School of Science and Engineering

Capstone Design

EGR 4402

IDENTIFYING MANIPULATED JPEG IMAGES

Capstone Report

Zineb Guessous

Supervised by Dr. Naeem Nizar Sheikh

Spring 2019

IDENTIFYING MANIPULATED JPEG IMAGES

Capstone Report

Student Statement:

I have applied ethics to the design process and in the selection of the final proposed design. I have held the safety of the public to be paramount and have addressed this in the presented design wherever may be applicable.

Zineb Guessous

Approved by the supervisor

Dr. Naeem Nizar Sheikh

Acknowledgments

First, I would like to express my sincere gratitude to my supervisor Dr. Naeem Sheikh, for his constant help, advising, and constructive feedback throughout this project.

I am also particularly grateful to my family for their financial and moral support. Finally, I would like to thank my friends who made my college journey a very memorable one.

Contents

List of figures

Abstract

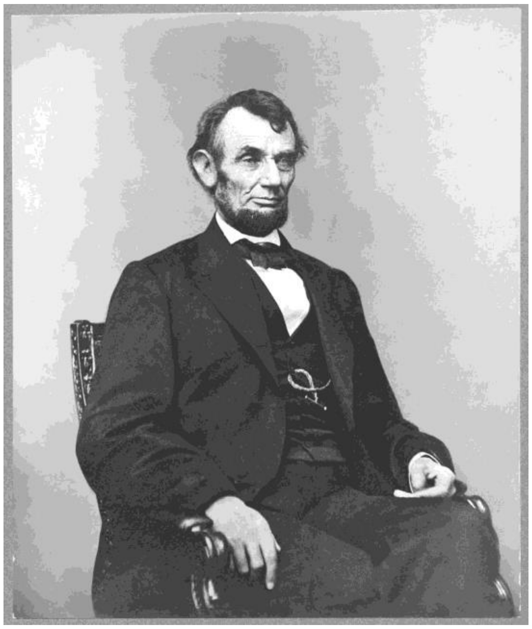
People are exposed to digital photographs more than ever before, as they are both creators and consumers of these images. Moreover, image editing software have no limits, and while it fosters creative artistic expression, it also allows for truth-deceiving content to be created. Error Level Analysis is an image forensic method which seeks to reveal manipulations in JPEG images. In this paper, we implement ELA as a fundamental technique of identifying manipulated images, and we discuss its performance and limitations. As we find that the method has a poor performance due to its noisy error level image results, we implement an improved version of ELA that handles that issue, and we evaluate its performance with regards to manipulation localization, based on metrics such as Precision, F1-score, and Matthews Correlation Coefficient (MCC). This method achieved a maximum of 81% for precision, 51% for the F1-score, and 54% for the MCC. Compared with state-of-the-art methods, this approach ranks among the techniques with the highest F1-scores, right below machine learning-based methods – the best of which reached an F1-score of 61% using fully convolutional networks (FCN). Yet, the method discussed in this paper outperforms all other techniques when it comes to the MCC score, even the machine learning-based ones – the best of which is the FCN-based technique with an MCC score of 48%.

1. Introduction

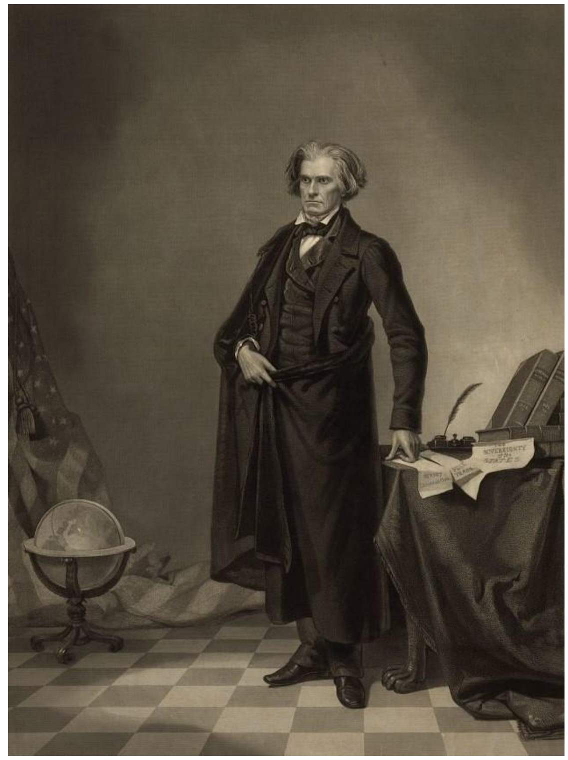
In the early 1860s, after the assassination of the U.S president Abraham Lincoln, a portrait painter created what was long thought to be a photo portrait of the late president, when in reality it was a composite print. [1] The painter combined Lincoln's head from a photograph, with the body of John Calhoun – a former U.S vice president – from a hand-drawn engraving, as shown in Figure 1. The words on the desk papers of the print were also modified, as Calhoun's political views were not exactly in tune with those of Lincoln's. Reasons as to why Lincoln's photo was faked are not entirely factual. It has been speculated that the main reason is the lack of heroic photos to commemorate the late president. [2] Even though this kind of image manipulation, back then, required special skillset, significant time, and utmost precision, it occurred quite frequently. Several political leaders have had painters or photographers doctor images for them, to mold the truth as they pleased and in whichever way bolstered their fame. [1]

While the means to image manipulation were scarce back then, they have become a lot more available in the late 80s to the early 90s, especially after the launch of the image processing tool Adobe Photoshop. The latter, further enhanced, is nowadays a powerful imaging and graphic design software that knows no bounds. With such tools, manipulating images becomes a fairly simple and quick task, that anyone can learn to do in a matter of hours. This creates far more opportunities for mischievous image manipulation, further spread through social media platforms. The issue lies in the fact that the human eye simply cannot distinguish an authentic image from a manipulated one. Therefore, to this day, many manipulated images go unnoticed before the eyes of digital media's heavy consumers, which are none but the billions of people on this planet. In this project, we explore the topic of identifying manipulations in digital photographs.

This project falls into the recent discipline of digital image forensics, which aims at analyzing and assessing the authenticity of a digital image. In this context, image processing and analysis techniques are used to reveal manipulation traces, if any. [3] Needless to say, each of these techniques tackles this topic from a different perspective, therefore, while they prove to be effective for certain situations, they might not be as efficient for others. With that in mind, we have chosen to approach this topic by implementing the Error Level Analysis technique improved with the block-based Discrete Cosine Transform, which seeks to expose splicing manipulations specifically in digital photographs saved under the JPEG format.



(a) Lincoln’s photograph portrait



(b) Calhoun’s hand-drawn engraving

(c) Resulting print of superimposing Lincoln’s head from the photo (a) into the engraving (b)

Figure 1: Lincoln's composite print

2. Background

2.1. JPEG

2.1.1. JPEG File Format

JPEG stands for Joint Photographic Experts Group, and is one of the most used image file formats, as it delivers a remarkable visual quality while reducing the file size as much as possible. [4] The adequate quality-to-size ratio of JPEG files makes it so that they are easily and quickly transferrable over websites and web applications. Hence why JPEG was able to reach unprecedented popularity.

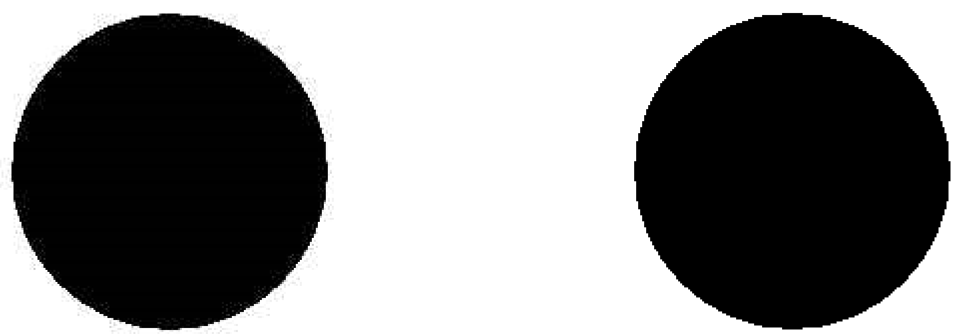
When an image is saved or shared as any file format, it is compressed following a certain method. The nature and steps of the compression method used by JPEG are key to achieving its satisfactory quality-to-size ratio. Compression methods can be either 'lossy' or 'lossless'.

