Digital Forensics Labwork 5 Report

Introduction

What is Digital Forensics

Digital forensics (sometimes known as digital forensic science) is a branch of forensic science encompassing the recovery and investigation of material found in digital devices, often in relation to computer crime. In this document, descriptions and examples will be provided to analyze Digital Forensics in Images Processing.

Techniques in detecting images tampering

A few techniques can be named such as:

1. Re-sampling.
2. Double JPEG Compression.
3. Luminance Non-linearities.
4. Signal to Noise Ratio.

The importance

Digital Forensics helps identifying direct evidence of a crime, digital forensics can be used to attribute evidence to specific suspects, confirm alibis or statements, determine intent, identify sources (for example, in copyright cases), or authenticate documents.

Methods

Re-sampling

Re-sampling an image is required considering the scenario when digital fogery is created by splicing multiples images together. Though re-sampling is imperceptible, specific correlations which represent evidence of tampering are detected in re-sampled images. First, detect re-sampling correlations in 1-D signal, depending on the re-sampling rate might the process introduce correlations of varying degrees between neigboring samples, then generalize 2-D images.

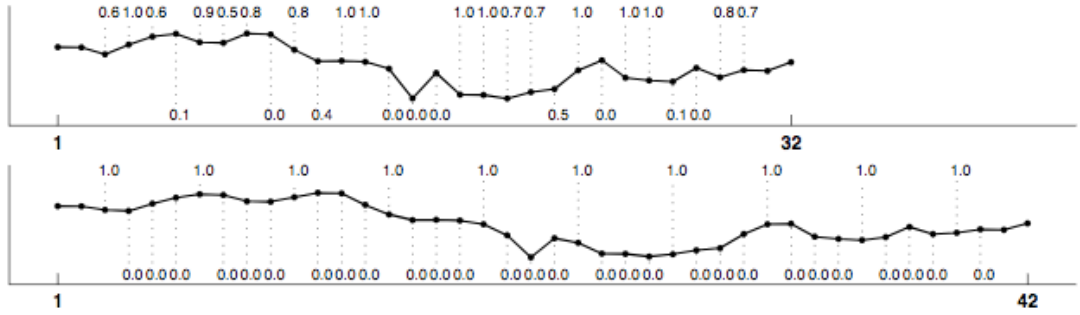


Fig. 1: A signal with 32 samples (top) and this signal re-sampled by a factor of 4/3 (bottom). Each sample is annotated with its probability of being correlated to its neighbors. Note that for the up-sampled signal these probabilities are periodic, while for the original signal they are not.

Given a signal that has been re-sampled, The expectation/ maximization algorithm (EM) is developed to simultaneously estimate a set of periodic samples that are correlated to their neighbors, and the specific form of these correlations. The extension to 2-D images is relatively similar and EM algorithm can also be used with 2-D images for generalization.

