REPORT

SOIL CLASSIFICATION AND MOISTURE INSPECTION

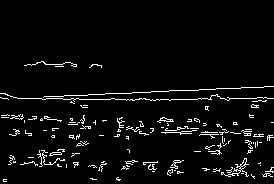
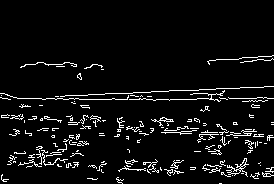
# DATASET

* 4 types of soil images are used to generate the dataset .
* Every type of soil (Red,Alluvial,Black,Clay) has about 400 images , and a total 1500 images are used .
* Converted the image to grayscale.



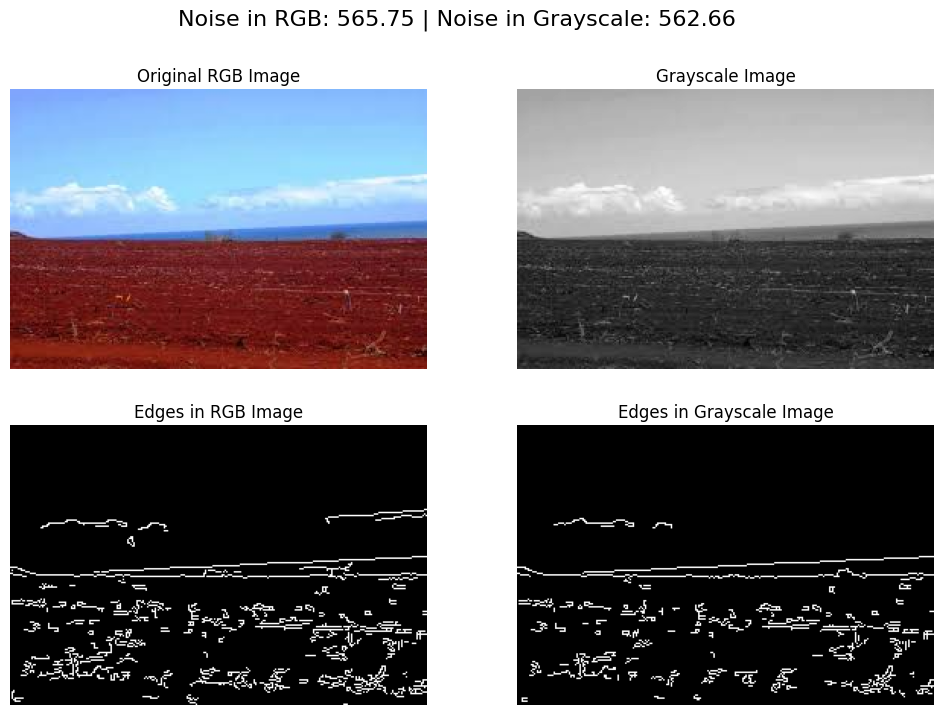
# IMAGE PROCESSING

* Edge Detection (Noise) using Laplacian operator



RGB IMAGE -565.75284 Gray -562.6632

* Visual Comparison –



* Applied 4 different method for noise reduction i.e (Gaussian Blur ,Median filtering ,Bilateral filtering , Non local means denoising)

## NOISE REDUCTION TECHNIQUE

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Four noise reduction techniques were implemented:

1. Gaussian Blur
2. Median Filter
3. Bilateral Filter
4. Non-Local Means Denoising
5. **Gaussian Blur**

A Gaussian Blur filter uses a Gaussian function to smooth images, reducing noise by averaging pixel values with their neighbors.

1. **Median Filter**

The Median Filter replaces each pixel's value with the median value of the intensities in its neighborhood, effectively reducing salt-and-pepper noise.

1. **Bilateral Filter**

The Bilateral Filter smooths images while preserving edges by combining domain and range filtering.

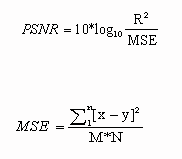
1. **Non-Local Means Denoising**

This technique reduces noise by averaging all pixels in an image weighted by their similarity to the target pixel.

### EVALUATION METRICS

**Peak Signal-to-Noise Ratio (PSNR)**

PSNR is a metric used to measure the quality of a reconstructed image compared to its original form. It is expressed in decibels (dB) and is based on the Mean Squared Error (MSE) between the original and reconstructed images. The formula for PSNR is:



where R is the maximum possible pixel value of the image (typically 255 for an 8-bit image) and MSE is the Mean Squared Error calculated as:

Where M and N are the dimensions of the images.

**Interpretation**

High PSNR----- Indicates that the reconstructed image is very close to the original image, implying high quality. Typical values for PSNR in image and video compression applications range between 20 dB to 50 dB.

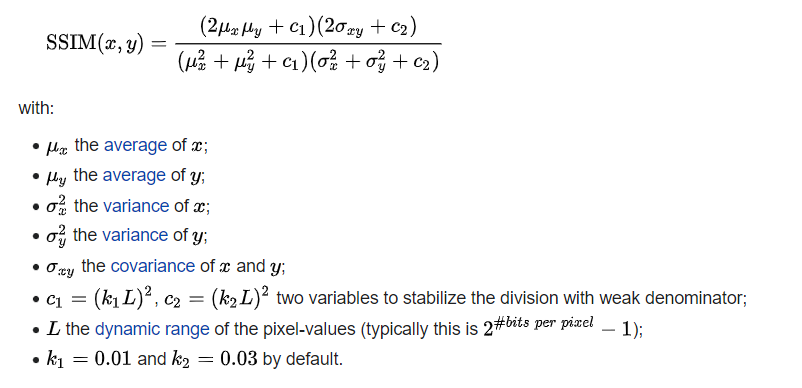
Low PSNR----- Indicates that the reconstructed image has a high level of distortion compared to the original image, implying lower quality.