2.1.2. 'Lossy' vs 'Lossless' Compression

A compression method is defined as 'lossy' when it erases an amount of pixels from the original image to allow for a significant file size reduction. [6] The lower the compression quality is, the more pixels are lost, and the smaller the file size gets. The pixels lost in the process of compression cannot be restored into the compressed image. Technically there is a difference of quality between the compressed image and the original one, but visually, that difference is hardly ever perceptible, unless the image has been compressed multiple times using low compression qualities. An example of a lossy compression is JPEG's compression method.

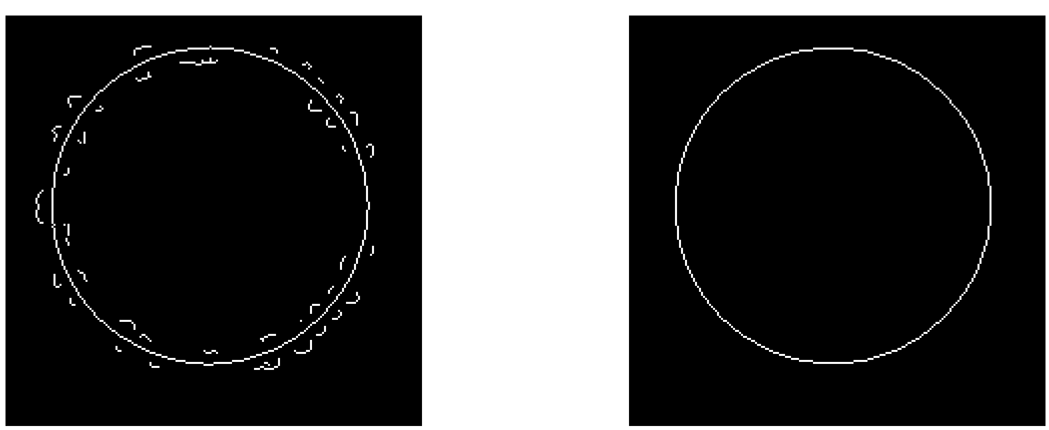
On the other hand, a compression is labeled as 'lossless' when it removes only unnecessary auto-generated data, such as metadata which represents information about the type of device and image editor used, in order to keep the quality intact while slightly reducing the file size. [6] Since no pixels are lost in the process of compression, it is possible to restore the uncompressed image. An example of a lossless compression is PNG's compression method.

Figure 2 below shows the difference between saving an image using each compression method. The 'circle' image is saved under JPEG and PNG in Figure 2 (a) and Figure 2 (b) respectively. Visually, the two images are similar, but statistically they are quite different. This shows when they are processed by an edge detector, as compression artifacts appear around the edge of the JPEG image in Figure 2 (c). Such compression artifacts would translate into blocky pixels that would be visible in the image when it has been saved under a low compression quality.



(a) 'circle.jpg' image

(b) 'circle.png' image



(c) 'circle.jpg' processed by an edge detector

(d) 'circle.png' processed by an edge detector

Figure 2: Difference between lossy and lossless compression

2.2. Image Manipulation

Image manipulation is a broad term that refers to any editing subjected to an image, or, in the context of this project, digital photographs. From slight corrections to significant alterations, there is a wide spectrum of manipulations that usually fall within the following categories: retouching, cloning, and splicing. The focus of this project is image splicing.

2.2.1. Image Retouching

Retouching an image means globally adjusting its characteristics such as brightness, contrast, or saturation, as well as locally correcting certain parts of the image such as aligning an uneven structure in a scenery, or smoothening a model's skin. [7] Such modifications are mostly used by photo studios, professional photographers, and magazines.

Opinions vary widely on whether retouching is considered as a harmful manipulation. The common ground is that retouching could be acceptable as long as it does not lead to conveying a false truth. For instance, taking a picture of the Leaning Tower of Pisa, changing its alignment, then passing it off as if the tower miraculously changed would be regarded as truth-deceiving. On the other hand, capturing a set of flowers, one of which moves at the moment the picture is taken, then retouching the photo to fix it would be tolerable.

An example of mischievous retouching is shown in Figure 3 below. A photo of the U.S president George Bush holding a book upside-down had gone viral over the internet in 2002. Later, it turned out it had been retouched; the book in the image had in fact been overturned. In this case, the purpose of retouching the image was to create political propaganda. Conversely, an example of moderate retouching is shown in Figure 4 below. It is a picture of a road on the country-side. Retouching is performed to fix the sky's overexposure – more brightness than it should be – and correct the colors overall.

Figure 3: President Bush Upside-Down Book (left: original, right: retouched)

Figure 4: Overexposure and color corrections (left: original, right: retouched)

2.2.2. Image Cloning

Cloning, also called copy-move, means duplicating a component of an image within the image itself. [7] Such kind of manipulation is usually used to heighten a part of an image, or multiply some objects within that image. Unlike retouching, there is no debate over copy-move manipulations; the intent behind it is clearly malicious. Below, Figure 5 is an example of a cloning manipulation. It is a 2008 photo of an Iranian missile test that was used in the front page of many newspapers such as NY Times and Chicago Tribune. The next day, it had been reported as a digitally altered photo by all newspapers that had shared it the day before. The added missile is said to be a copy of the one in its left, and the added smoke is cloned from the one on its right. [8]

Figure 5: Iranian Missile Test (left: original, right: manipulated)

2.2.3. Image Splicing

Splicing involves selecting an object from an image and inserting it into a different one to create a new composite image. [7] It is one of the most commonly used image manipulation schemes, and just like cloning, it is a harmful manipulation with the purpose of spreading false information. An example of that is shown below in Figure 6. In 2014, the photo of the Malaysia Airlines flight 17 crash had been posted on Twitter and, within a few hours, had over 3,600 retweets. Later that day, it had been reported as a manipulated screenshot from an episode of the TV series Lost. [9]

Figure 6: Malaysia Airlines Crash (left: original, right: manipulated)

3. Methodology

3.1. Error Level Analysis

Image manipulation detection techniques are divided into two major categories: active methods that depend on pre-embedded image information such as digital watermarks or signatures, and passive methods that rely on the variations of an image's statistical properties caused by the digital manipulations. Error Level Analysis is a passive pixel-based manipulation detection method that considers the underlying statistical inconsistencies, introduced at the pixel level of an image, in order to uncover manipulations in images.

3.1.1. ELA Overview

Dr. Neal Krawetz' image forensic method, Error Level Analysis, was first introduced in his paper, "A picture's worth…," as such:

> "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"." [10, Page 16]

Hence, in order to analyze the authenticity of an image using the ELA method, that image has to be resaved at a certain quality level, after which an error level image is generated by computing the absolute difference between the initial image and the resaved one. The ELA method is suitable for JPEG images, as the 'lossy' nature of the JPEG compression scheme is what allows the difference in quality levels between genuine pixels and foreign ones, in a manipulated image, to be highlighted.

Theoretically, the error level image could be expressed as such:

$$I_{ELA}(n_1, n_2) = |I_{initial}(n_1, n_2) - I_{resaved}(n_1, n_2)|$$

where n_1 and n_2 are row and column indices of the image, I_{ELA} is the error level image, $I_{initial}$ is the image to be analyzed, and $I_{resaved}$ is the resaved image. This is computed for each channel in RGB.

