(Student) - Younis Ghirfani

Final Report

[Draw your reader in with an engaging abstract. It is typically a short summary of the document.   
When you’re ready to add your content, just click here and start typing.]

Table of Contents

[Introduction 3](#_Toc70697010)

[Problems 3](#_Toc70697011)

[Python 3](#_Toc70697012)

[Achieved and not Achieved 3](#_Toc70697013)

[Achieved: 3](#_Toc70697014)

[Analyse JPEG files and apply Benford’s Law to the file 3](#_Toc70697015)

[User Friendly GUI 3](#_Toc70697016)

[Create a live graph and progress bar 3](#_Toc70697017)

[Not Achieved: 4](#_Toc70697018)

[A.I 4](#_Toc70697019)

[Threading 4](#_Toc70697020)

[PyGame 4](#_Toc70697021)

[Multi OS platform support 4](#_Toc70697022)

[Learning Outcomes 4](#_Toc70697023)

[Benford’s Law 4](#_Toc70697024)

[JPEG Compression methods 4](#_Toc70697025)

[DCT Coefficients distribution 4](#_Toc70697026)

[Project Planning 5](#_Toc70697027)

[Time Management 5](#_Toc70697028)

[Testing 6](#_Toc70697029)

[Conclusion 7](#_Toc70697030)

[Plagiarism Declaration 7](#_Toc70697031)

# Introduction

The idea was to have an application that allows a user to enter an image and verify its legitimacy. More specifically to see if it’s a Deep fake image. Using Benford’s Law curve as a guide and using DCT comparison to Benford’s Law. A result should look in the range of less 1’s and 2’s found in the total DCT Coefficients.

# Problems Encountered

## Python

Python became a huge learning curve, structured differently to other programming languages. When installing packages, it became clear it wasn’t downloading into the correct directory. After some research and some changes, I fixed that issue as I work along the project. Sometimes a random pip install.….. command would glitch and not install properly. Also missed a crucial python setup command that allowed some packages to install with no problems. Some people have version issues when they use python, I had some issues but most of it was just changing the name of the package folder or the method I was using

## Tkinter and Matplotlib

Having Tkinter and matplotlib is great, except having them work together caused some problems for me. I found that adding a canvas before matplotlib and

# Achieved and not Achieved

## Achieved:

### Analyse JPEG files and apply Benford’s Law to the file

Using the dct function allowed python to handle the quantization of the grey number. But handling the DCT numbers to graph and allowing checking Etc. was the main part of my work including GUI.

### User Friendly GUI

The project is made as a universal vision. Buttons are large, image display is large, the window is a fixed size to prevent scaling issues and consistency.

### Create a live graph and progress bar

I found matplotlib to be very useful, it was relatively easy to setup with some issues with placement and live display. The slowest part of the image processing is the live graphing. I believe its because its live updating on the canvas too and updating all numbers from 1-9 instead of a single point. I though having some number array displayed wouldn’t be easy to follow especially for a person who doesn’t know much about Benford’s curve and how to know if it’s real or fake

## Not Achieved:

### A.I

Using A.I can make the reading more accurate and allows more reliable results. I haven’t research much into using A.I unfortunately ran out of time. Having A.I would open a new part of DCT and Benford’s Law in the future

### Threading

I have tried many times to make threading work with my project, but never seems to work. After some research, it seems Tkinter is the issue. As its live updating, Tkinter must be threading as well as the processing. I’m not sure if it will make the project faster in processing, in the future that’s something I can look at fixing.

### PyGame

For a brief time, I though adding a small snake game would help let the time pass which its processing the image, not essential but could be implicated

### Multi OS platform support

While an initial feature, I didn’t see the benefit this would bring considering my time range. It would be a good feature for the future to look at

# Learning Outcomes

## Benford’s Law

I never heard of a math equation like Benford’s Law until I watched an episode on Netflix, I was initially confused to how accurate it is. But after reading some articles and thesis.

## JPEG Compression methods

JPEG compression is the key to verifying an image, it uses DCT which uses cosine functions to find its data. As an image gets more compressed, the quality “Q” will be lower (100 meaning full quality). Resulting in lower number of 1’s. As image processing, isn’t really a subject of my subject I found understanding JPEG very new compared to Secure applications or Reverse Engineering and Malware Analysis

## DCT Coefficients distribution

One of the most important part of my project, I found DCT a bit difficult to understand as a lot of it is math equations changing data through graphs.

# Project Planning

## Time Management

I was delayed for a few weeks after charismas, just with documentation and software. But within about 2 week I caught up with some assignments being postponed.

My original schedule:

23rd October – 6 November-

* Collect photo database reference
* Understand Bedford’s Law and learn how to implement it into Deep fake
* Create layout for Research Document
* Setup GitHub
* Start multi-user interface

6th November – 20th November

* **13th Research Manual**
* Have functional Website that allows upload of images
* Store current and previous images
* Work on GUI

20th November – 4th December

* **27th Functional Specification**
* Have good bit on total document done
* Work on my analysis aspect of probability
* Finish GUI

4th December – 17th December

* **11th Design Manual**
* Have an uploaded working application uploaded to GitHub
* First draft of my document
* Finish multiple user GUI and backend

18th December – 1st January

* **18th Presentation**
* Final testing
* Upload updates to GitHub
* Compare results to other experiment

1st January – 14th January

* Final document draft
* Complete test of the project

14th January – 18th April

* **Second Presentation**
* Final testing and tuning
* Finished document, project, and code