Conclusions from PSNR Values----

* A higher PSNR value generally means better image quality.
* PSNR values above 30 dB are usually considered acceptable in terms of image quality.
* PSNR values below 20 dB typically indicate significant degradation in image quality.

**Structural Similarity Index (SSIM)**

SSIM is a perceptual metric that quantifies the image quality degradation caused by processing such as data compression or transmission losses. Unlike PSNR, which is based on pixel differences, SSIM considers changes in structural information, luminance, and contrast. The SSIM index is calculated as:



SSIM values range from -1 to 1:

* 1: Perfect similarity between the original and reconstructed images.
* 0: No similarity (completely different images).
* Negative values: Generally indicate that the images are negatively correlated, which is uncommon in practice.

Conclusions from SSIM Values:

* Higher SSIM values: Closer to 1 indicate better image quality and preservation of structural information.
* Lower SSIM values: Indicate poorer image quality with significant structural distortion.
* SSIM values above 0.9 are generally considered very good, indicating that the processed image is almost indistinguishable from the original.

**Conclusion Based on PSNR and SSIM**

When comparing different noise reduction techniques, both PSNR and SSIM are important:

1. PSNR

- Helps quantify the overall fidelity of the reconstructed image.

- A higher PSNR value indicates lower error and higher similarity to the original image.

2. SSIM:

- Provides a perceptual measure of image quality by considering structural changes, luminance, and contrast.

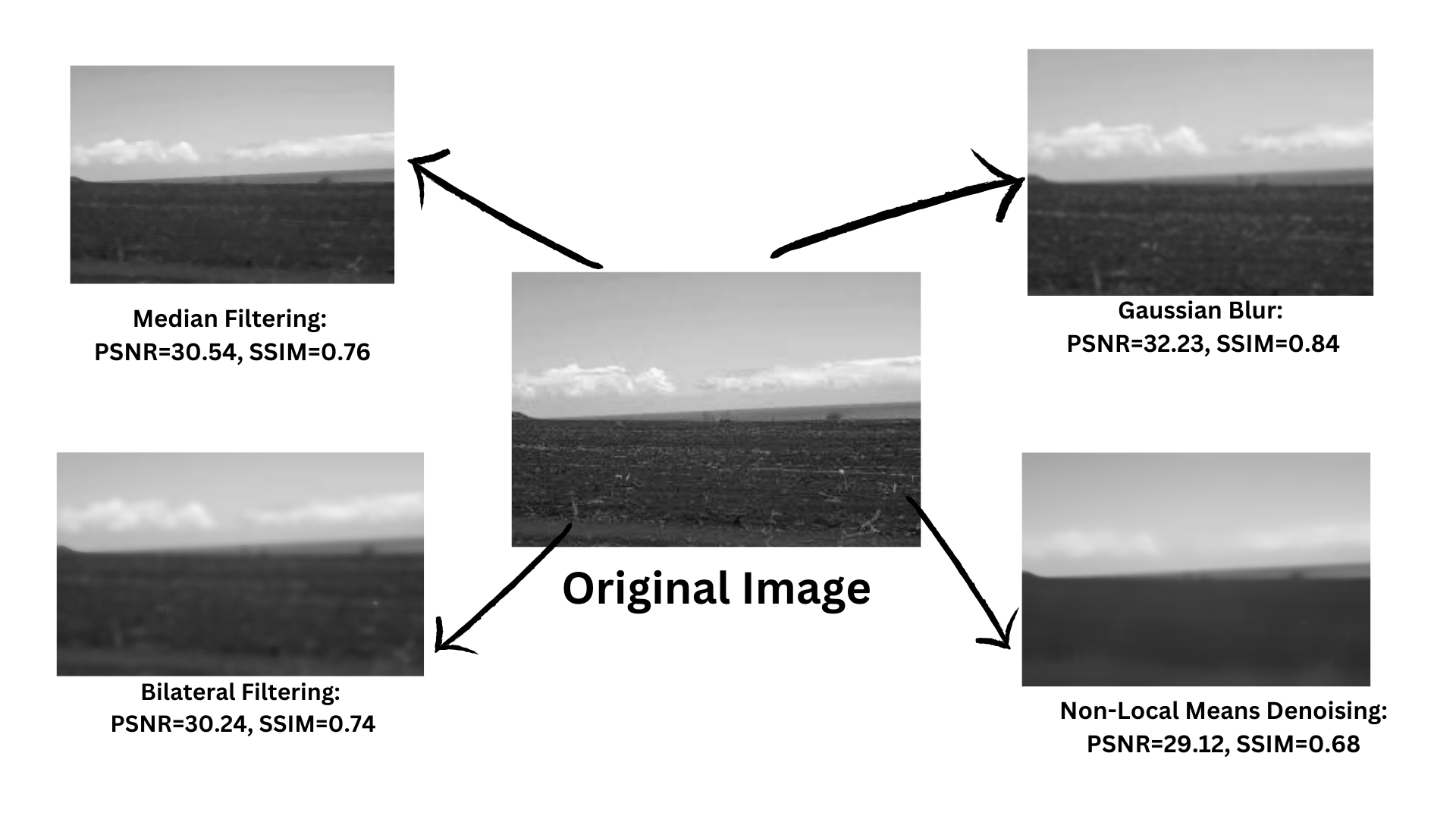
- A higher SSIM value indicates that the structural integrity of the image is well preserved.

**Overall Conclusion:**

- Techniques with higher PSNR and SSIM values are preferred as they indicate better preservation of image quality.

- Comparing PSNR and SSIM together provides a more comprehensive evaluation of the noise reduction techniques, balancing both pixel-wise fidelity (PSNR) and perceptual quality (SSIM).