3.1.2. Literature Review

In this section, we review the previous work we found regarding manipulation detection using Error Level Analysis.

ELA with Laplacian Edge Detection [11] In this paper, a manipulation detection approach using Error Level Analysis and Laplacian Edge Detection is briefly discussed. The paper suggests that after the error level image is generated, Laplacian Edge detection is applied to further highlight the manipulated parts in comparison with the rest of the image, yet no implementation details are discussed. Photos used for testing were collected from Nikon and Canon cameras as raw images, then saved under JPEG; the splicing manipulations were carried out using Paint. Results of a couple of photos are presented, but no evaluation regarding the accuracy of those results is performed.

ELA with Semi-Automatic Wavelet Soft Thresholding [12] In this paper, another manipulation detection approach is put forward. The paper discusses that using the ELA method often introduces some noise in the error level image, especially if the image analyzed had already been compressed multiple times. This makes it harder to differentiate between areas in the image which are in fact manipulated, and the ones that are not. In order to reduce the amount of noise, the paper proposes to apply a semi-automatic wavelet thresholding acting as a lowpass filter on the error level image. The method is tested on a couple of spliced photos, and the results are presented. Yet, no evaluation as to the results' accuracy is given.

ELA with Discrete Cosine Transform [13] In this paper, a different approach is proposed. The paper addresses the noise issue that arises when using the ELA method on heavily compressed images, such as those shared over social media platforms. Therefore, it suggests an improved version of the ELA method which uses the block-based Discrete Cosine Transform to discard the initial image's noisy elements, consequently lowering the noise proportion in the error level image generated. The development of such a method is discussed, and a database of photos is used for testing. Results of a few photos are presented, yet again no evaluation was done about how accurate those results were.

3.2. Chosen Method

After careful investigation into the literature briefed in the previous section, we have selected to implement and evaluate the DCT-improved ELA method. This is because the paper [13] addressing this approach provided a detailed development, and out of all the papers reviewed, it had the most compelling results. The implementation part of this project involves an implementation of the ELA algorithm as described in the original paper [10], followed by an implementation of ELA with block-based DCT.

3.2.1. Discrete Cosine Transform

In the context of this project, the 2-dimensional Discrete Cosine Transform is a mathematical function that represents the image data as a sum of cosine functions at different frequencies, in order to transform the digital image from the spatial domain into the frequency domain. [13] This is done by splitting the image into N by N pixel blocks, each of which is reconstructed into an N by N frequency block as such:

$$B_{frequency}(u,v) = \alpha(u).\alpha(v)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} B_{pixel}(x,y).\cos\left[\frac{(2x+1).u.\pi}{2N}\right].\cos\left[\frac{(2y+1).v.\pi}{2N}\right]$$

where

$$\alpha(i) = \begin{cases} \sqrt{\frac{1}{N}} & if\ i = 0 \\ \sqrt{\frac{2}{N}} & otherwise \end{cases}$$

$B_{frequency}(u,v)$ is the u,v^{th} element of each frequency block, N is the size of the block, and $B_{pixel}(x,y)$ is the x,y^{th} element of each pixel block representing the image. This is computed for each channel in RGB.

An N by N frequency block is shown in Figure 7 below. The top left element is called the Direct Current (DC) coefficient, and all the other elements are called Alternating Current (AC) coefficients. We can see below the frequency distribution over the N by N block; the DC coefficient along with the first few closest coefficients represent the low frequency coefficients, which have the most contribution to the original pixel block. In other words, if we take only the top left coefficients and discard the rest, we would still be able to reconstruct the original N by N pixel block corresponding to that N by N frequency block.

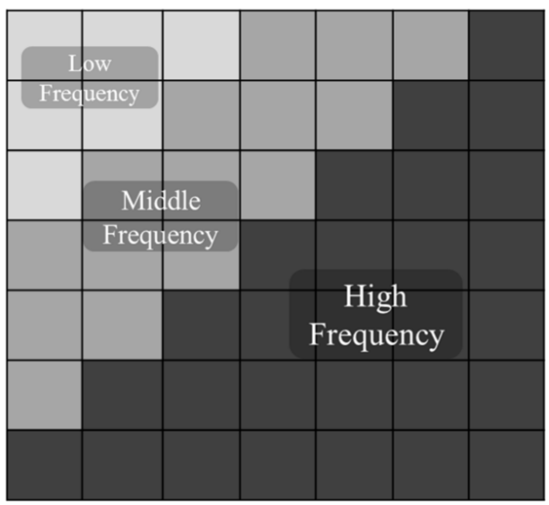


Figure 7: Frequency Distribution in a DCT Block

3.2.2. ELA with DCT

The idea behind using the 2D-DCT function in the ELA algorithm is so that we would be able to generate an error level image with as little amount of noise as possible. Simply put, the block-based 2D-DCT function will be applied on the image to be analyzed and the resaved one, after which the absolute pixel by pixel difference will be computed to generate the error level image. Finally, the inverse 2D-DCT will be applied to transform the error level image back to the spatial domain. The theoretical expression for the error level image is then the following:

$$I_{ELA}(n_1, n_2) = DCT_{inverse}\left(\left|DCT_{forward}\left(I_{initial}(n_1, n_2)\right) - DCT_{forward}\left(I_{resaved}(n_1, n_2)\right)\right|\right)$$

3.2.3. Technology Enablers

The work presented in this paper has been written in Python using Jupyter Notebook as a development environment, for it allows to have inline code and visualizations, which is useful given the nature of this project. We have used the Pillow library, formerly known as PIL (Python Imaging Library), for image-related functions, as it provides powerful image processing capabilities. We have also used Scipy library and more specifically the fftpack module that provide a 1-dimensional DCT implementation, which we used for our implementation of the block-based 2-dimensional DCT.

3.3. Experimental Evaluation

Evaluation is a crucial part of every project, as it provides valuable insights and meaning to the results. In our simulations, we evaluate the performance of our work by running it through a dataset of spliced images. Then, we use evaluation measures to compare the results and draw in well-founded conclusions.

3.3.1. Dataset and Experimental Setup

For testing purposes, we have used a set of spliced images that we created from authentic images found in the CASIA dataset. [14] The latter includes images of different sizes and a wide range of categories, such as animals, architecture, and nature. We have carried out realistic and challenging splicing manipulations using Adobe Photoshop, and we have created ground truth masks for each spliced image in the set. Our dataset consists of 25 images in total, including 15 spliced images and 10 authentic images. Error level images have been generated, stored, and evaluated by means of pixel classification, which assesses how accurate our implementations can locate the manipulated area of a tampered image.

3.3.2. Pixel Classification

As its name implies, pixel classification involves classifying the pixels within the error level image into genuine and foreign pixels. This is done by comparing the output error level images to the ground truth masks. Each pixel can be either a true positive (TP), a true negative (TN), a false positive (FP), or a false negative (FN). A true positive is a foreign pixel that showed up as such in the error level image. Likewise, a true negative is a genuine pixel that showed up as such in the error level image. On the other hand, a false positive is a genuine pixel that showed up as foreign in the error level image. Similarly, a false negative is a foreign pixel that showed up as genuine in the error level image.

The evaluation measures explained below, namely the F1-score and the Matthew's Correlation Coefficient, are calculated in order to evaluate the ability of the two methods implemented to correctly classify pixels within an image.

F1-score This is a harmonic mean of **precision** and **recall**. Precision is the number of pixels correctly identified as foreign out of all the pixels classified as foreign. While, recall is the number of pixels correctly identified as foreign out of all the pixels correctly classified.

- Precision = TP / (TP + FP)
- Recall = TP / (TP + FN)
- F1 = 2 * Precision * Recall / (Precision + Recall)

Matthew's Correlation Coefficient MCC is a correlation coefficient between predicted labels and actual labels. It is a measure that includes all quadrants of confusion matrix, which is why it is considered as a balanced measure. The formula is the following:

$$MCC = \frac{TP * TN \ - \ FP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

4. Implementation

4.1. First Algorithm: ELA

We have implemented the original ELA method previously described, it is summed up in the diagram represented in Figure 8 below. The algorithm has been defined as a function which takes as input the path of the image to be analyzed as well as the quality compression at which the image will be resaved, and outputs the error level image.