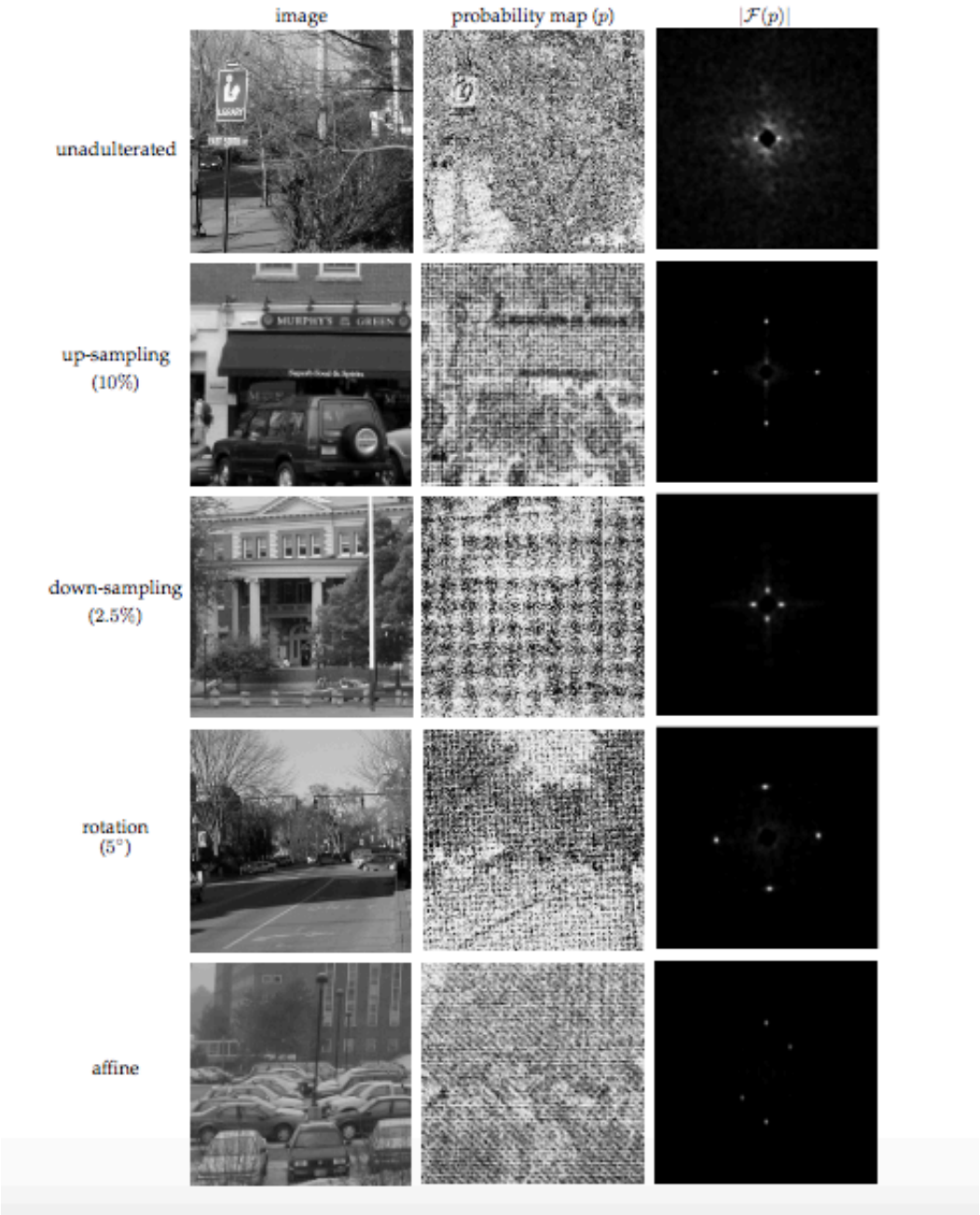


Figure: Shown in the top row is an unadulterated image, and shown below are images re-sampled with different parameters. Shown in the middle column are the estimated probability maps that embody the spatial correlations in the image. The magnitude of the Fourier transforms of these maps are shown in the right-most column. Note that only the re-sampled images yield periodic maps.

