Computer Vision I: Python Exercise / project #1 (total 10 points)

Due October 8, 2024, 2 pm

[Boyuan Chen] [2200017816]

This is a small project designed to confirm certain characteristics of natural image statistics, as previously discussed in Chapter 2. Among the materials, you will find three distinct image sets:

1. Set A consists of four natural images, including two captured in urban settings and two in countryside scenes.
2. Set B comprises two man-made images, specifically paintings by Vincent van Gogh.
3. Set C encompasses a random noise image.

Python Library: To complete the project, please install the required dependencies provided in the *environment.yml* file (including cv2, PIL, numpy, scipy, matplotlib, tqdm). You are also welcome to utilize any libraries of your choice, **but please report them in your report (for autograder)! Again, report any customized library in the report (do not go too crazy as this will add a significant burden to TAs).**

What to hand in: Please submit both a formal report and the accompanying code. For the report, kindly provide a PDF version. You may opt to compose a new report or complete the designated sections within this document, as can be started by simply loading the tex file to Overleaf. Your score will be based on the quality of **your results**, **the analysis** (diagnostics of issues and comparisons) of these results in your report, and your **code implementation**.

Notice: Hint code is provided, please **don't change** the name of the functions for automatic scoring, but feel free to add new functions.

1 Problem 1

(High kurtosis and scale invariance, 4 points). *Please read through this question description before you start. For this problem, please consider 3 sets: set A, set B, set C.*
For the sake of computational efficiency,, convert the images to grayscale and subsequently rescale the intensity levels to fall within the range [0, 31], i.e. 32 grey levels. Please note that this rescaling operation is not necessary for the random image selected from set C. Then convolve the images with a gradient filter denoted as $\nabla_x I$, which signifies the intensity difference between two adjacent pixels along the horizontal axis. Accumulate the histograms by computing the averages over all images within each respective set. To illustrate, the process of accumulating the histogram for "set A" entails the aggregation of histograms from all four natural images found within "set A". Subsequently, complete the following steps for all sets.

1.1 Subproblem 1

Question: Plot the histogram $H(z)$ for the difference against the horizontal axis $z \in [-31, +31]$. Then do a log-plot $\log H(z)$. [Some bins will be zero, you can assign ϵ for such bins in the log-plot].

Answer: As shown in Figure 1 and Figure 2, we visualized the histograms and their corresponding logarithmic values for different datasets. The observations and related analysis are as follows:

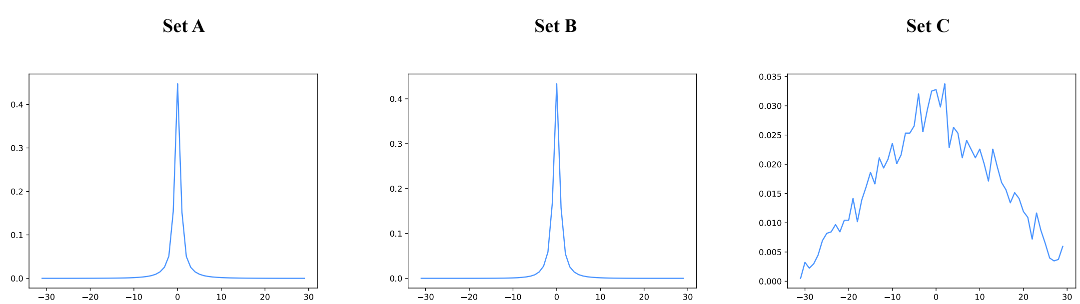

Figure 1: Histogram $H(z)$ of different dataset.

- **Set A (Natural Images)**: We observed a steep peak at $z = 0$, as natural images tend to have smooth transitions between adjacent pixels, resulting in small gradient differences.