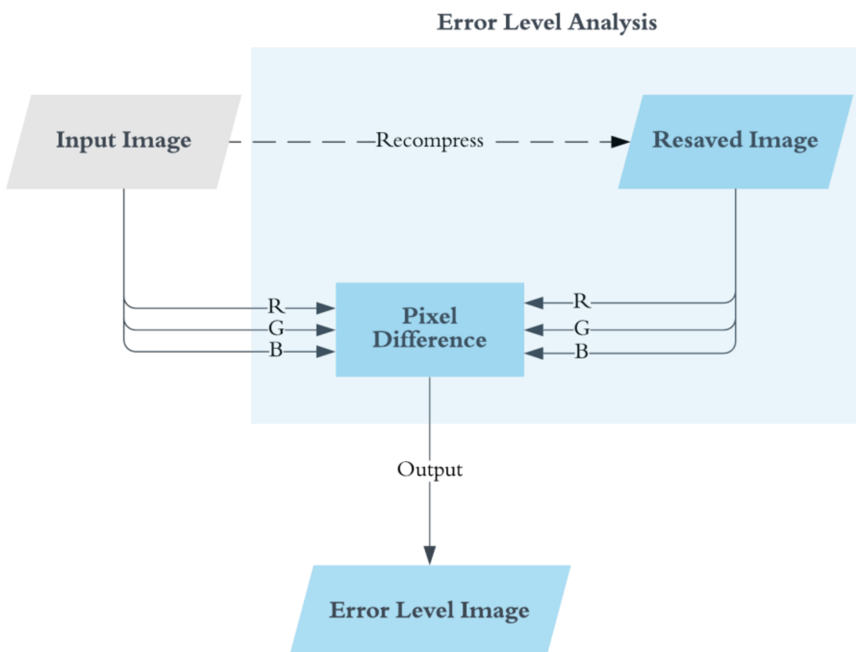


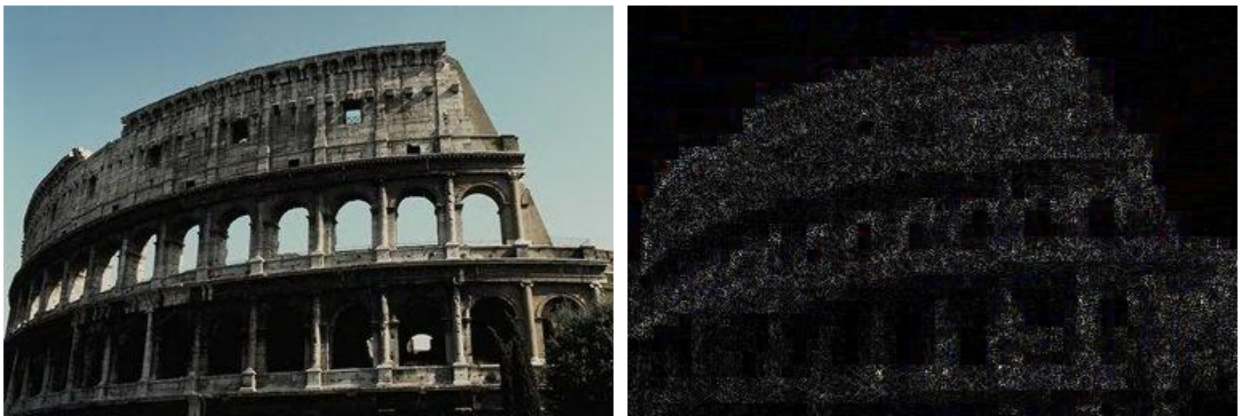
Figure 8: ELA Algorithm

The ELA function was later used in our testing script, which takes the path of the directory containing our dataset, and iterates over the images to generate error level images. Concerning the quality for resaving the input image, we have observed experimentally that different manipulations are better highlighted at different compression qualities. Yet, we have noticed that using a compression quality below 70% always introduces noise artifacts that would mislead us in the analysis of the error level image. Hence, for each image in our dataset, we have generated corresponding error level images at compression qualities ranging from 70% to 95% with an increment of 5% each time.

4.1.1. Results

In the following, we present the best and worst results, for some of the authentic and manipulated images in our dataset, regarding the error level images generated using our implementation of the original ELA algorithm.

Authentic Images Results



(a) Authentic Image (b) Error Level Image at a resave quality of 70%

Figure 9: Authentic Image - ELA Best Result

(a) Authentic Image (b) Error Level Image at a resave quality of 90%

Figure 10: Authentic Image - ELA Worst Result

The best and worst results for authentic images are shown in Figure 9 and Figure 10 above, respectively. Figure 9 is considered as a best-case example because even though an amount of change is highlighted in the error level image, it is distributed uniformly. That amount of change is actually related to the noise issue addressed in the literature review. Such noise is usually present around high frequency areas of an image such as textures or grass, which explains why in our example, the noise is showing around the building and the trees but not the sky. On the other hand, in Figure 10, the noise is more problematic as there is an abrupt change of noise distribution specifically around the center of the image. Such a difference of noise would be more likely to mislead us into thinking that the image is manipulated, and that the manipulation occurred in the area where the noise changed. Hence why this was considered as a worst case example. Out of the other 8 authentic images (excluding the examples shown) in our dataset, 7 error level images behaved similarly to the best-case example, while 1 of them behaved more like the worst-case example.

Manipulated Images Results

(a) Manipulated Image

(b) Error Level Image at a resave quality of 80%

(c) Ground Truth Mask

Figure 11: Manipulated Image - ELA Best Result - Example 1

(a) Manipulated Image

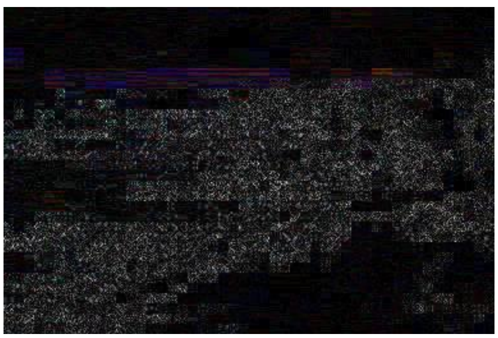
(b) Error Level Image at a resave quality of 75%

(c) Ground Truth Mask

Figure 12: Manipulated Image - ELA Best Result – Example 2

The best case results for manipulated images are shown in Figure 11 and Figure 12 above. These are considered as best case examples because the most highlighted areas in the error level images, are in fact the manipulated areas as per their corresponding ground truth masks. Moreover, the amount of noise in the error level images is relatively minor. While, Figure 13 and Figure 14 below show worst case results in which a portion of the manipulated areas of the images are highlighted, and the surrounding noise is much more intensified. Out of the other 11 spliced images (excluding the examples shown) in the dataset, 4 error level images behaved like the best case examples, while the 3 remaining behaved like the worst case examples.