By using both metrics, We can make more informed decisions about which noise reduction technique is most effective for our specific application.



| **Noise Reduction Technique** | **PSNR Value** | **SSIM Value** |
| --- | --- | --- |
| Gaussian Blur | 32.23 | 0.84 |
| Median Filtering | 30.54 | 0.76 |
| Non-Local Means Denoising | 29.12 | 0.68 |
| Bilateral Filtering | 30.24 | 0.74 |

## BEST TECHNIQUE AND PARAMETER

------------------------------------------------------------------------------

We evaluated different noise reduction techniques applied to grayscale images and identified the best performing method and its optimal parameters. Four denoising techniques were compared: Non-Local Means (NLM), Gaussian, Median, and Bilateral filtering. Each technique was tested with various parameter sets, and their performance was assessed using Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM).

### METHOD:

**Non-Local Means (NLM) Denoising**:

* Parameters: (h, hForColor, templateWindowSize)
* Tested sets: (5, 3, 15), (10, 3, 20), (15, 5, 25), (20, 7, 30), (25, 10, 35), (30, 15, 40)

**Gaussian Denoising**:

* Parameters: (kernel\_size, sigma)
* Tested sets: (3, 1), (5, 1.5), (7, 2), (9, 2.5), (11, 3), (13, 3.5)

**Median Denoising**:

* Parameters: (kernel\_size)
* Tested sets: (3), (5), (7), (9), (11), (13)

**Bilateral Filtering**:

* Parameters: (diameter, sigmaColor, sigmaSpace)
* Tested sets: (5, 25, 25), (7, 50, 50), (9, 75, 75), (11, 100, 100), (13, 125, 125), (15, 150, 150)

### RESULT:

The performance of each technique was evaluated by calculating the average PSNR and SSIM values for each parameter set across the dataset of 1500 images. The results are summarized in the table below:

| **Technique** | **Parameters** | **Average PSNR** | **Average SSIM** |
| --- | --- | --- | --- |
| nlm | (5, 3, 15) | 47.1912 | 0.9950 |
| nlm | (10, 3, 20) | 33.7452 | 0.9560 |
| nlm | (15, 5, 25) | 28.4377 | 0.8611 |
| nlm | (20, 7, 30) | 25.3884 | 0.7537 |
| nlm | (25, 10, 35) | 23.4362 | 0.6429 |
| nlm | (30, 15, 40) | 22.2458 | 0.5529 |
| gaussian | (3, 1) | 27.1733 | 0.8764 |
| gaussian | (5, 1.5) | 24.9362 | 0.7899 |
| gaussian | (7, 2) | 23.8952 | 0.7310 |
| gaussian | (9, 2.5) | 23.2358 | 0.6854 |
| gaussian | (11, 3) | 22.7567 | 0.6476 |
| gaussian | (13, 3.5) | 22.3965 | 0.6163 |
| median | (3) | 26.6077 | 0.8488 |
| median | (5) | 24.4904 | 0.7558 |
| median | (7) | 23.5003 | 0.6936 |
| median | (9) | 22.8851 | 0.6482 |
| median | (11) | 22.4540 | 0.6114 |
| median | (13) | 22.1353 | 0.5818 |
| bilateral | (5, 25, 25) | 30.9738 | 0.9420 |
| bilateral | (7, 50, 50) | 26.5826 | 0.8409 |
| bilateral | (9, 75, 75) | 24.6908 | 0.7608 |
| bilateral | (11, 100, 100) | 23.4941 | 0.6911 |
| bilateral | (13, 125, 125) | 22.8132 | 0.6429 |
| bilateral | (15, 150, 150) | 22.3481 | 0.6051 |

### CONCLUSION:

The results indicate that the **Non-Local Means (NLM) denoising technique with parameters (5, 3, 15)** achieves the highest average PSNR of 47.1912 and the highest average SSIM of 0.9950. This suggests that NLM is the most effective denoising method for the given dataset, providing superior noise reduction while maintaining the structural integrity of the images.

#### **Findings**

1. **Best Denoising Technique**: Non-Local Means (NLM)
2. **Best Parameters**: (5, 3, 15)
3. **Average PSNR**: 47.1912
4. **Average SSIM**: 0.9950

# 

# 

# 

# 

# 

# 

# ISOLATING SOIL REGION

--------------------------------------------------------

STEP 1: Convert to Grayscale

* The input image is first converted to grayscale to simplify further processing

STEP 2: Apply Thresholding

* A binary threshold is applied to the grayscale image to highlight the soil region. In this process, pixel values below 127 are set to 255 (white), and those above 127 are set to 0 (black).

STEP 3: Find Contours

* Contours are detected in the binary image. Contours are curves joining all the continuous points along the boundary of a shape, having the same color or intensity.