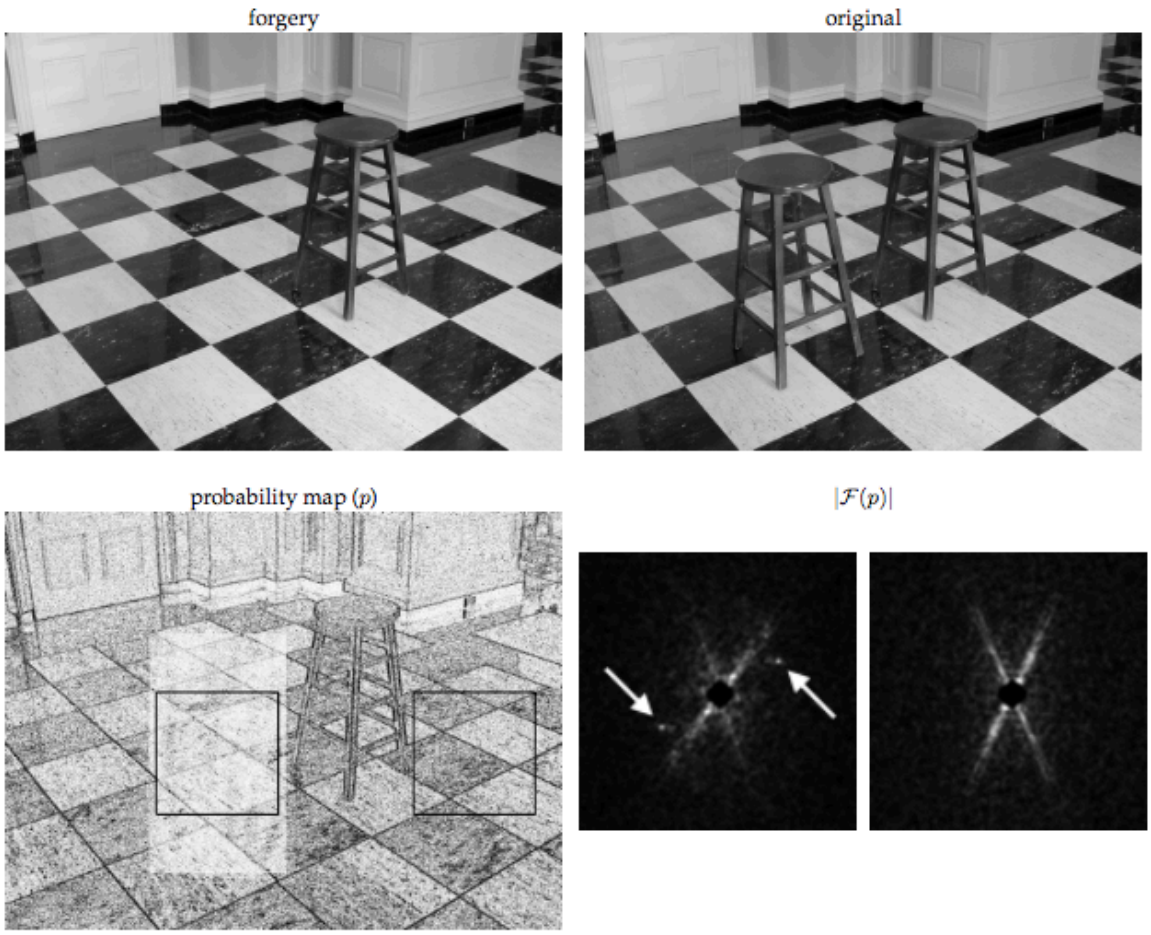


Fig. 3: Shown along the top row is a forgery and the original image. The forgery consists of removing a stool and splicing in a new floor taken from another image (not shown here) of the same room. Shown below is the estimated probability map (p) of the forgery, and the magnitude of the Fourier transform of a region in the new floor (left) and on the original floor (right). The periodic pattern (spikes in $|\mathcal{F}(p)|$) in the new floor suggest that this region was re-sampled.

Double JPEG Compression

Tampering digital images requires editing software like Adobe Photoshop. In a JPEG image, any editing is carried out undergoes re-compression. Detecting evidence of double JPEG compression may provide useful information on revealing traces of images tampered or fabricated manipulation.

What is JPEG Compression