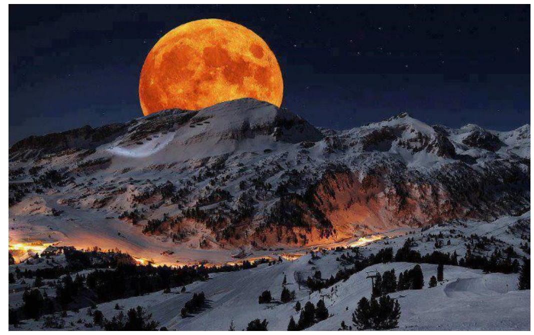
(a) Manipulated Image



Error Level Image at a resave quality of 70%

(c) Ground Truth Mask

Figure 13: Manipulated Image - ELA Worst Result - Example 1



(a) Manipulated Image

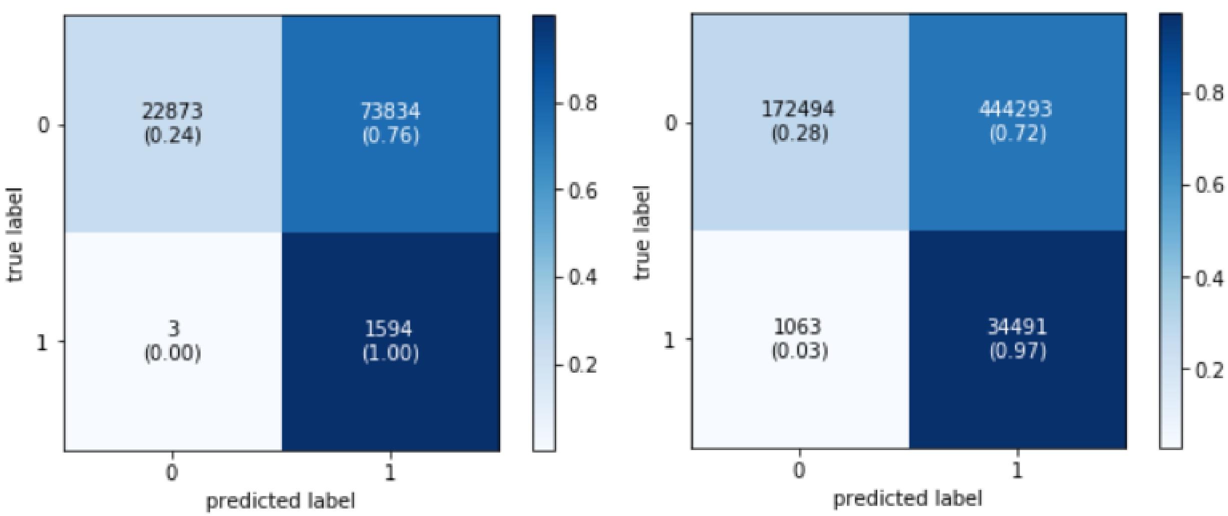
Error Level Image at a resave quality of 85%

(c) Ground Truth Mask

Figure 14: Manipulated Image - ELA Worst Result - Example 2

4.1.2. Analysis

As expected from the literature review, error level images generated from the original ELA algorithm are very difficult to interpret, mainly because of the noise introduced, whether be it when analyzing authentic images or manipulated ones. Regarding authentic images, the large amount of noise does not make it any easier to identify authentic images from manipulated ones, and would most probably lead to inaccurate results, should the ELA method be used as a classification technique. As for manipulated images, the large amount of noise makes it difficult to pinpoint the location of the manipulation. In this case, pixel classification proves to be completely useless, as all the noisy pixels are classified as foreign, which leads to an extremely high rate of false positives. For instance, the confusion matrices in Figure 15 below show the number of true and false positives as well as true and false negatives for the error level image in Figure 11 and Figure 12 respectively. The high rate of false positives is due to the amount of noise in the error level images.



Confusion Matrix for Figure 11 (b)

Confusion Matrix for Figure 12 (b)

Figure 15: Confusion Matrix for ELA Best Results

Conclusion The original ELA method on its own does not effectively achieve the objectives of this project, yet it serves as a reference and as a basis for the algorithm discussed in the next section.

4.2. Second Algorithm: DCT-Improved ELA

In order to overcome the noise issue, we use the DCT to transform each of the input image, and the resaved image into the frequency domain. We do so by applying the 2D-DCT over every 8 by 8 blocks of an image. Then we apply a threshold function which keeps the necessary low frequency coefficients and discards the high frequency ones responsible for the noisy components. Error levels are then computed before the inverse 2D-DCT is applied in order to output the error level image. These steps are summed up in Figure 16 below. The DCT-improved ELA function is used in a similar script as the one used in the first algorithm, using the same range for compression qualities.

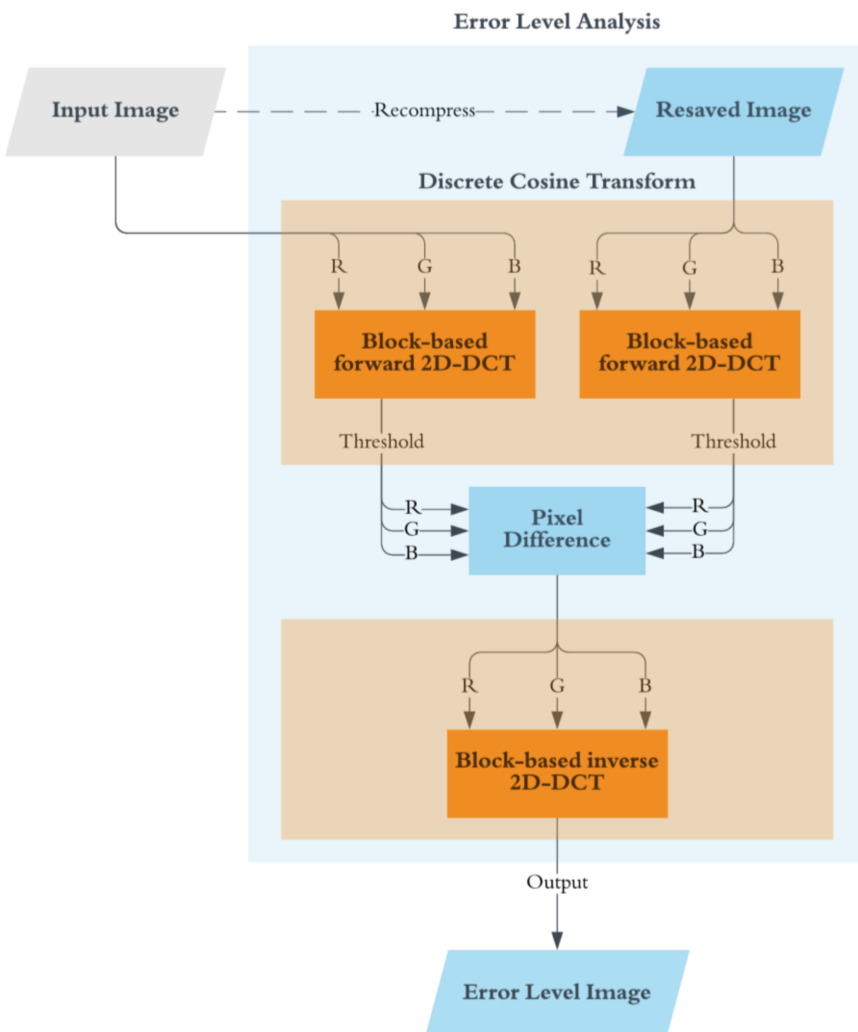


Figure 16: DCT-Improved ELA Algorithm

4.2.1. Results

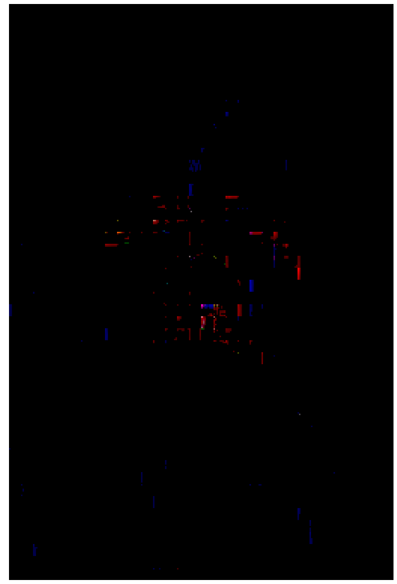
In the following, we present the best and the worst of the resulting error level images, generated from the DCT-improved ELA implementation, for a couple of authentic and manipulated images from our dataset.