STEP 4: Get the Largest Contour

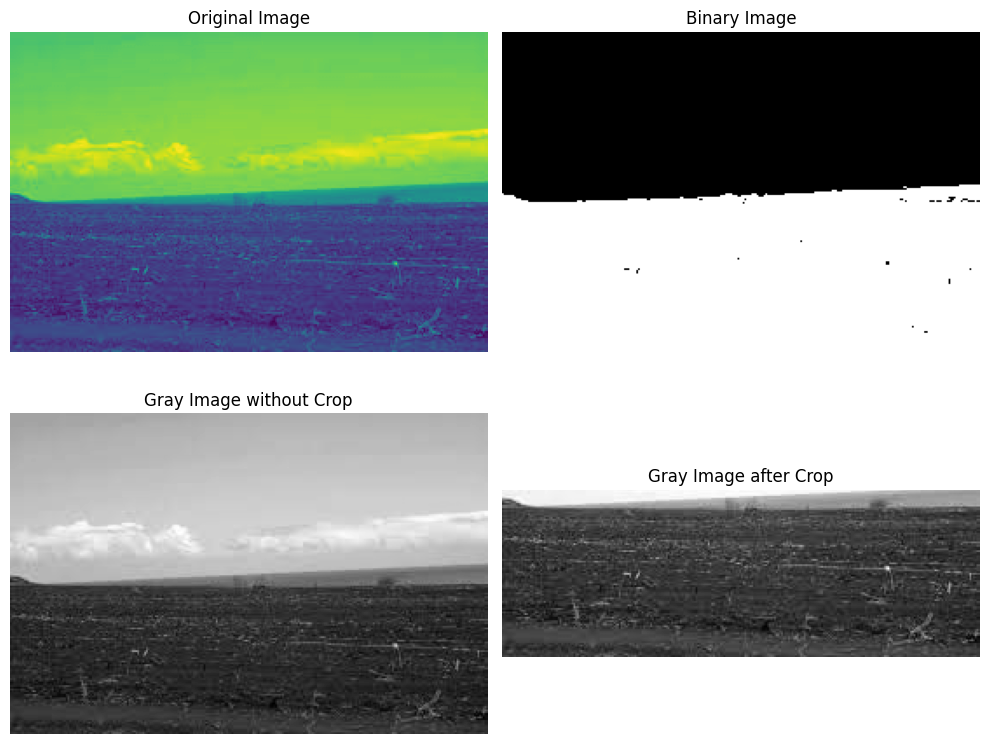
* Among the detected contours, the largest contour is identified. This contour is assumed to represent the soil region.

STEP 5: Get the Bounding Box of the Largest Contour

* The bounding box for the largest contour is calculated. This box will be used to crop the image.

STEP 6: Crop the Image Using the Bounding Box

* The original image is cropped using the coordinates of the bounding box obtained in the previous step.



GLCM PROPERTIES COMPARISON

—---------------------------------------------------

| GLCM Properties | Cropped Image | Original Gray Image |
| --- | --- | --- |
| Contrast | 95.91662851037852 | 51.69798136645962 |
| Correlation | 0.9639465677035237 | 0.9947484702526976 |
| Energy | 0.030726996120950815 | 0.0636933913938309 |
| Homogeneity | 0.24433910162196476 | 0.524726349301551 |

CONCLUSION

--------------------

The comparative analysis of GLCM properties between the cropped and original grayscale images provides valuable insights into the effectiveness of our preprocessing:

* **Higher Contrast in Cropped Image:** Indicates increased detail and potential noise within the region of interest.
* **Slightly Lower Correlation:** Suggests a minor loss in the spatial relationship between pixels in the cropped image.
* **Lower Energy and Homogeneity:** Implies the presence of more variations and potential residual noise in the cropped image compared to the original.

Overall, while the preprocessing techniques have successfully isolated the region of interest, there is room for improvement in noise reduction to enhance the uniformity and homogeneity of the cropped images. Future work is focused on refining noise reduction methods to achieve a better balance between detail preservation and noise elimination, ultimately improving the quality of the region of interest in the images.

# GLCM PROPERTIES

---------------------------------------------------

## Definition of GLCM Properties

1. **Contrast**:

-**Definition**: Contrast is a measure of the intensity contrast between a pixel and its neighbor over the whole image. It is calculated as the sum of squared differences between gray levels.

- **Application**: High contrast values indicate significant variations in texture, which can help identify soils with rough or uneven surfaces. Low contrast values suggest smoother textures.

2. **Correlation**:

- **Definition**: Correlation measures how correlated a pixel is to its neighbor over the whole image. It calculates the linear dependency of gray levels of neighboring pixels.

- **Application**: High correlation values indicate that pixel pairs have a predictable relationship, which can help identify homogeneous soil types. Low correlation values suggest more randomness in texture.

3. **Energy**:

- **Definition**: Energy, also known as Angular Second Moment, measures the sum of squared elements in the GLCM. It indicates textural uniformity.

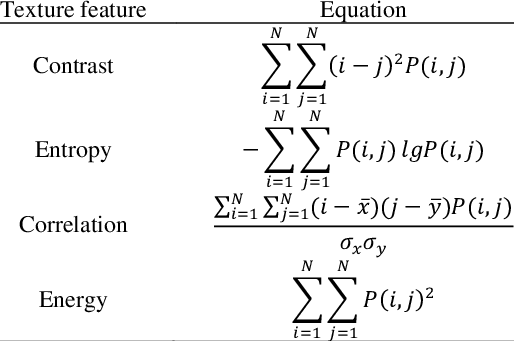
- **Application**: High energy values suggest that the image has a uniform texture, which can be helpful in identifying soils with consistent particle size. Low energy values indicate more complex textures.

4. **Homogeneity**:

- **Definition**: Homogeneity measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal. It evaluates how similar each pixel is to its neighbor.

- **Application**: High homogeneity values indicate that pixels are similar to their neighbors, which can be useful for identifying fine-grained soils. Low homogeneity values suggest more coarse or heterogeneous textures.

## FORMULA :



## Application of GLCM Properties in Soil Classification

The GLCM properties are essential for classifying soil types based on their textural characteristics. Here are the applications and benefits of using these properties:

1. **Soil Texture Analysis**:

- Different soils (e.g., sandy, clayey, loamy) exhibit distinct texture patterns. By analyzing GLCM properties, we can differentiate between these soil types based on their textural features.

2. **Soil Quality Assessment**:

- The texture of soil affects its properties such as water retention, nutrient availability, and root penetration. GLCM properties provide a quantitative measure of texture, aiding in soil quality assessment.