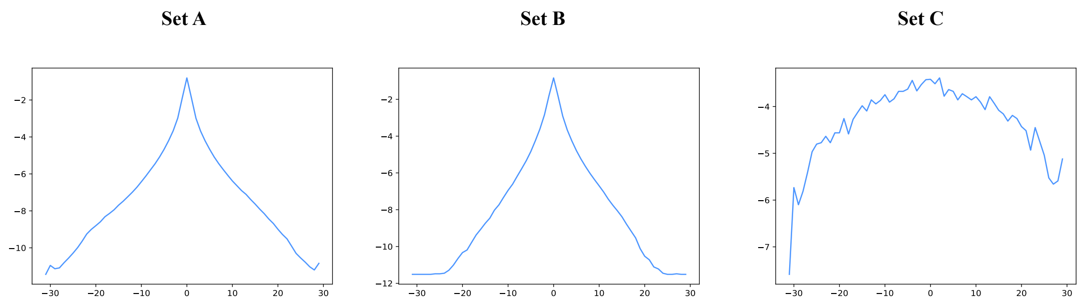

Figure 2: log Histogram $H(z)$ of different dataset.

- **Set B (Van Gogh Paintings)**: The gradient smoothing trend is almost identical to that of Set A, except for some abrupt changes. This indicates that Van Gogh paintings exhibit a similar gradient behavior to natural scenes in this processing.
- **Set C (Random Noise Image)**: The histograms are evenly distributed, as there is no structured relationship between adjacent pixel intensities in a random noise image.

1.2 Subproblem 2

Question: Compute the mean, variance, and kurtosis for this histogram [Report the numeric numbers in your report].

Answer: As shown in Table 1, we calculated the mean, variance, and kurtosis for the histograms of different images, and it was observed that the statistical characteristics vary across the datasets.

Table 1: Statistical characteristics.

Image	Mean	Variance	Kurtosis
Set A	-0.0007997729117050767	6.380297660827637	18.811634063720703
Set B	0.0005937975365668535	4.391456604003906	14.579910278320312
Set C	-0.010416666977107525	171.1199188232422	2.390532970428467

Based on the provided table, here's the revised analysis of the results:

- **Set A (Natural Images)**:
  The mean gradient difference is close to zero, indicating very smooth transitions between adjacent pixels, which is typical for natural images. The variance of approximately 6.38 reflects moderate variations in pixel intensity across the image,

suggesting that most areas undergo gradual intensity changes. The high kurtosis (18.81) indicates that most of the gradient differences are near zero, contributing to a sharp peak in the histogram and highlighting the smooth, continuous nature of natural scenes.

- **Set B (Van Gogh Paintings)**:
  The mean is slightly positive, indicating a slight overall shift in pixel intensity differences, but still close to zero, reflecting overall smooth transitions with some localized abrupt changes (e.g., brushstrokes). The variance is lower (4.39), suggesting even more gradual changes compared to Set A. The kurtosis (14.58) is also high, indicating a significant concentration of gradient differences near zero, despite some sharp transitions, which aligns with the textured but mostly smooth nature of Van Gogh's paintings.

- **Set C (Random Noise Image)**:
  The mean is much lower than in Sets A and B, at -0.010, indicating that random noise images exhibit significantly larger and more varied gradient differences. The variance is extremely high (171.12), which reflects highly irregular and random intensity changes between adjacent pixels. The kurtosis (2.39) is much lower, implying a flatter histogram with more dispersed gradient values, which is characteristic of unstructured random noise where pixel differences are distributed more uniformly.

1.3 Subproblem 3

Question: Fit this histogram to a Generalized Gaussian distribution $e^{|z/\sigma|^\gamma}$ and plot the fitted-curves super-imposed against the histogram. What is the value of γ in the fitted generalized Gaussian?

Answer: As shown in Figure 3, we fit the histogram to a Generalized Gaussian distribution $e^{-|z/\sigma|^\gamma}$ and plot the fitted curves superimposed on the histogram. The curves for the former two image sets are generally similar, while Set C shows a noticeably flatten tendency.

We also computed the corresponding values of σ and γ for the fitted Generalized Gaussian Distribution (GGD), as shown in Table 2. Notably, the fitted σ for Set C is significantly larger than those for Sets A and B, indicating much greater variability in the pixel intensity differences, which is characteristic of random noise. The γ value for Set C is negative and close to zero, reflecting the flat and unstructured nature of the noise histogram. In contrast, Sets A and B have relatively small σ and positive γ, consistent