Authentic Images Results

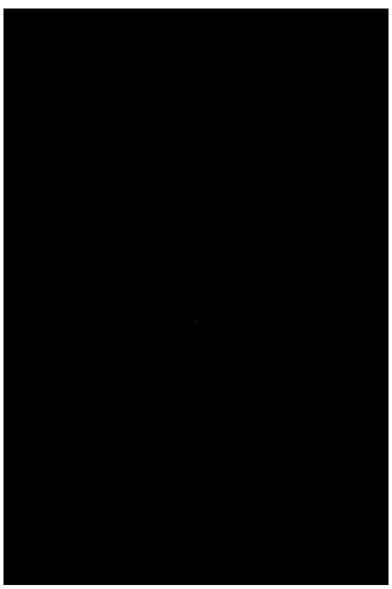
With regards to authentic images, we have noticed that for 4 images in the set, their corresponding error level images are pitch black starting 85% resave quality, as shown in Figure 17 below. For 4 others, it was starting 90% resave quality, as shown in Figure 18.

The remaining 2 images had some amount of noise at all the resave qualities as shown in Figure 19.

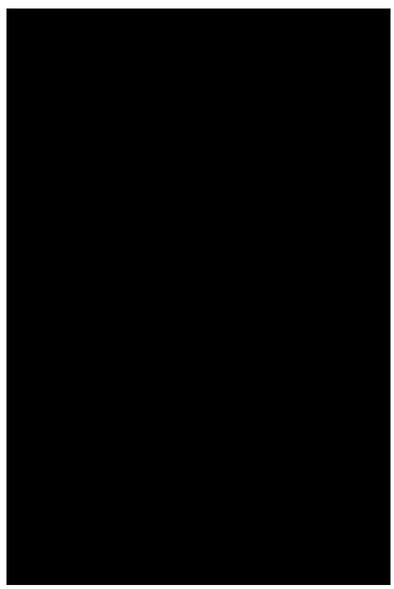
(a) Authentic Image



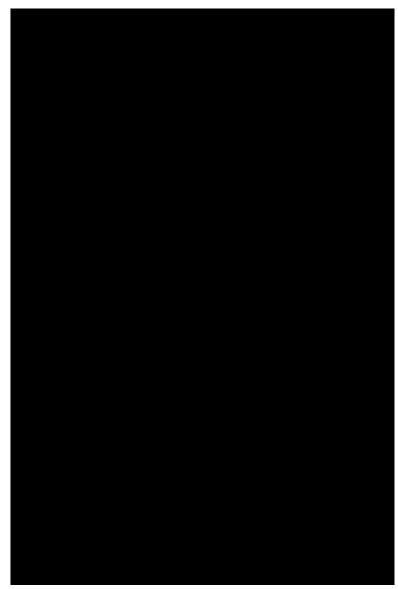
(b) Error Level Image at resave quality of 80%



(c) Error Level Image at resave quality of 85%



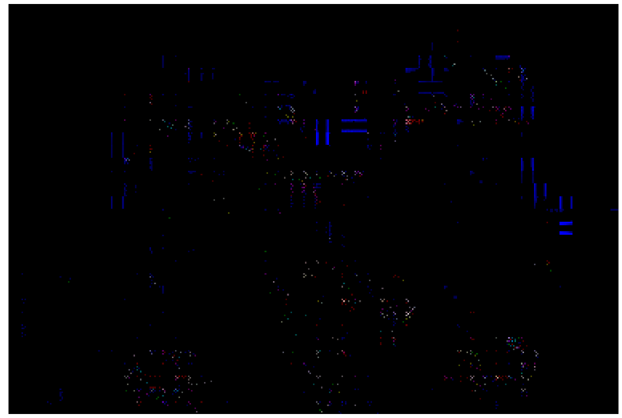
(d) Error Level Image at resave quality of 90%



(e) Error Level Image at resave quality of 95%

Figure 17: Authentic Image - DCT-Improved ELA - Example 1

(a) Authentic Image



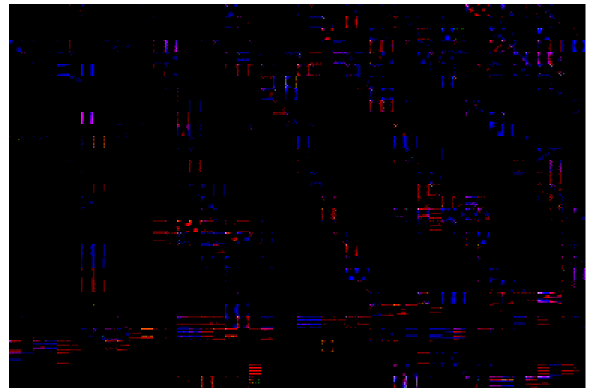
(b) Error Level Image at resave quality of 85%

(c) Error Level Image at resave quality of 90%

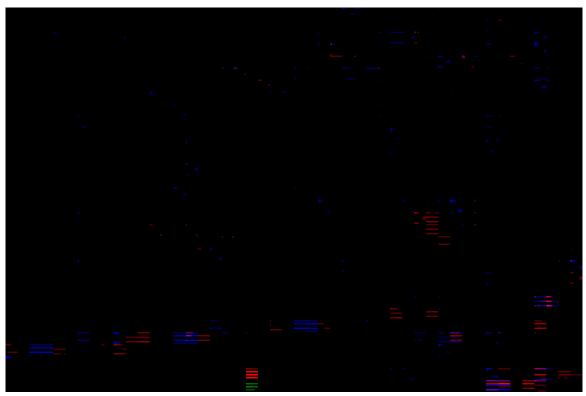
(d) Error Level Image at resave quality of 95%

Figure 18: Authentic Image - DCT-Improved ELA - Example 2

(a) Authentic Image



(b) Error Level Image at resave quality of 90%



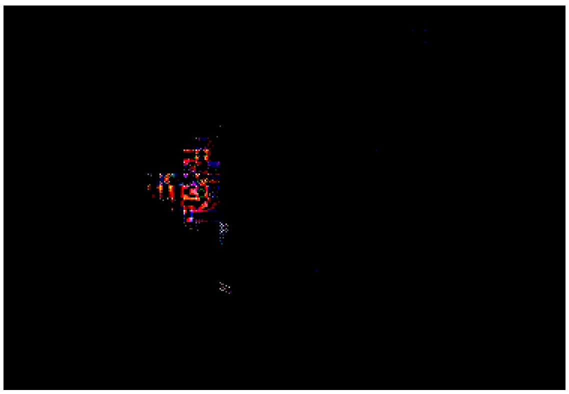
(b) Error Level Image at resave quality of 95%

Figure 19: Authentic Image - DCT-Improved ELA - Example 3

Manipulated Images Results

As we revisit the images from Figure 11 and Figure 12, we can see a remarkable difference in the error level images below in Figure 20 and Figure 21, as most of the noise has been significantly reduced. These are considered as best results. Moreover, manipulation locations are further highlighted and are almost perfectly matching the ground truth masks. As for the images from Figure 13 and Figure 14, their error level images, as shown below in Figure 22 and Figure 23, are still not ideal, but they are definitely less noisy, and they are able to give a general idea about where the manipulations have occurred. These are considered worst results. Out of the 15 spliced images in our dataset, 9 error level images had the correct location of the manipulations, while the remaining 4 were a bit ambiguous.