3. **Agricultural Applications**:

- Farmers and agronomists can use texture analysis to determine the suitability of soil for various crops. Fine-textured soils (high homogeneity and energy) are generally better for water retention, while coarse-textured soils (high contrast) may drain water more quickly.

## Why GLCM Properties are Helpful

1. **Quantitative Analysis**:

- GLCM properties provide a numerical representation of texture, making it easier to compare and classify different soil samples objectively.

2. **Sensitivity to Texture Variations**:

- These properties are sensitive to subtle variations in texture, which might not be visible to the naked eye. This sensitivity is crucial for accurate soil classification.

3. **Complementary Information**:

- Each GLCM property captures different aspects of texture (e.g., contrast measures variability, homogeneity measures uniformity). Using multiple properties together provides a comprehensive analysis of soil texture.

4. **Non-Destructive Testing**:

- The analysis of GLCM properties is a non-destructive method, meaning the soil samples can be preserved for further testing or usage.

In conclusion, GLCM properties are powerful tools for soil classification, providing detailed insights into the texture of soils. Their application spans various fields, including agriculture, environmental monitoring, and remote sensing, making them invaluable for both scientific research and practical soil management.

## EXTRACTED GLCM PROPERTIES

GLCM is a powerful tool for texture analysis, providing statistical measures that describe the texture of an image. The GLCM was computed for each preprocessed image at multiple angles (0°, 45°, 90°, and 135°) and distances to capture diverse texture patterns. The following GLCM properties were extracted:

1. Dissimilarity
2. Correlation
3. Homogeneity
4. Contrast
5. ASM
6. Energy

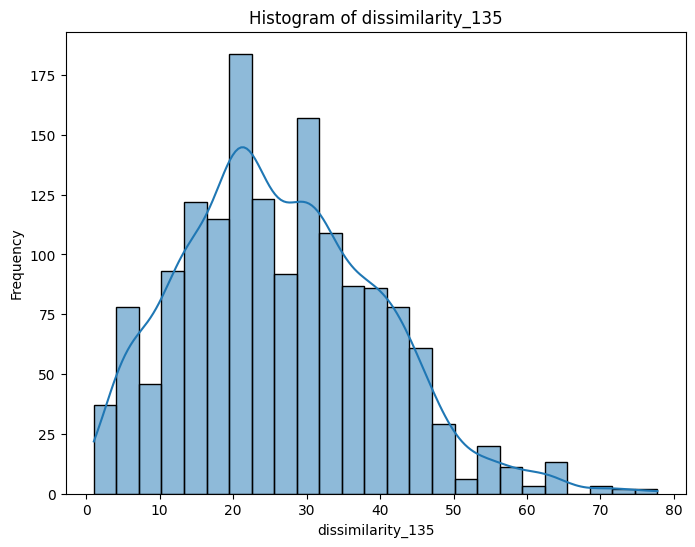
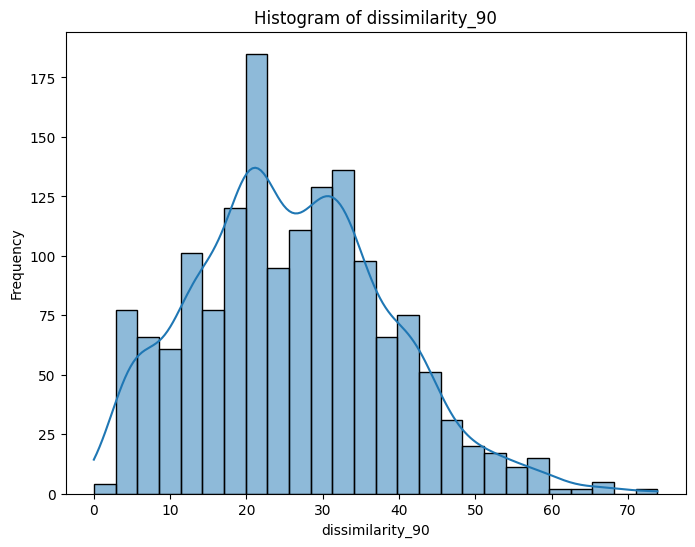
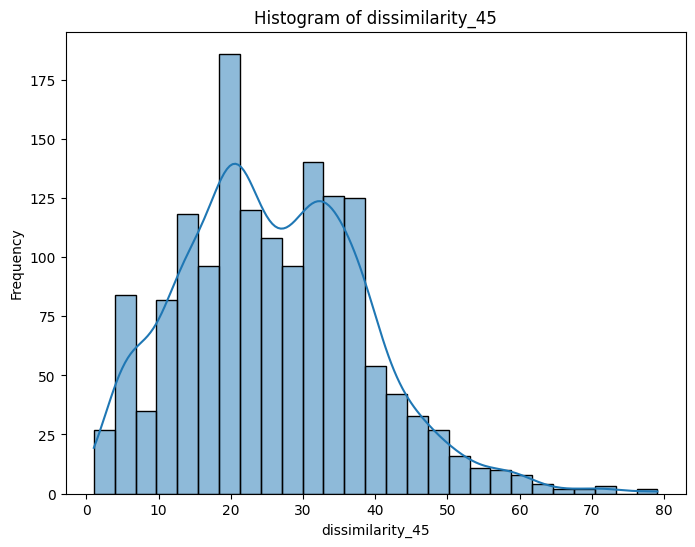
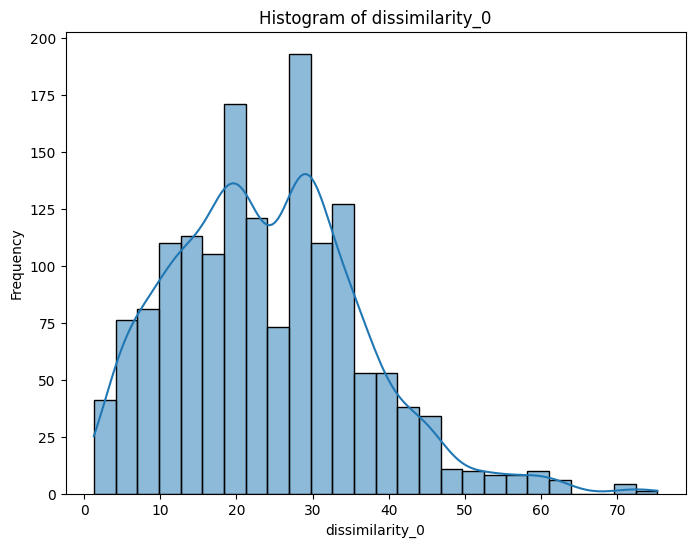
## GLCM DATASET OVERVIEW