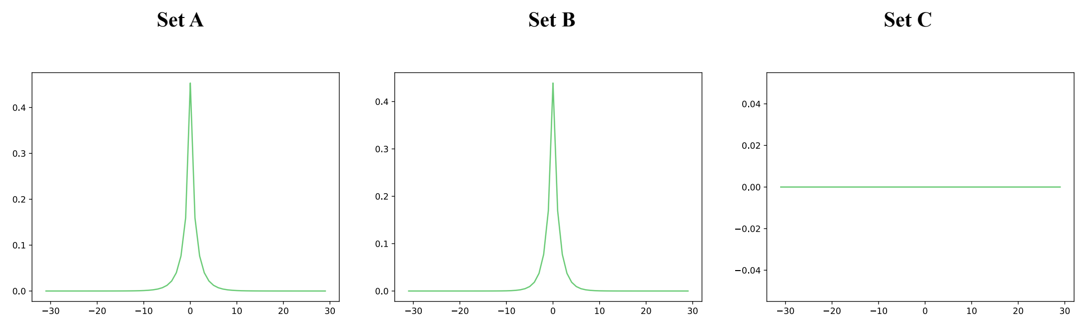

Figure 3: Fitted GGD curves.

with the smoother and more structured intensity changes typical of natural and artistic images.

Table 2: Fitted GGD related values.

Image	Fitted σ	Fitted γ
Set A	0.9450746326994838	0.7690207510720445
Set B	1.0585722611868367	0.8659580647731308
Set C	6.810281106208066	-0.08963886991500347

1.4 Subproblem 4

Question: Plot the Gaussian distribution using the mean and the variance in step (2), and super-impose this plot with the plots in step (1) above (i.e. plot the Gaussian and its log plot, this is easy to do in python with matplotlib).

Answer: As shown in Figure 4, we plot the Gaussian distribution and its log plot using the mean and variance from step (2), and superimpose this plot with the results from step (1).

In Figure 4, the lower part displays the corresponding log plot, contrasting the Gaussian distribution curve fitted with the mean and variance against the original data.

1.5 Subproblem 5

Question: Down-sample your images by a 2×2 average (or simply sub-sample) the image. Plot the histogram and log histogram, and impose with the plots in step 1, to compare the difference. Repeat this down-sampling process 2-3 times.

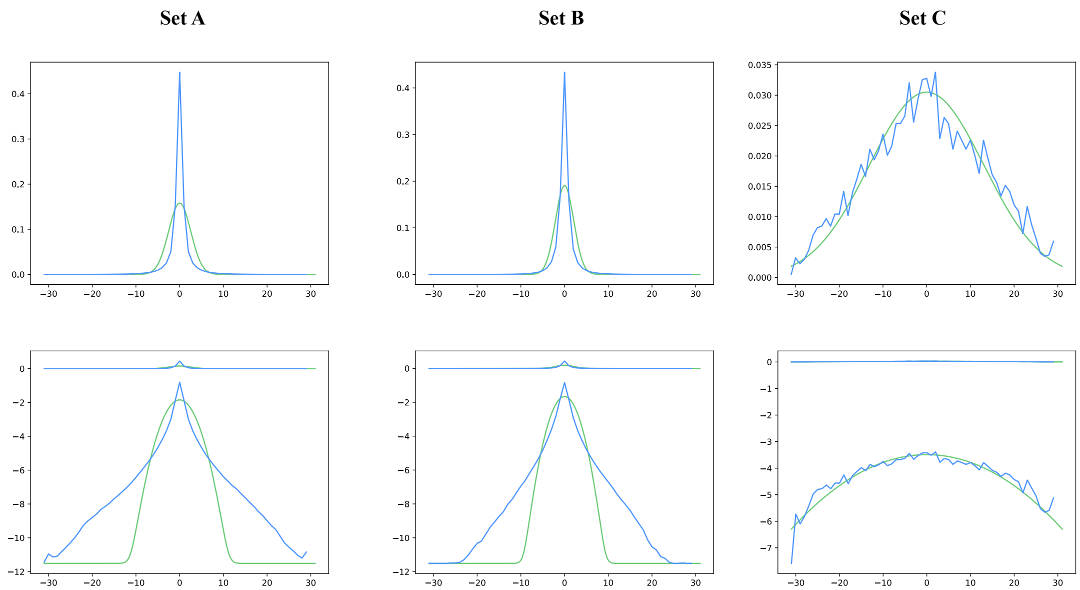

Figure 4: (log) Gausian Distribution and super-imposed with the above histograms.