(a) Manipulated Image

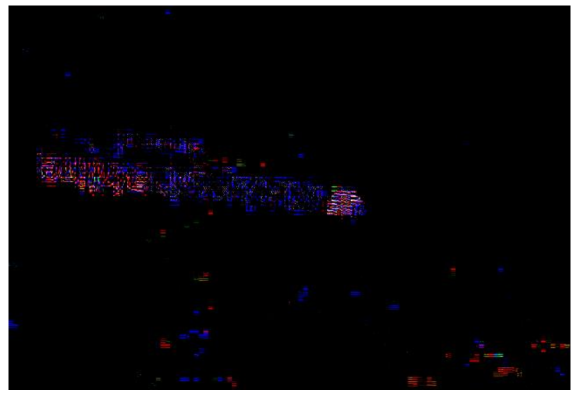


(b) Error Level Image at resave quality of 80%

(c) Ground Truth Mask

Figure 20: Manipulated Image - DCT-Improved ELA Best Result - Example 1

(a) Manipulated Image

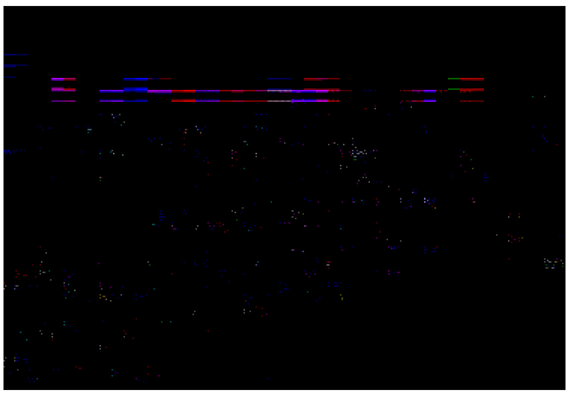


(b) Error Level Image at resave quality of 75%

(c) Ground Truth Mask

Figure 21: Manipulated Image - DCT-Improved ELA Best Result - Example 2

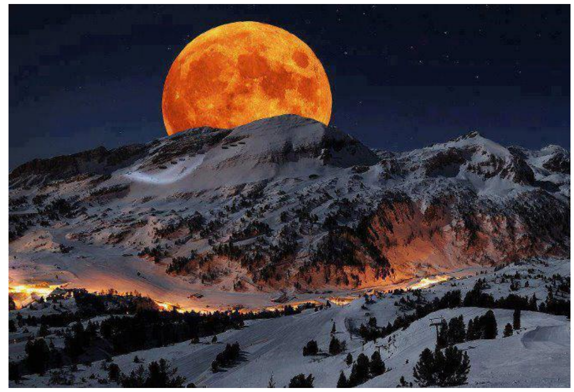
(a) Manipulated Image



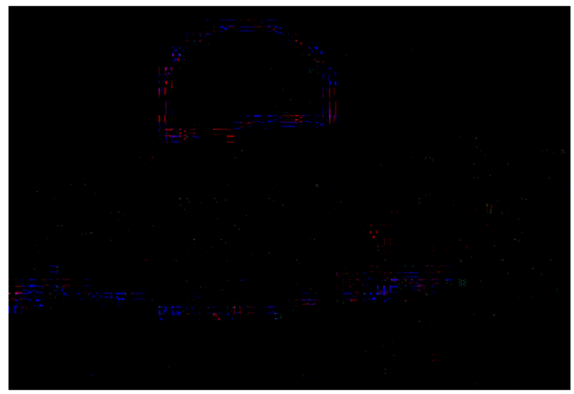
Error Level Image at a resave quality of 80%

(c) Ground Truth Mask

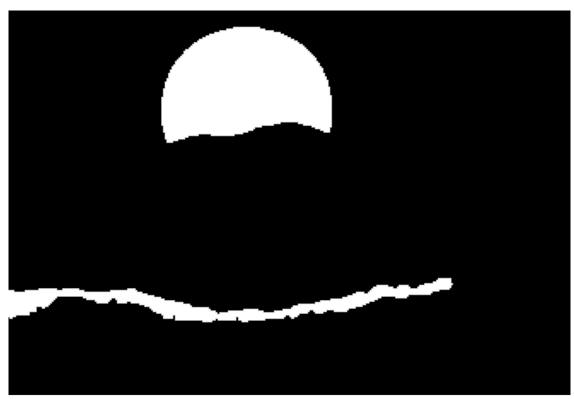
Figure 22: Manipulated Image - DCT-Improved ELA Worst Result - Example 1



(a) Manipulated Image



Error Level Image at a resave quality of 90%



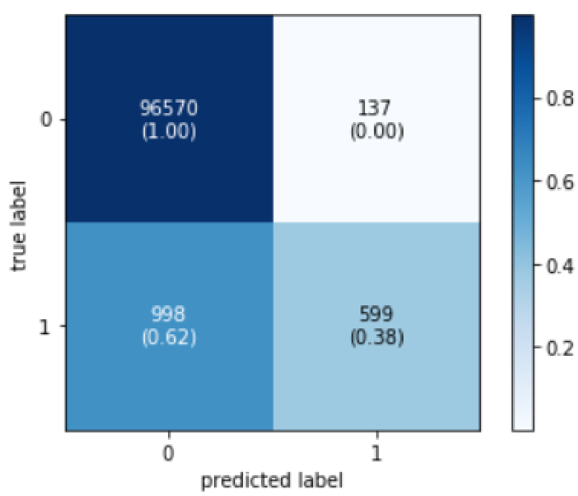
(c) Ground Truth Mask

Figure 23: Manipulated Image - DCT-Improved ELA Worst Result - Example 2

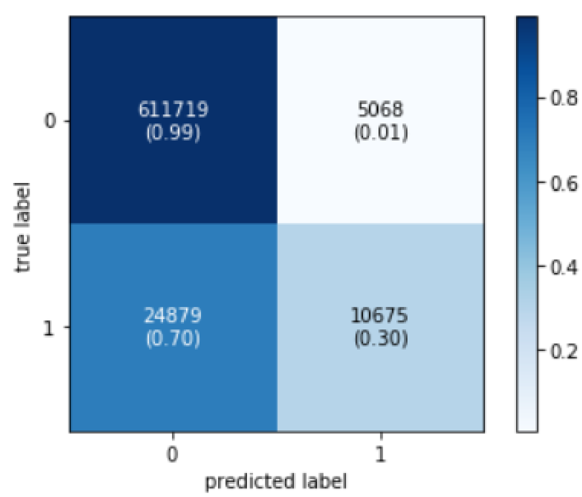
4.2.2. Analysis

The DCT-improved ELA method has proven its effectiveness with regards to identifying authentic images from manipulated ones. More importantly, it has efficiently tackled the noise issue associated with the original ELA method. We were able to evaluate the results of the second algorithm using pixel classification and evaluation metrics such as precision, the F1-score, and the MCC.

Below are the confusion matrices for error level images in Figure 20 and Figure 21, respectively. We can clearly see that the number of false positives has become completely insignificant; the highest it gets is 4%. We also notice a high number of false negatives, which represents foreign pixels that were not classified as such. But, from analyzing the error level images, we notice that it only becomes problematic when the rate of false negatives goes beyond 90%; otherwise, it means that at least the perimeter of the manipulation was highlighted. In fact, the error levels images, we had deemed as somewhat ambiguous, had a false negative rate of 90% or over, while the rest had between 62% and 82%.



Confusion Matrix for Figure 20 (b)



Confusion Matrix for Figure 21 (b)

Figure 24: Confusion Matrix for DCT-Improved ELA Best Results

From the confusion matrix of each error level image, we computed the precision, the F1-score, and the MCC. The results for the 9 error level images considered as best-case examples are as follows:

- Precision: ranged from 45% to 81%, with an average of 76%
- F1-score: ranged from 23% to 51%, with an average of 40%
- MCC: ranged from 28% to 54%, with an average of 42.5%

Conclusion The DCT-improved ELA method was able to achieve impressive qualitative and quantitative results; especially given that it does not rely on any machine learning algorithm.

5. STEEPLE Analysis

The STEEPLE analysis aims at identifying the implications and factors, or lack thereof, associated with a certain project from different aspects, more specifically the Social, Technological, Economic, Environmental, Political, Legal, and Ethical aspects. As such, it is substantial that each of the aforementioned elements is investigated meticulously.

