records the date and time the picture was taken. With the help of timestamp, we can identify the time when the image was created and modified.

### 3. MIME type and File Size:

MIME stands for Multipurpose Internet Mail Extensions. It is a standard originally developed to extend e-mails to be able to support more formats like non-ASCII text and attachments in form of image, audio, video or executable files. The MIME Type is part of the header of the MIME and specifies the type of media contained in an e-mail. It is also referred to as media type or MIME content type. The usage of MIME is not limited only to e-mail though. It is used on the internet to determine the type of a file. It works similarly to the file extension on a computer. Web servers and browsers contain a list of MIME Types that help them to identify and thus interpret all kinds of files, independent of the operating system and hardware used by the user. Itis structured in a certain way, which contains a type and a subtype. For example, the MIME Type of a .png image is image/png with “image” being the type and “png” being the subtype. A slash (/) is used to separate type from subtype.

The size of a file is the amount of space it takes up on your hard drive. The File Size is measured in bytes as opposed to bits. One byte consists of 8 bits. The value is thus given out in bytes, kilobytes (KB), megabytes (MB), gigabytes (GB) and so on. One megabyte is 1024 kilobytes, one kilobyte is 1024 bytes, etc. The display of the File Size might wary by a few bytes depending on the hard drive it is saved on. Some files might take up a little more space on an external drive like a USB stick as compared to on the hard drive. Usually, it is displayed in the most economical way. Meaning that, in many cases, a long number like 55366 bytes will be shown as 54 KB. The File Size plays a crucial role when it comes to calculating and saving space on your hard drive, sending files or uploading them to the internet.

### 4. Error Level Analysis:

Error Level Analysis is an examination of compression artifacts in computerized data with lossy compression such as JPEG. Lossy compression is usually uniformly applied to a set of data, such as an image, resulting in a uniform level of compression artifacts. Alternatively, the data may comprise of parts with different levels of compression artifacts. This distinction may arise from the different parts having been repeatedly subjected to the same lossy compression a different number of times, or the different parts having been subjected to different kinds of lossy compression. A difference in the level of compression artifacts in different parts of the data may therefore shows that the data have been edited.

In case of JPEG, even a composite with parts subjected to similar compressions will have a distinction in the compression artifacts. So, to make the typically weak compression artifacts more clearly visible, the data to be examined is subjected to an additional round of lossy compression, this time at a known uniform level, and the result is subtracted from the original data under the investigation. The resulting distinct image is then inspected manually for any change in the level of compression artifacts. In 2007, N. Krawetz denoted this method "error level analysis". Digital data formats like JPEG sometimes include metadata describing the specific lossy compression used. If such data, in the noticed compression artifacts differ from those expected from the given metadata description, then the metadata may not describe the actual compressed data, and thus indicate that the data have been edited.

By nature, data with no lossy compression, such as a PNG image, cannot be subjected to error level analysis. Since editing could have been performed on data without lossy compression with lossy compression applied uniformly to the altered, composite data, the presence of a uniform level of compression artifacts doesn't rule out altering of the data. Furthermore, any non-uniform compression artifacts in a composite may be removed by subjecting the composite to repeated, uniform lossy compression. Also, if the image color space is reduced to 256 colors or less, for example, by conversion to GIF, then error level analysis will generate incompetent results. More notable, the actual description of the level of compression artifacts in a given segment of the data is subjective, and the assurance of whether altering has occurred is therefore not vigorous.

### 5. Image Hash:

Perceptual hash algorithms describe a class of comparable hash functions. Features in the image are used to generate a different fingerprint, and these fingerprints are comparable. Perceptual hashes are a distinct concept compared to cryptographic hash functions like MD5 and SHA1. With cryptographic hashes, the values of hash are random. Data which is used to generate the hash acts like a random seed, so the similar data will generate the same result, but different data will create distinct results. Comparing two SHA1 hash values only tells you two things which is whether the hashes are different, then the data is different. And if the hashes are same, then the data will be same. In contrast, perceptual hashes can be compared giving you a sense of similarity between the two data sets. In this module we can generate hash value of an image which is unique for each file or image. Image hash can be of SHA512, SHA256, SHA1 and MD5.

### 6. Compare Image hash pixel by pixel:

The general idea is very simple pixel-by-pixel comparison. The comparison engine gets the color of pixels that have the same coordinates within the image and compares this color. If the color of each pixel of both images coincides, tool considers the two images to be identical.

### 7. Steganography :

Steganography is a method of concealing secretive data within an ordinary looking file or message in order to avoid detection. The hidden data can then be extracted at its destination. The use of steganography can be used with encryption as an extra step for hiding or protecting data. Steganography can be used to hide any type of digital data, including text, image, video or audio content; the data to be hidden can be hidden inside almost any other type of digital file. The file content to be concealed through steganography is called *hidden text* is most often encrypted before being included into the ordinary-seeming *cover text* file or data stream. If unencrypted, the hidden file is processed in some way in order to increase the difficulty of detecting the secret content.

Steganography is practiced by those wishing to communicate a secret message or code. There are also many legitimate uses for steganography, malware developers have also been found to use steganography to obscure the transmission of malicious code. Different forms of steganography have been used for centuries and include almost any technique for hiding a secret message in an otherwise harmless container. For example, using invisible ink to hide secretive messages in ordinary looking files; hiding documents recorded on microdot which can be small as 1 mm in diameter on or inside legitimate seeming correspondence; and even by using multiplayer gaming environments to share information.

## Future Scope

This tool is command line based and does not support GUI, this might be worked upon in the future. More features related to image forensics will be added in the future. There are various future scopes for the tool, firstly there is Image extraction and detection from drive automatically, secondly Scanning for hidden strings, data and implement more forensics option like mobile disk image extraction.

## Conclusion

In conclusion, I have learned that as technology continues to rapidly advance not only with computers but with mobile devices, the need for forensics also increases. The project is open source image forensic tool which will provide various image forensic features.

## References

[1] https://whatis.techtarget.com/definition/image-metadata

[2] https://en.wikipedia.org/wiki/Timestamp

[3] https://www.metadata2go.com/file-info/mime-type

[4] https://searchsecurity.techtarget.com/definition/steganography

[5] https://www.phash.org/