Answer: As shown in Figure 5, we performed several rounds of downsampling (2×2 average) on the three datasets and then plotted their respective histograms and log histograms for comparison. It was observed that for natural images, the number of downsampling rounds does not significantly affect the results. For random noise images, as the number of downsampling rounds increases, the distribution becomes less uniform.

This can be explained by the fact that natural images contain smooth pixel transitions, so averaging adjacent pixels during downsampling preserves the overall structure and smoothness, resulting in minimal changes to the histogram. However, random noise images are highly sensitive to downsampling because the original pixel intensities are random and uncorrelated. As downsampling increases, these pixel values are averaged, reducing the overall randomness and causing the histogram to become more structured.

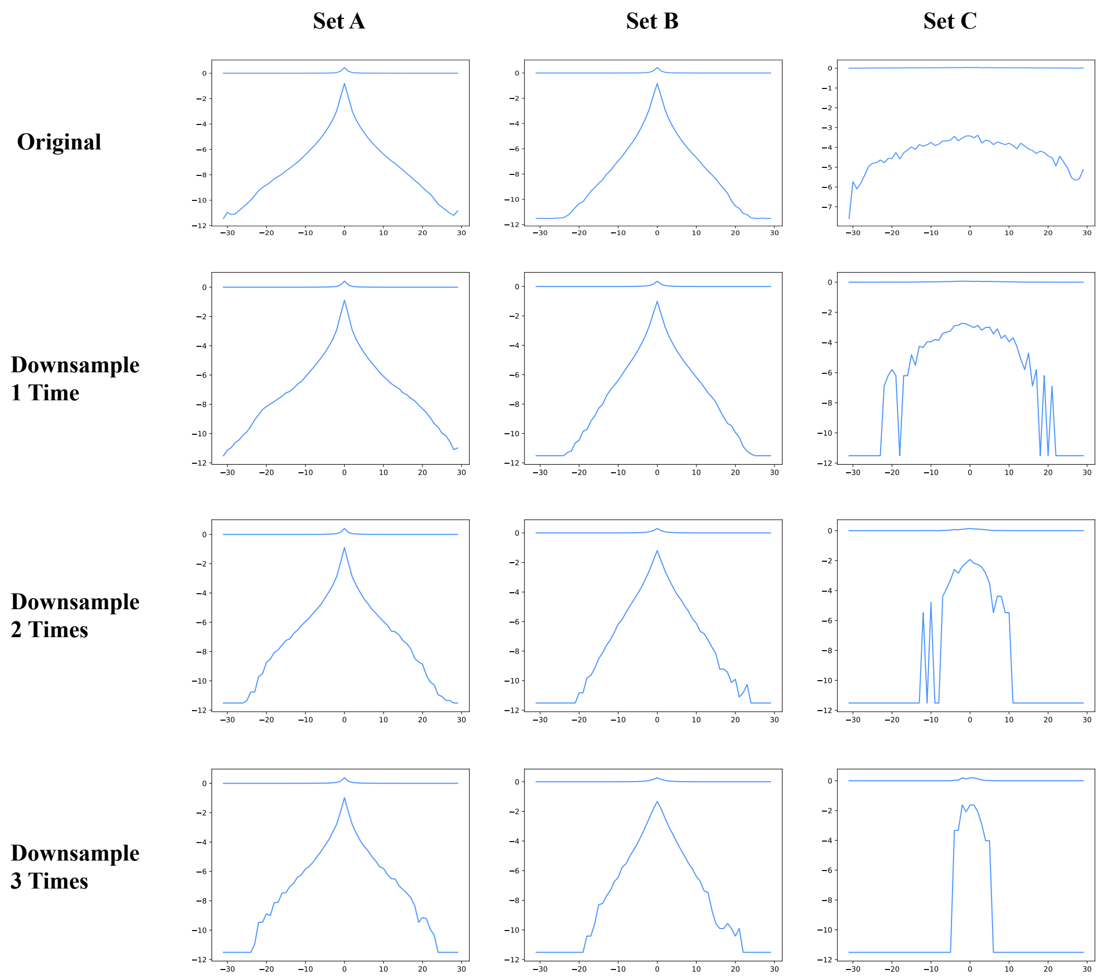

Figure 5: The effect of different downsampling rounds on the histogram and log histogram distributions.