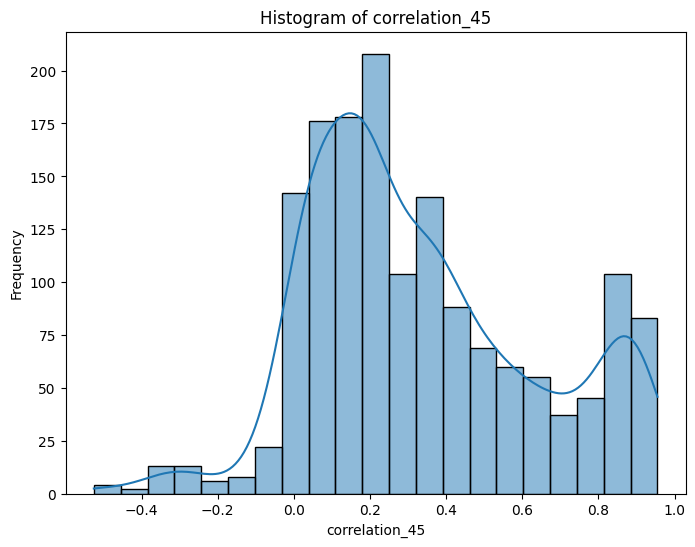
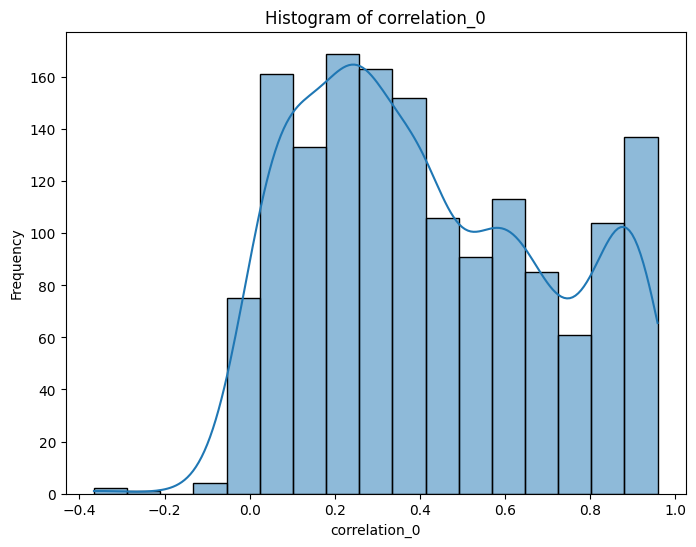
----------------------------------------------------------------

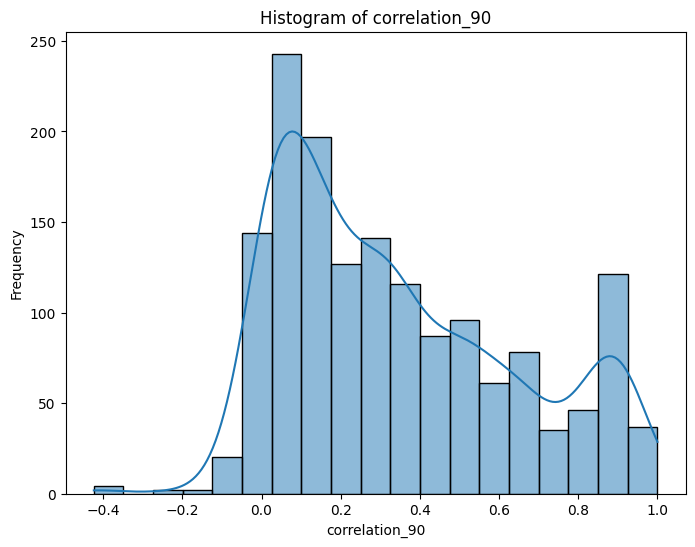
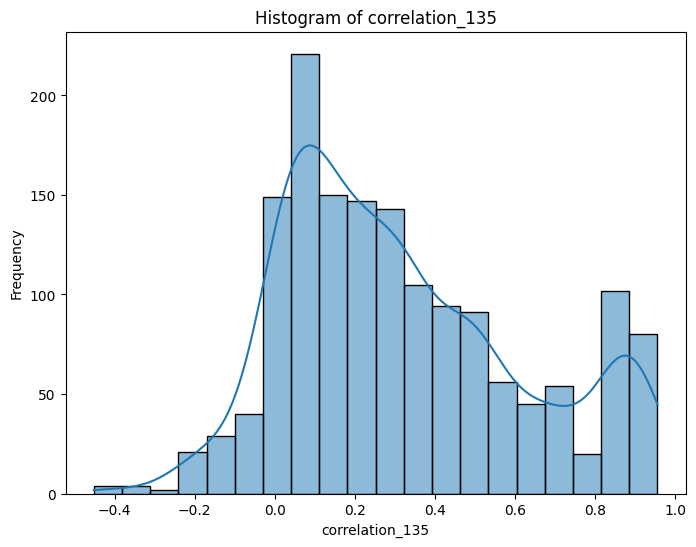
The dataset consists of 1557 entries with 25 columns, where each row represents an image sample of soil. The columns include various GLCM (Gray Level Co-occurrence Matrix) properties calculated at four different angles (0°, 45°, 90°, and 135°), along with a label indicating the soil type.