JPEG is a standardized image compression procedure proposed by Joint Photographic Experts Committee (JPEG). Two compression schemes are needed for the JPEG standard to be generally applicable: a lossless predictive scheme and a lossy scheme based on the Discrete Cosine Transform (DCT). Encoding an image involves three steps:

- 1. Discrete Cosine Transform.
- 2. Quantization.
- 3. Entropy Encoding.

The decoding of a compressed image is also involves the above three steps but in reversed order.

What is Double Quantization

Quantization is a point-wise operation described by a one-parameter family of functions:

$$q_a(u) = \left\lfloor \frac{u}{a} \right\rfloor ,$$

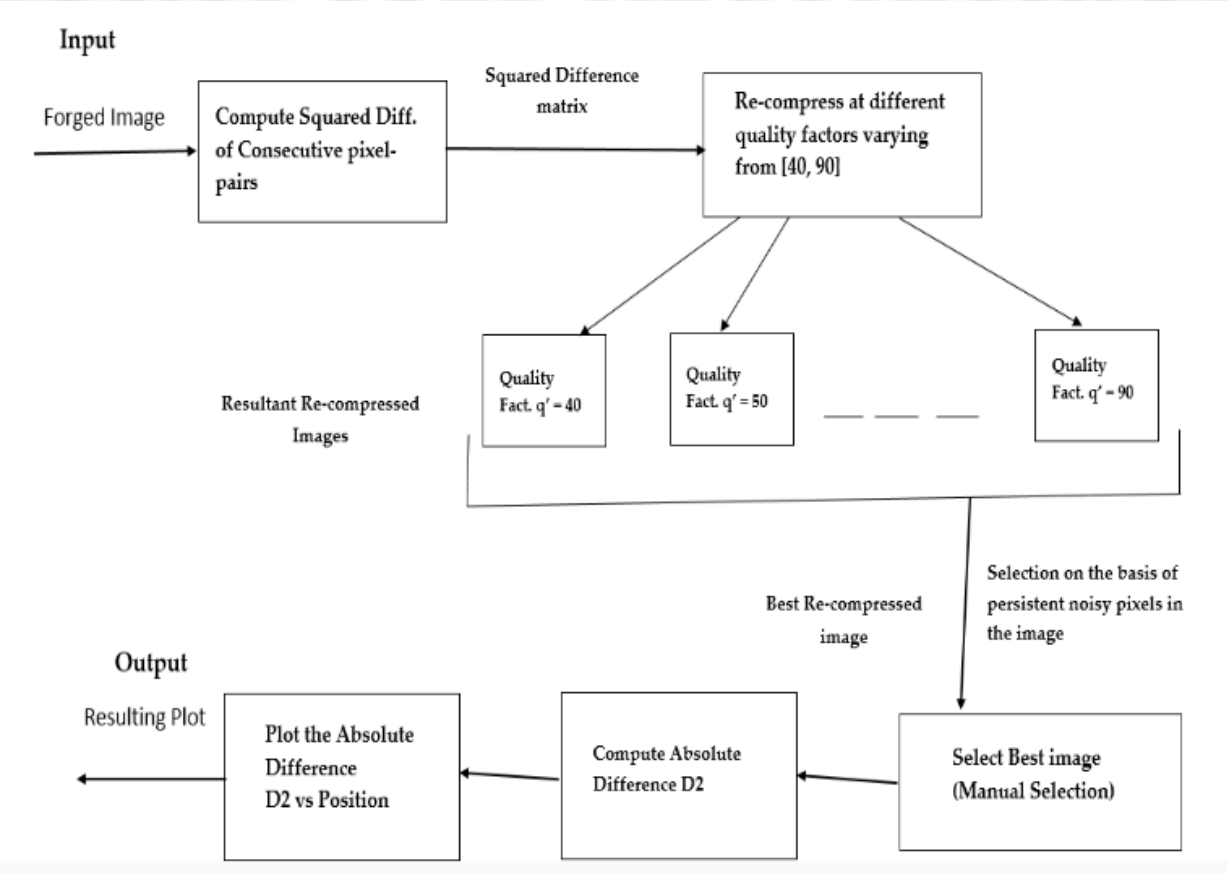
Where a is the quantization step (a strictly positive integer), and u denotes a value in the range of x[t] – a generic discrete 1-D signal. Double quantization is the same but two-parameters:

$$q_{ab}(u) = \left\lfloor \left\lfloor \frac{u}{b} \right\rfloor \frac{b}{a} \right\rfloor ,$$

Proposed methods for the detection of JPEG Double Compression

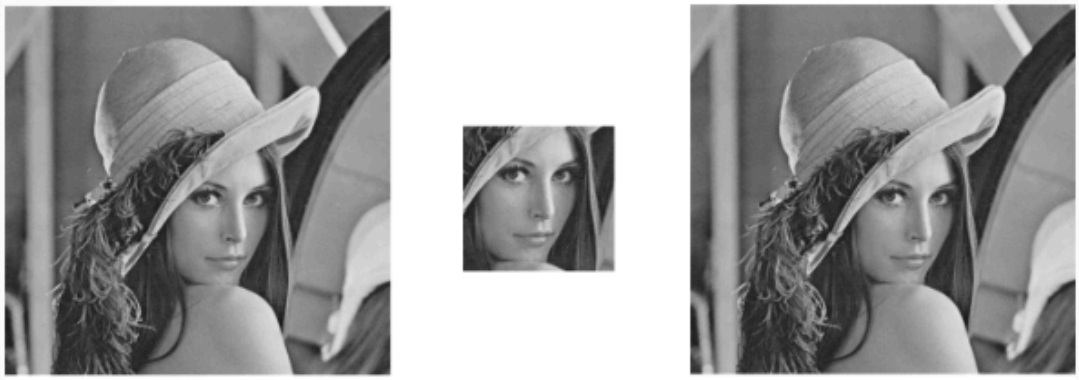
Method 1: Multi-compression based on JPEG Forgery Detection

Since JPEG image undergoes re-compression when editing, it is exploitable to detect any illegal modifications via the algorithm:



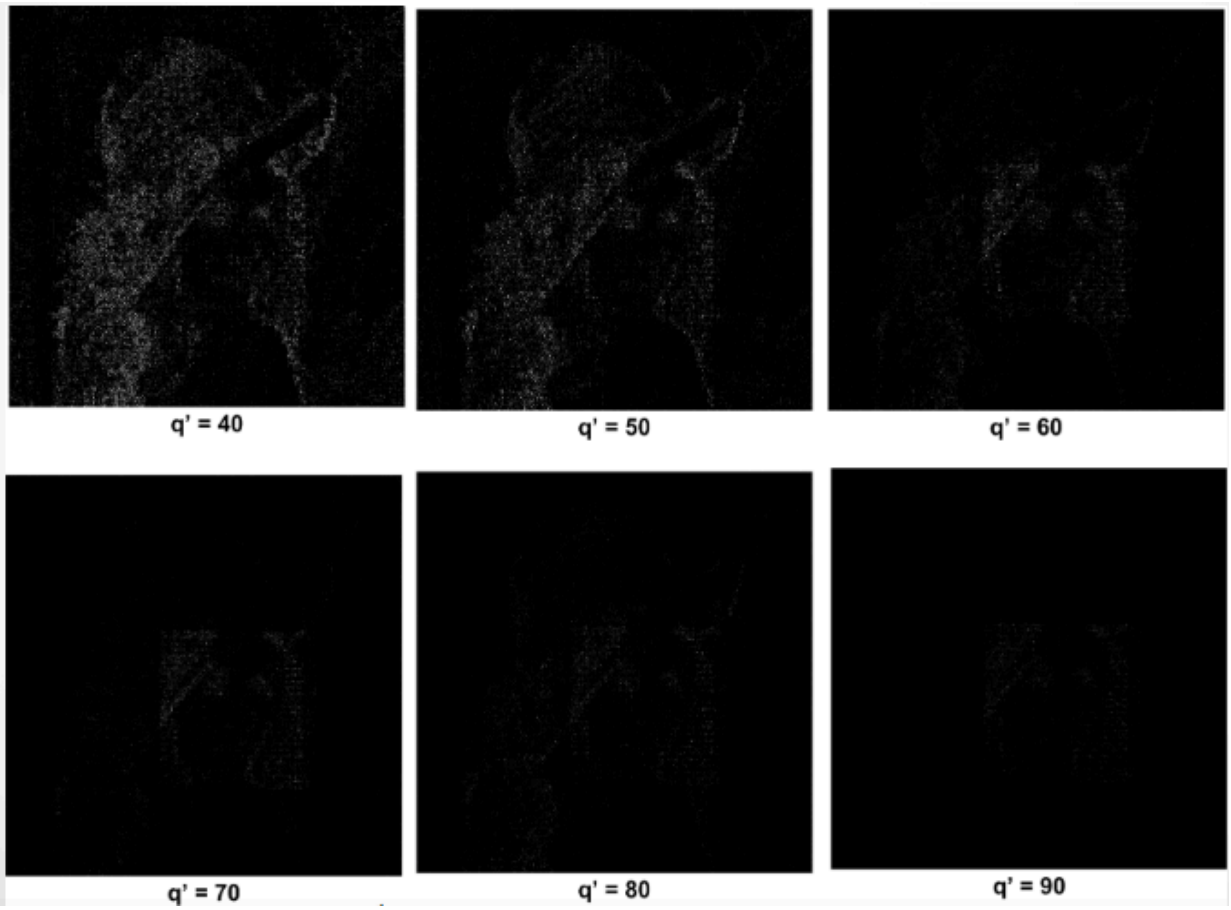
Results:

(Manually) Tampered Lena Image



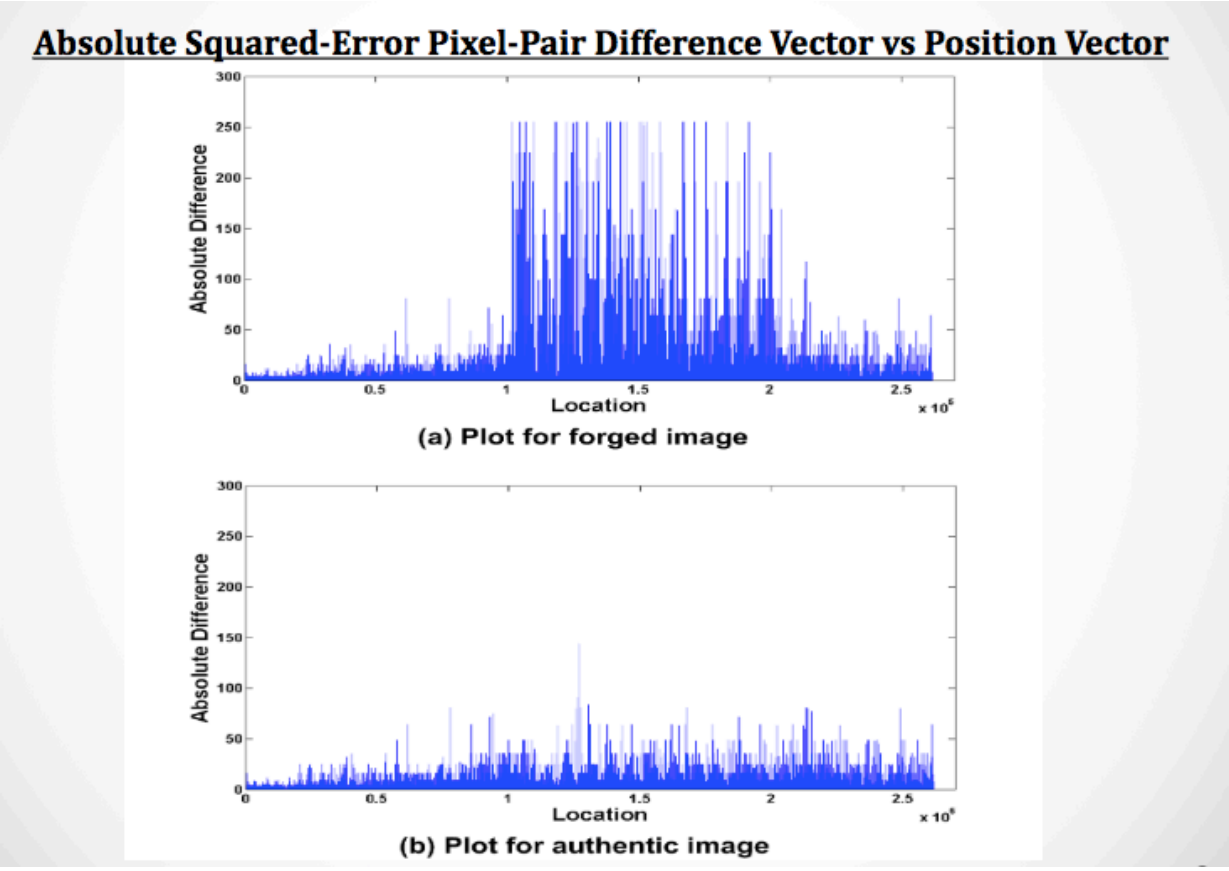
(a) (b) (c)