2 Problem 2

(Verify the 1/f power law observation in natural images, 3 points). *Please read through this question description before you start. For this problem, please only consider set A.* Perform a Fast Fourier Transform (FFT) on the grayscale image denoted as I. This operation will yield a Fourier image denoted as $\hat{I}(\xi, \eta)$, which is a complex-number matrix indexed by horizontal and vertical frequencies (ξ, η). Subsequently, compute the amplitude (modulus) of each complex number, denoted as $A(\xi, \eta)$, and it can be calculated as follows: $A(\xi, \eta) = |\hat{I}(\xi, \eta)|$. Denote the frequency $f = \sqrt{\xi^2 + \eta^2}$, and to facilitate further analysis, convert the data to polar coordinates. In this coordinate system, calculate the total Fourier power, denoted as $A^2(f)$, for each frequency. To achieve this, discretize the values of f and compute $A^2(f)$ averaged over the respective ring corresponding to each value of f. It is important to terminate the process when the circle reaches the boundary of the Fourier image.

2.1 Subproblem 1

Question: Plot $\log A(f)$ against $\log f$. This should be close to a straight line for each image. Plot the curves for the 4 images in one figure for comparison. [Hint: First check the magnitude spectrum $\log |\hat{I}(\xi, \eta)|$, $|\hat{I}(0, 0)|$ typically is the largest component of the spectrum. The dc component is usually moved to the center of the frequency rectangle.]

Answer: In natural images, the **Fourier amplitude** $A(f)$ often follows a power law distribution of the form:

$$A(f) \propto \frac{1}{f^\alpha}$$

where f is the frequency, and α is a constant that determines the steepness of the curve. This is known as the **1/f power law**.

To linearize this power law relationship, we take the logarithm of both sides:

$$\log A(f) = -\alpha \log f + \log C$$

Thus, if the Fourier amplitude $A(f)$ follows the 1/f power law, the plot of $\log A(f)$ against $\log f$ should yield a **straight line**.

Natural images tend to have spatial structures at different scales, leading to a roughly scale-invariant distribution of frequency components, resulting in the **1/f behavior**.

As shown in Figure 6, all four natural images exhibit a linear relationship, confirming the observation of the 1/f power law.

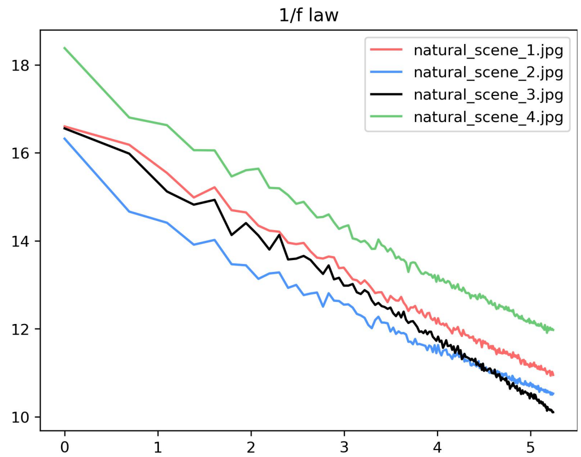

Figure 6: 1/f power law observation.

2.2 Subproblem 2

Question: Compute the integration (summation in discrete case) of $S(f_0) = \int_\Omega A^2(\xi, \eta) d\xi d\eta$ over the domain

$$\Omega(f_0) = \{(\xi, \eta) : f_0 \leq \sqrt{\xi^2 + \eta^2} \leq 2f_0\}$$

Plot $S(f_0)$ over f_0, the plot should fit to a horizontal line (with fluctuation) as $S(f_0)$ is supposed to be a constant over f_0.

Answer:

$$S(f_0) = \int_{\Omega(f_0)} A^2(\xi, \eta)\, d\xi\, d\eta$$

where $\Omega(f_0) = \{(\xi, \eta) : f_0 \leq \sqrt{\xi^2 + \eta^2} \leq 2f_0\}$.