19th April

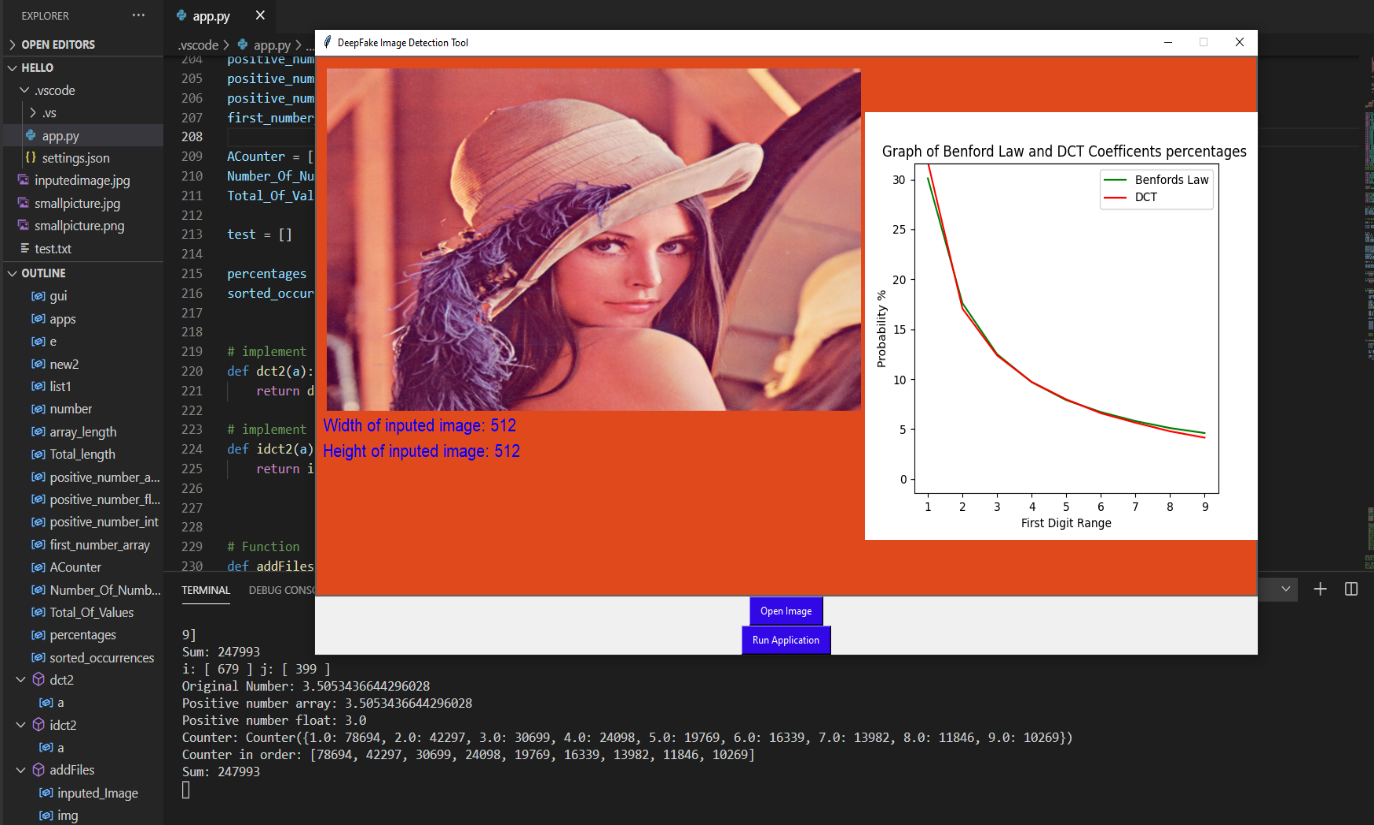
* **Technical Manual & Final Report**

From looking at this, I can see I was ahead of schedule for the first few months, but as assignment was assigned, I fell behind a for a bit.

# Testing



Lenna is the most commonly used in image forensics since 1973. For my project I will use it as my control image.



The counter in order is the numbers representing 1-9. These number are made in a percentage of the array

Benford’s Law: 30.1%

78694 / 247993 \* 100 = 31.7%

Very similar result to Benford’s Law. if the number 1’s is 25% or lower, it is considered to be a deep fake image. The compression process causes the DCT Coefficients to create a smaller number of 1’s

# Conclusion

Overall, I’m happy with my project, I applied dct and used Benford’s Law. The dct function has many types and different version/iterations. I tried to find data/ research on what type of dct is used but most of the time it referenced to MATLAB and how its applied to their software. But other than dct function changes, I’m happy with the GUI processing. I don’t think its possible to make it more efficient as I tried skipping the first step as a test and it didn’t work well. I would say to my old self to get more work done over Christmas before more assignments after Christmas

I think with threading and more accurate reading on if its real or fake. This project has potential to be a faster convenient way to see if an image is real or fake

# Plagiarism Declaration

I declare that all material in this submission is entirely my own work except where duly

acknowledged. I have cited the sources of all quotations, paraphrases, summaries of information,

tables, diagrams, or other material; including software and other electronic media in which intellectual property rights may reside. I have provided a complete bibliography of all works and sources used in the preparation of this submission. I understand that failure to comply with the Institute’s regulations governing plagiarism constitute a serious offence.

*Name: Younis Ghirfani*

*Student number: c00229681*

*Date:30/4/21*