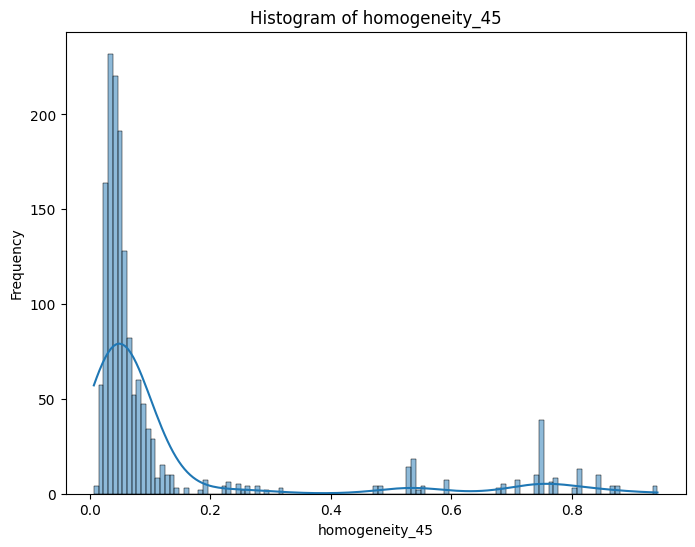
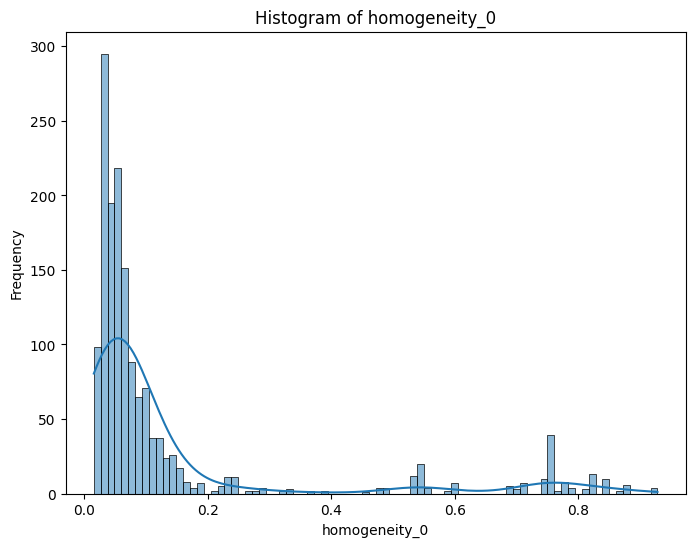
#### **Columns Description**

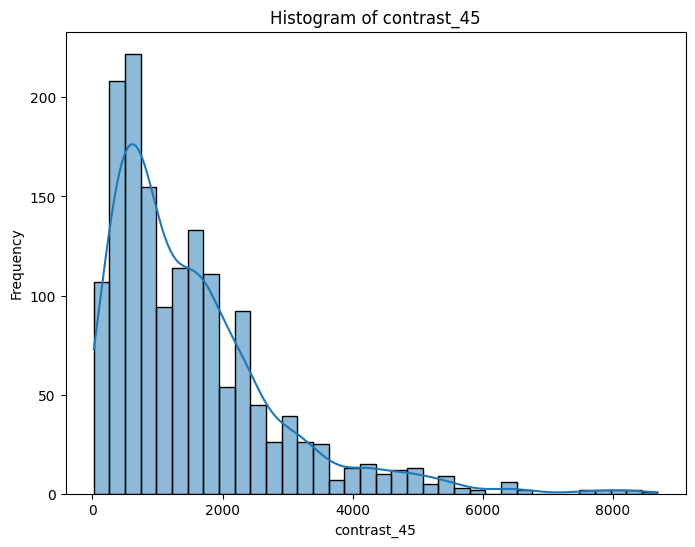
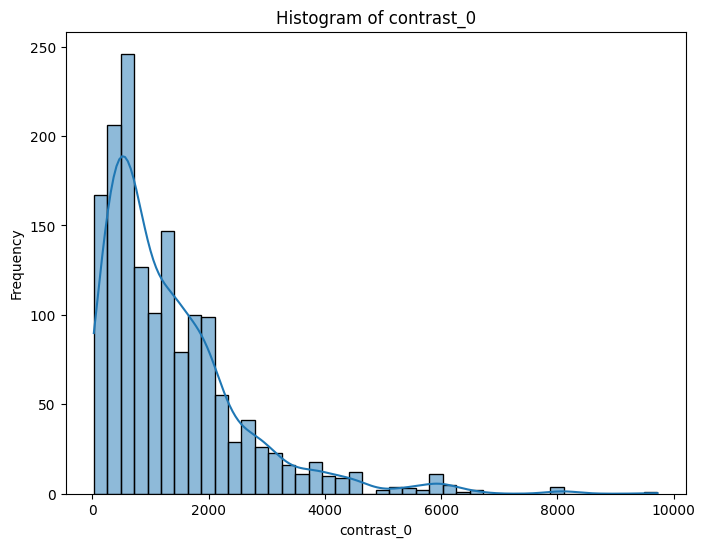
1. **dissimilarity\_0**: Measures the variation of gray-level pairs in the GLCM at 0°.
2. **dissimilarity\_45**: Measures the variation of gray-level pairs in the GLCM at 45°.
3. **dissimilarity\_90**: Measures the variation of gray-level pairs in the GLCM at 90°.
4. **dissimilarity\_135**: Measures the variation of gray-level pairs in the GLCM at 135°.
5. **correlation\_0**: Assesses the correlation of pixels at 0°.
6. **correlation\_45**: Assesses the correlation of pixels at 45°.
7. **correlation\_90**: Assesses the correlation of pixels at 90°.
8. **correlation\_135**: Assesses the correlation of pixels at 135°.
9. **homogeneity\_0**: Evaluates the homogeneity of the GLCM at 0°.
10. **homogeneity\_45**: Evaluates the homogeneity of the GLCM at 45°.
11. **homogeneity\_90**: Evaluates the homogeneity of the GLCM at 90°.
12. **homogeneity\_135**: Evaluates the homogeneity of the GLCM at 135°.
13. **contrast\_0**: Quantifies the contrast in the GLCM at 0°.
14. **contrast\_45**: Quantifies the contrast in the GLCM at 45°.
15. **contrast\_90**: Quantifies the contrast in the GLCM at 90°.
16. **contrast\_135**: Quantifies the contrast in the GLCM at 135°.
17. **ASM\_0**: Measures the Angular Second Moment (ASM) at 0°, indicating uniformity.
18. **ASM\_45**: Measures the Angular Second Moment (ASM) at 45°, indicating uniformity.
19. **ASM\_90**: Measures the Angular Second Moment (ASM) at 90°, indicating uniformity.
20. **ASM\_135**: Measures the Angular Second Moment (ASM) at 135°, indicating uniformity.
21. **energy\_0**: Reflects the sum of squared elements in the GLCM at 0°.
22. **energy\_45**: Reflects the sum of squared elements in the GLCM at 45°.
23. **energy\_90**: Reflects the sum of squared elements in the GLCM at 90°.
24. **energy\_135**: Reflects the sum of squared elements in the GLCM at 135°.
25. **label**: The categorical label indicating the type of soil.