First, in polar coordinates, let $r = \sqrt{\xi^2 + \eta^2}$ and θ be the angle. The Jacobian for the coordinate change is r, so the integral becomes:

$$S(f_0) = \int_{f_0}^{2f_0} \int_0^{2\pi} A^2(r, \theta)\, r\, d\theta\, dr$$

Assume $A(r, \theta) \approx A(r)$, so the integral simplifies to:

$$S(f_0) = 2\pi \int_{f_0}^{2f_0} A^2(r)\, r\, dr$$

Assume $A(r) \propto \frac{1}{r^\alpha}$, then:

$$S(f_0) = 2\pi \int_{f_0}^{2f_0} \frac{1}{r^{2\alpha}} r\, dr = 2\pi \int_{f_0}^{2f_0} \frac{1}{r^{2\alpha - 1}} dr$$

Compute the integral $\frac{1}{r^{2\alpha-1}}$:

$$\int \frac{1}{r^{2\alpha-1}}\, dr = \frac{r^{2(1-\alpha)}}{2(1-\alpha)}$$

Substitute the limits of integration and the final result is :

$$S(f_0) = \frac{\pi}{1-\alpha} f_0^{2(1-\alpha)} \left(2^{2(1-\alpha)} - 1\right)$$

For values of α close to 1 (as is typical in natural images), the term $2^{2(1-\alpha)} - 1$ becomes small, which explains why $S(f_0)$ tends to remain relatively constant across different f_0 values.

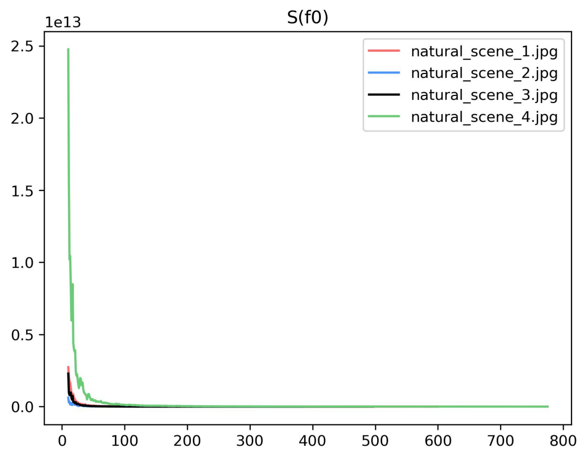

Figure 7: $S(f_0)$ over f_0.

As shown in Figure 7, the plot of $S(f_0)$ against f_0 appears as a horizontal line with minor fluctuations, reflecting the expected behavior that $S(f_0)$ remains stable over f_0. This occurs because natural images generally follow a scale-invariant distribution of frequencies, with their Fourier spectra displaying consistent power across frequency bands.

3 Problem 3

(A 2D scale invariant world, 3 points). *Please read through this question description before you start.*

Let's consider the simulation of a 2D world in which the images consist of only 1D line segments. Each line segment in an image can be characterized by its staring point (x_i, y_i), orientation θ_i, and length r_i. The line segments are independently distributed with uniform probability for their centers and orientations. The length of the line segments follows a probability distribution denoted as $p(r)$, which is proportional to $1/r^3$, representing a cubic power-law distribution. [Hint: How to sample r from $p(r)$? Calculate the Cumulative Distribution function of $p(r)$, then draw a random number in [0,1].]

If you did it right [Please try!], the 6 images must look the same (i.e., you should not be able to tell what scale the 6 images are cropped from). As a result, this 2D world is scale-invariant.

3.1 Subproblem 1

Question: Simulate 1 image I_1 of size 1024×1024 pixels with a total N lines. (You need to truncate long lines and hide (discard) lines shorter than a pixel.)

Answer: As shown in Figure 8, we set $N = 10000$ and simulate 1 image I_1 of size 1024×1024 pixels with a total N lines.

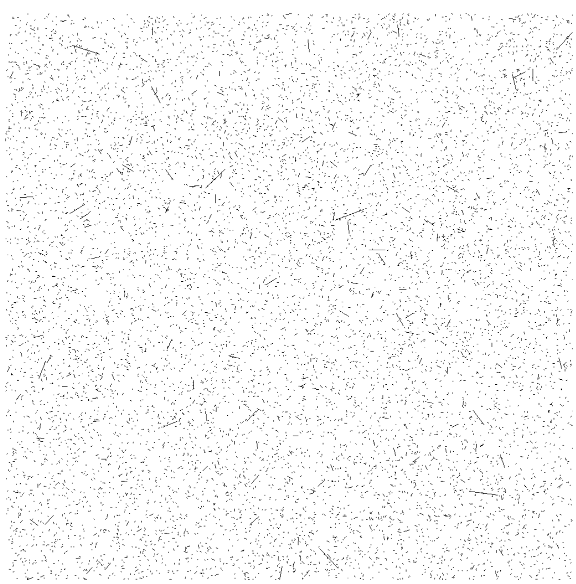

Figure 8: Simulated Image 1024×1024.