(a) Original 512×512 image
(b) Central 200×200 portion, re-saved at a different degree of compression
(c) Forged image having its central portion modified



q' = 40 q' = 50 q' = 60

q' = 70 q' = 80 q' = 90



Method 2: Detection based on Convolutional Neural Networks

The proposed Convolutional Neural networks(CNNs)-based method does not use histogram features as input and provides end-to-end detection capability. The network proposed architechture named JPEG-CNN involves these:

- Take raw DCT coefficients of JPEG image as input without any handcrafted preprocessing operation (histogram).
- Adapt to JPEG DCT sub-band structure via a multi-branch architecture.
- Some layers such as Absolute value (ABS) layers and Batch Normalization (BN) layers are used to improve network performances.

Method 3: Detection based on Histograms.

Luminance Non-linearities

Mordern imaging devices normally would be able to enhance images' quality via some parameters of Luminance Non-linearities, which depends on cameras or scenes. Since images can be taken by different conditions, we can use tools from polyspectral analysis to detect luminance non-linearities that also introduce some special artifacts named correlations in the Fourier domain.

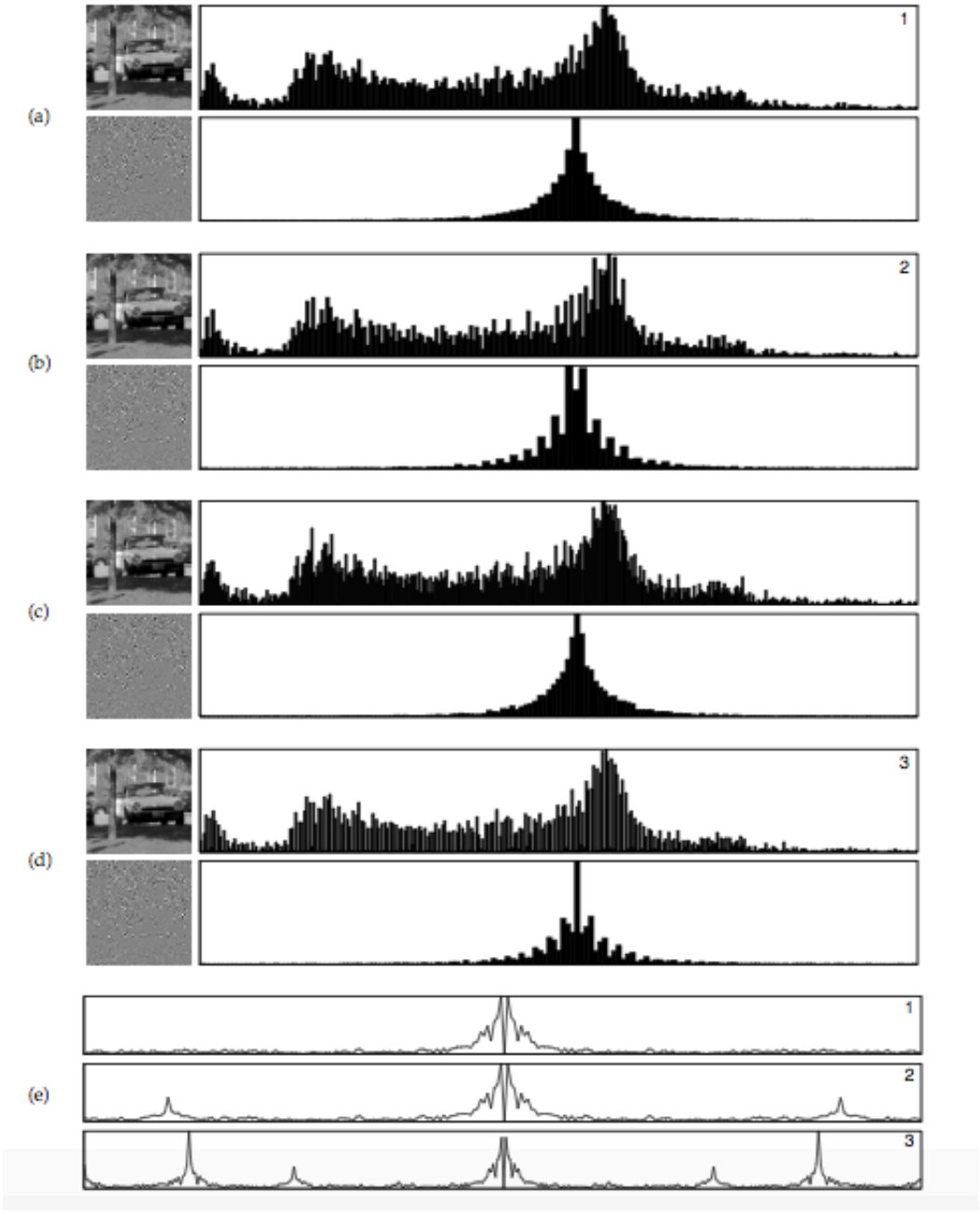


Figure: Shown in the top four panels are DCT coefficients for two frequencies ((1,1) and (2,2)), and their histograms for single and double compressed JPEG images: (a) single JPEG compression with quality 75, (b) double JPEG compression with quality 85 followed by 75, (c) single JPEG compression with quality 85, (d) double JPEG compression with quality 75 followed by 85. Shown in panel (e) are the Fourier transforms of three zero-meaned histograms. Note the periodic artifacts introduced by double quantization (panels 2, 3) reflected by the high frequency peaks in the Fourier transforms.

