CSCI 4150/ CSCI 6050/ DASC 6050: Programming Assignment 01, Image Forensics

Farnoosh Koleini

Nov 30, 2022

Goal: The overarching goal of this assignment is to determine if an image is tampered using Benford's law.

After understanding the concepts of Benford's law. Now, it was time to take things more visual. So, I took an image (repeated on 4 other different images) and started to think what numeric information could be extracted from it. As we all know, images are made of pixels and the pixel value is typically an 8-bit data value (with a range of 0 to 255). Below is one of the images that I took for analysis.

I had to evaluate whether the pixels of this image followed Benford's Law or not. Using OpenCV2 and NumPy libraries from Python3, I read the image and evaluated the discrete cosine transform as I had to remove all the background noise from the image for analysis. The discrete cosine transform (DCT) helps separate the image into parts (or spectral sub-bands) of differing importance (with respect to the image's visual quality). The DCT is similar to the discrete Fourier transform: it transforms a signal or image from the spatial domain to the frequency domain.

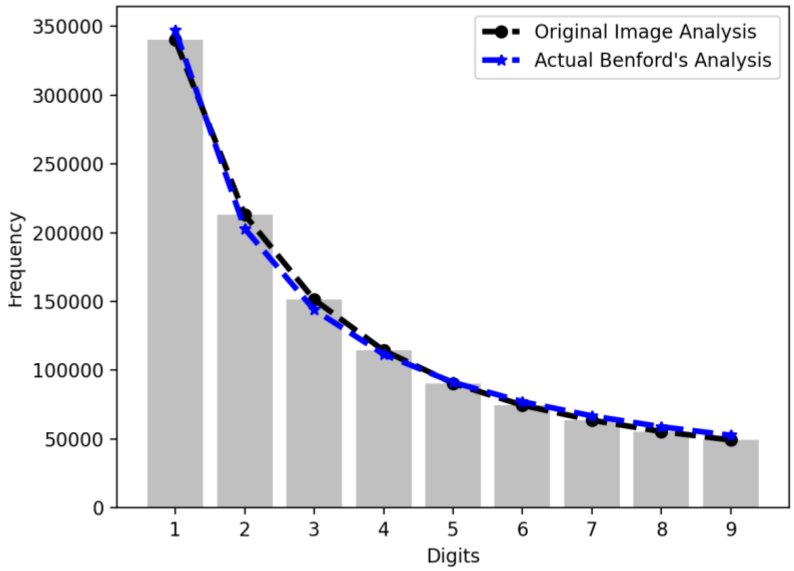
The general equation for a 2D (N by M image) DCT is defined by the following equation:

$$F(u,v) = \left(\frac{2}{N}\right)^{\frac{1}{2}} \left(\frac{2}{M}\right)^{\frac{1}{2}} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \Lambda(i).\Lambda(j).cos\left[\frac{\pi.u}{2.N}(2i+1)\right] cos\left[\frac{\pi.v}{2.M}(2j+1)\right].f(i,j)$$

and the corresponding inverse 2D DCT transform is simple F-1(u,v), i.e.:

$$\Lambda(\xi) = \begin{cases} \frac{1}{\sqrt{2}} & \text{for} \xi = 0 \\ 1 & \text{otherwise} \end{cases}$$

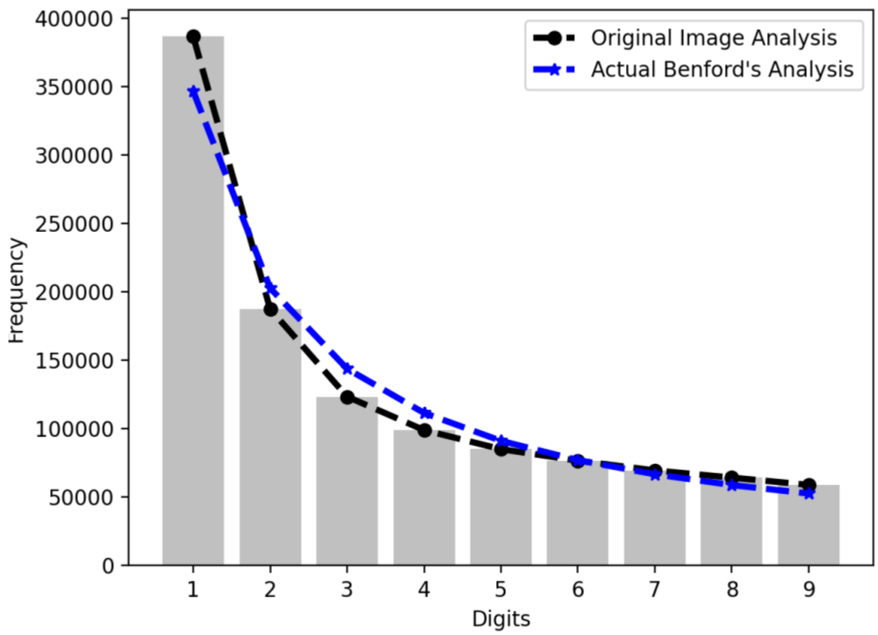
Next course of action was to flatten the 2D matrix into a 1D vector to ease the analysis of pixel numbers. After the complete analysis of the 1D pixel vector, I was shocked again. The feature vector clearly followed Benford's algorithm. The distribution of the digits among the pixel values in the vector was the same as suggested by Benford's Law. Below is the graph of the distribution of the digits in the feature vector with the actual distribution of the digits that Benford predicted.



The graph clearly says that the image is original/true as it follows Benford's Law. At that time, I just wanted to clap looking at the beauty of numbers.

After recovering from the previous shock, I did a little experiment. I modified the original image, by applying some filters on it. Now, the image is tampered and is no more original. Below is the tampered image that I will analyze to check whether it is a tampered image or not (based on Benford's Law).

I followed the same methodology as before, by firstly applying DCT (discrete cosine transform) on the image pixels and then, generating a 2D matrix of the same. Then, flattened the image 2D vector and generated a 1D vector for analysis. After analysis, the results totally shocked me again. It was crystal clear that the image was tampered as the distribution of the digits in the 1D vector did not follow Benford's Law. Below is the graph representing the distributions of digits in 1D vector of the tampered image with the Benford predicted distribution.



Well, as it is clear from the difference in distribution, Benford's Law was able to detect the tampered image.

This all looks fascinating, now one can easily differentiate between an original image and a tampered image. This can clearly help in the scenarios where there are chances of frauds w.r.t the identity of a person.