Question: Simulate 2 new images I_2 and I_3 of size 512×512 and 256×256 pixels respectively. I_2 and I_3 are down-sampled version of I_1 and are generated by shortening the N line segments in I_1 by 50% and 25% respectively (discard lines shorter than 1).

Answer: We simulate 2 new images I_2 and I_3 of size 512×512 and 256×256 pixels respectively. We set $N = 10000$.

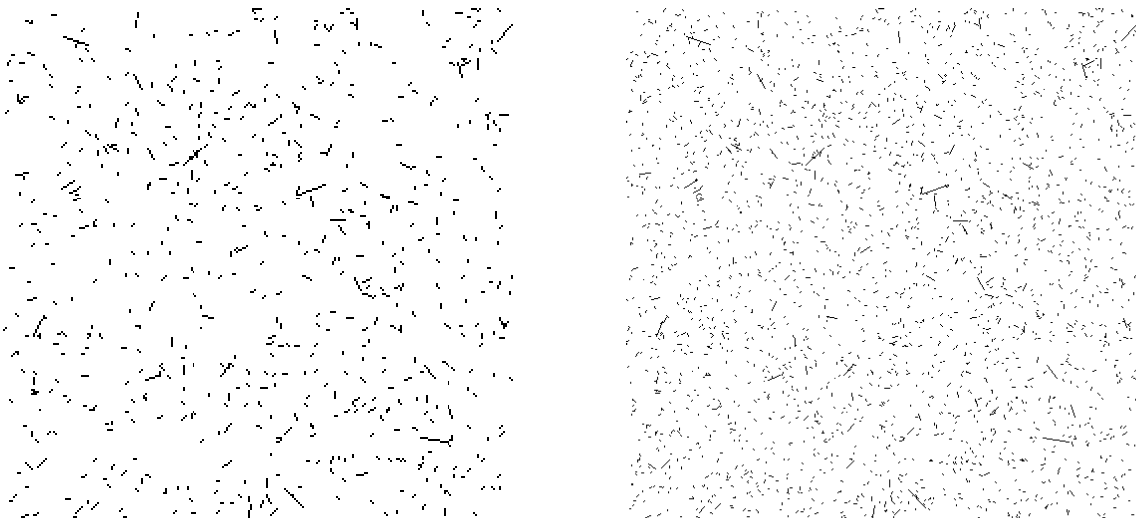

Figure 9: **Left:** Simulated Image 256×256. **Right:** Simulated Image 512×512.

Question: Crop 2 image patches of size 128×128 pixels randomly from each of the three images I_1, I_2, I_3 respectively. Plot these six images [draw the line segments in black on white background].

Answer: We randomly cropped two image patches of size 128×128 pixels from each of the three images I_1, I_2, and I_3, as required by the task. As shown in Figure 10, the six images appear visually identical (cannot tell what scale the 6 images are cropped from), demonstrating the scale-invariant nature of this 2D world.

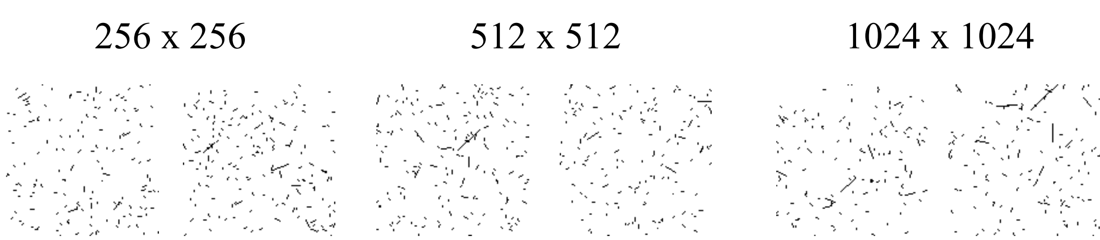

Figure 10: Six Cropped Images.