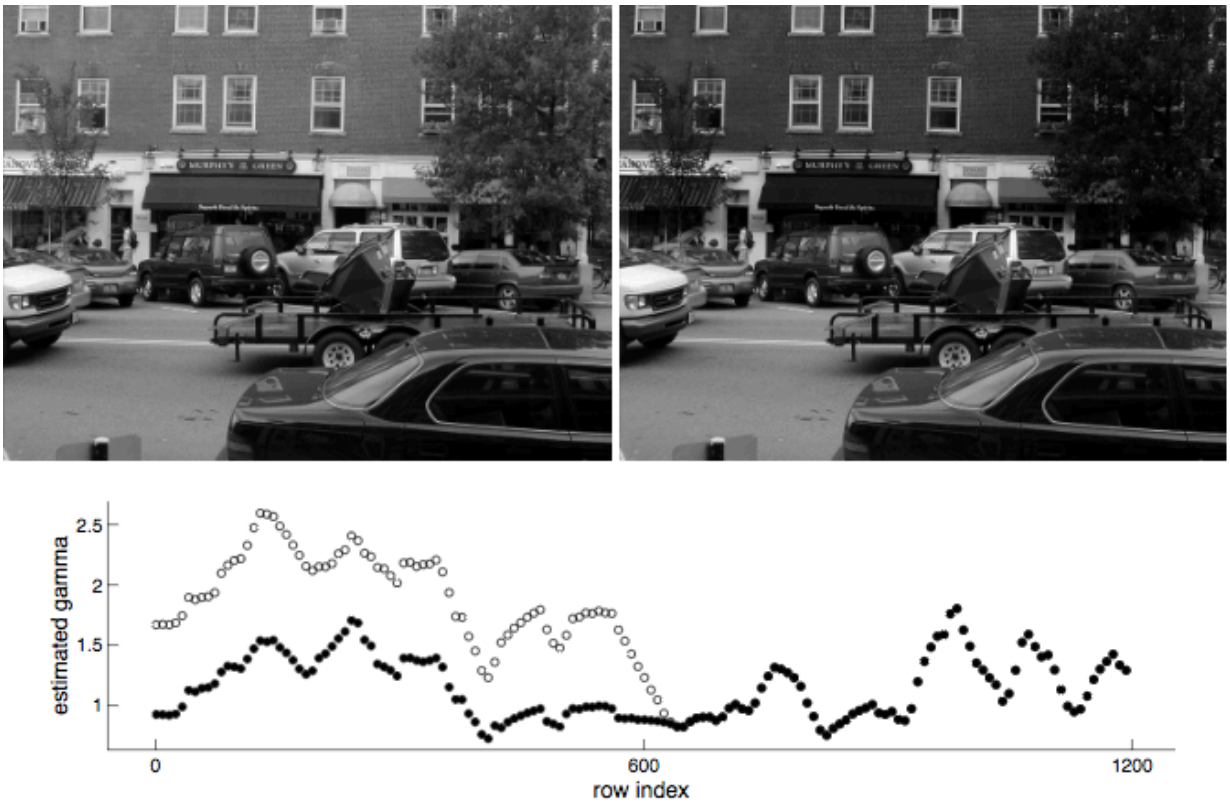


Fig. 6: Top panel: a natural image (left), and the same image whose top portion was gamma corrected with $\gamma = 1.8$ (right). The images are 1200×1600 pixels in size. Bottom panel: Estimated gamma values from horizontal scan lines, where the black dots correspond to estimates from the unadulterated image, and the white dots correspond to estimates from the image whose upper half has been gamma corrected. Each data point corresponds to a running average over 60 scan lines.

Signal to Noise Ratio

Image processing or digital compression produces noices to images and these amount can vary between images. Detection of noices in multiple spliced images or variations in the signal to noise ratio (SNR) can reveal traces of tampering.

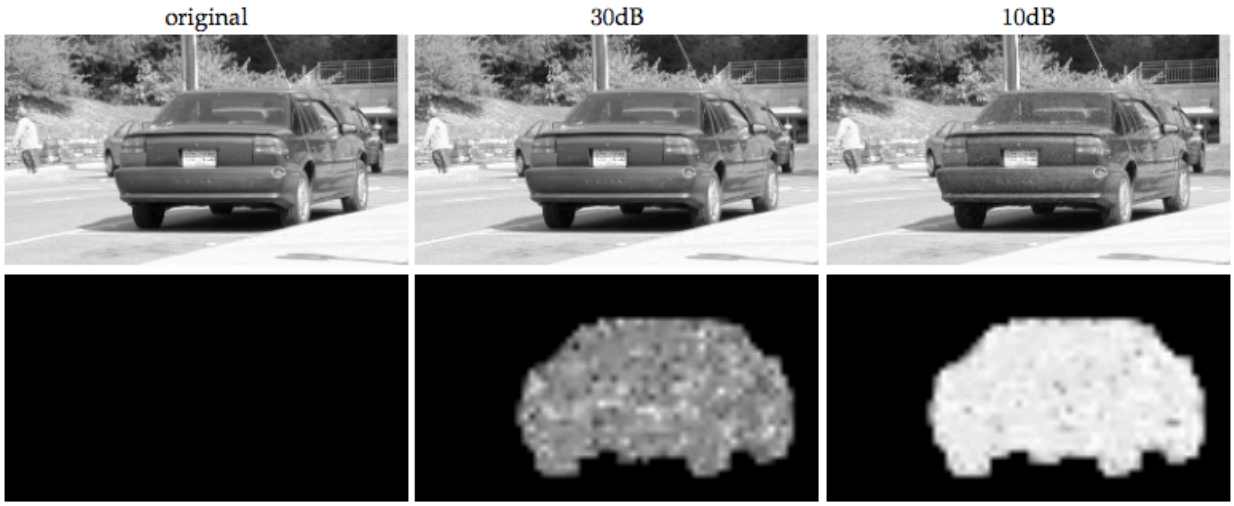


Fig. 7: Shown on the top row is an original image and this image with noise added locally to the car. Shown on the bottom row are the locally estimated noise variances (on the same log scale).

References

1. Wikipedia: Digital Forensics.
2. Alin C. Popescu, Hany Farid, '*Statistical Tools for Digital Forensics*'
3. Babak Mahdian, Stanislav Saic, '*Image Tampering Detection Using Methods Based on JPEG Compression Artifacts: A Real-Life Experiment*'
4. Bin Li, Hu Luo, Haoxin Zhang, Shunquan Tan, Zhongzhou Ji, '*A multi-branch convolutional neural network for detecting double JPEG compression*'
5. Pankaj Malviya, Ruchira Naskar, '*Digital Forensic Technique for Double Compression based JPEG Forgery Detection*'