As far as this project is concerned, the following implications were identified with regards to the STEEPLE analysis, most notably the Social, Technological, Political, Legal, and Ethical aspects.

5.1. Societal

Digital media is becoming an indispensable aspect of our daily lives. From the visual content we purposely scour for in websites and social media platforms, to the subliminal digital ads planted in between, we are hardly able to tell the difference between what is real and what is staged, unless explicitly mentioned – which is generally not the case. With this, media outlets and corporate companies are able to get off lightly with spreading fallacies that further disturb social ideations and standards. As such, this project aims to raise awareness about the issue of uncontrolled image manipulation, and attempts to provide a solution.

5.2. Technological

Recently, the field of Digital Image Forensics has been given a lot of attention from researchers in the scientific community. These efforts come in the context of reducing the risks associated with malicious use of, widely available, advanced image processing technologies. This project, built upon existing, although somewhat limited, research attempts to contribute to such efforts.

5.3. Economic

This project is not for profit, and so it is does not have any economic implications nor it is affected by economic factors.

5.4. Environmental

This project is not affected by environmental factors, as its deployment has little to no effect over natural resources.

5.5. Political

Image manipulation has been used extensively to create political propaganda, from the infamous manipulations, in the beginning of the 1900s, at the hands of political leaders such as Joseph Stalin, to the recent accounts of tampering, such as Osama Bin Laden's fake death photo in 2011, broadcasted by the Pakistani TV, and used in the front page of many renowned British newspapers. [5] While major media outlet, nowadays, enforce stricter rules on photojournalists regarding photo manipulation, a photographer's pledge is still the only source to the authenticity of photos. This project aims to establish a means to proving or denying the authenticity of photographs as accurately as possible.

5.6. Legal

There are few legal restraints that could possibly be applicable to the image manipulation issue, such as trademark law, false light, or defamation, all of which primarily pertain to the plaintiff's property, emotional well-being, or reputation, respectively. Yet, none of these laws are directly relevant to the question of the photograph's authenticity and truthfulness. The lack of such laws is due to the fact that there exist few to no reliable systems to authenticate images. Potentially, this project would contribute to the making of such systems.

5.7. Ethical

Current photo editing software provide a wide spectrum of possibilities for amateurs and professionals alike. Their intent and purpose behind the use of such tools largely impact how the photo manipulation issue unfolds. An ethical code should be established to draw a line between the kind of photo editing that is harmless, and the one that is not, following unified ethical and moral standards.

6. Final Remarks

6.1. Challenges and Limitations

The main challenges faced were related to finding a suitable splicing dataset with corresponding ground truth masks. I had to manually create spliced images along with their ground truth data, which is why our dataset was not of a significant size. A dataset larger than the one we used would provide better insights with regards to the accuracy of the implemented method.

6.2. Future Work

While looking into the Discrete Cosine Transform, I came across an interesting paper that discusses the use of DCT in a different manipulation detection setting. The paper discusses an approach that could be used to identify splicing as well as copy-move manipulation. First, the RGB image is converted to the Luminance & Chrominance channels, and the Chrominance channel is extracted and taken into consideration for further analysis, as that is where most of the manipulation traces are hidden. Local Binary Pattern (LBP) is then used to capture texture deviations, resulting from manipulations, in overlapping blocks of the image's chrominance channel. After that, the image's chrominance blocks are transformed into the frequency domain using the 2D-DCT for less computational power. Finally, the standard deviation is computed for all frequency blocks, which is used as feature that is fed to a Support Vector Machine for data classification. Even though this approach seemed to be more concerned with classification rather than locating manipulations, it could be used precisely for classification purposes, then the DCT-improved ELA method that we implemented would be focused solely on identifying the manipulation traces in tampered images.

6.3. Conclusion

As we, as a society, are moving towards an entirely digital age, we need to be aware of the repercussions associated with such a lifestyle. This project sheds light on a contemporary societal and ethical issue; that of image manipulation. We explore this topic by looking into state-of-the-art manipulation detection techniques, and we dive specifically into the Error Level Analysis method. We provided an implementation, and analysis of the original method and its improved version, followed by an evaluation of the results. We have concluded that Error Level Analysis as a standalone method faces a lot of limitations, but when combined with other image processing techniques, it could provide remarkable results.

7. References

[1] H. Farid, “Digital Doctoring: How to tell the real from the fake,” *Significance*, vol. 3, no. 4, pp. 162-166, Dec. 2006. [Online] Available: Royal Statistical Society, DOI: 10.1111/j.1740-9713.2006.00197.x

[2] Bronx Documentary Center, *United States Circa 1865*, Bronx Documentary Center Altered Images. [Online] Available at http://www.alteredimagesbdc.org/lincoln

[3] J.A. Redi, W. Taktak, and JL. Dugelay, “Digital Image Forensics: A Booklet for Beginners,” *Multimedia Tools Applications*, vol. 51, no. 1, pp. 133-162, Oct. 2010. [Online]. DOI: https://doi.org/10.1007/s11042-010-0620-1

[4] O. Shehata, “Unraveling the JPEG,” *Parametric Press*, issue 01: Science + Society, May 2019. [Online]. DOI: 10.5281/zenodo.2655041

[5] M. Davis, *Faking Photos: The Ethics of Photo Manipulation*, The Center for Digital Ethics and Policy. [Online]. Available: https://www.digitalethics.org/essays/faking-photos-ethics-photo-manipulation

[6] L.J. Karam, *The Essential Guide to Image Processing*, MA: Academic Press, 2009. [E-book]. Available: ScienceDirect, https://www.sciencedirect.com

[7] V.D. Shalini, and S.D. Suresh, “Detection of Image Forgery using LBP and DCT Techniques,” *IJSART*, vol. 2, no. 8, pp. 76-86, Aug. 2016. [Online]. Available: Academia, https://www.ripublication.com/ijaer17/ijaerv12n22_11.pdf

[8] M. Nizza, and P.J. Lyons, *In an Iranian Image, a Missile Too Many*, The New York Times, July 2008. [Online]. Available: https://thelede.blogs.nytimes.com/2008/07/10/in-an-iranian-image-a-missile-too-many/

[9] M. Novak, *This isn’t the Malaysia Airlines crash. It’s Photoshopped from Lost*, Gizmodo, July 2014. [Online]. Available: https://gizmodo.com/this-isnt-the-malaysia-airlines-crash-its-photoshopped-1606662221

[10] N. Krawetz, “A picture’s worth…,” *Hacker Factor Solutions*, 2007. [Online]. Available: http://www.hackerfactor.com/papers/bh-usa-07-krawetz-wp.pdf

[11] E. Ramadhani, “Photo Splicing Detection using Error Level Analysis and Laplacian Edge Detection Plugin on GIMP,” *Journal of Physics: Conference Series*, vol. 1193, conf. 1, 2019. [Online]. Available: IOP Science, DOI: 10.1088/1742-6596/1193/1/012013

[12] Daniel C. Jeronymo, Yuri C. Campbell Borges, and Leandro Coelho, “Image Forgery Detection by Semi-Automatic Wavelet Soft-Thresholding with Error Level Analysis,” *Expert Systems with Applications: An International Journal*, vol. 85, issue C, pp. 348-356 Nov. 2017. [Online]. Available: ACM Digital Library, DOI: 10.1016/j.eswa.2017.05.044

[13] A. Hilal, and S. Abou Chakra, “Development of The Error Level Analysis Forensic Tool for Images Shared over Messaging and Social Networking Applications,” *Proceedings of Researchfora 28th International Conference*, Aug. 2018. [Online]. Available: World Research Library, http://www.worldresearchlibrary.org/up_proc/pdf/1846-15390596449-13.pdf

[14] J. Dong, W. Wang, and T. Tan, “CASIA Image Tampering Detection Evaluation Database”. *Proceedings of the IEEE China Summit & International Conference on Signal and Information Processing (ChinaSIP)*, pp. 422–426. July 2013.