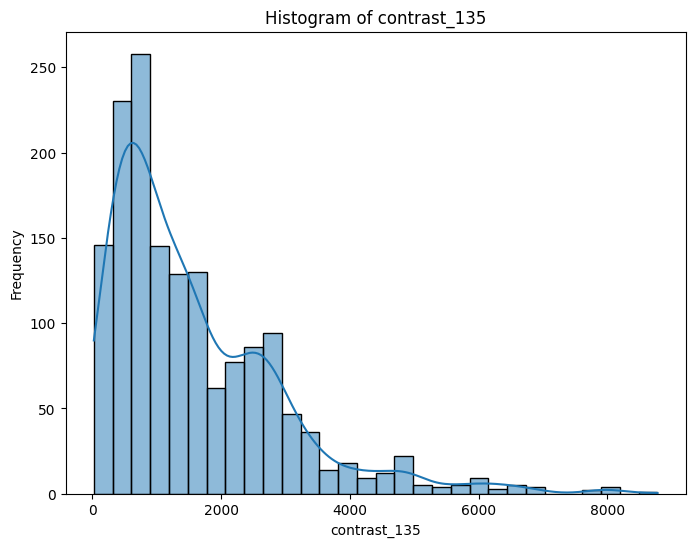
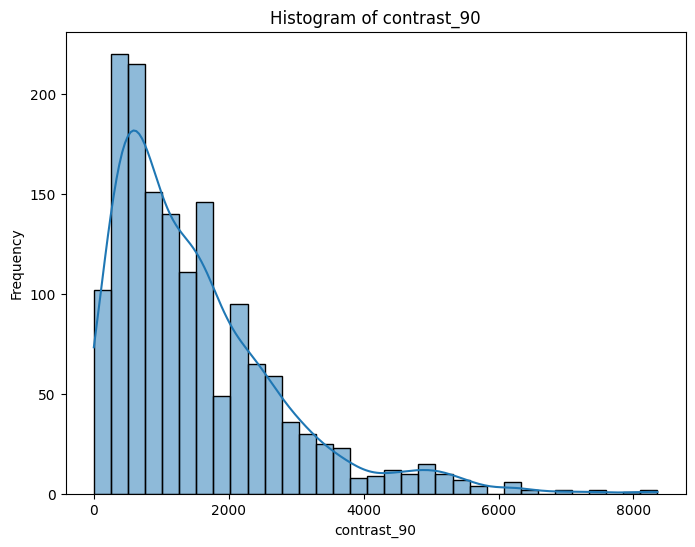


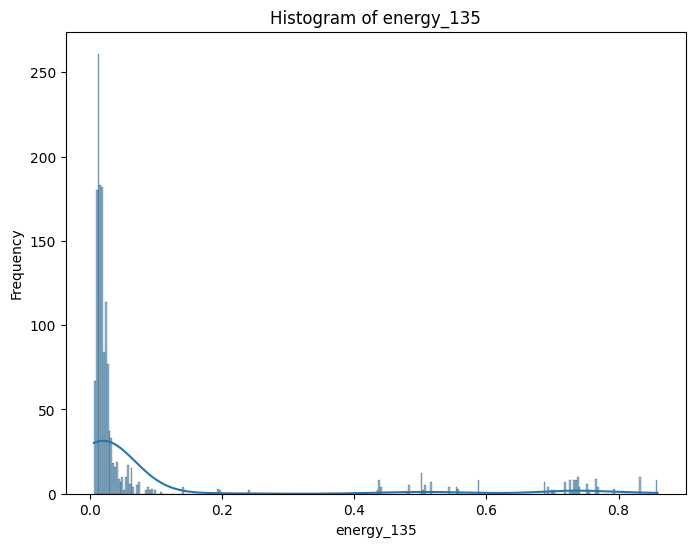
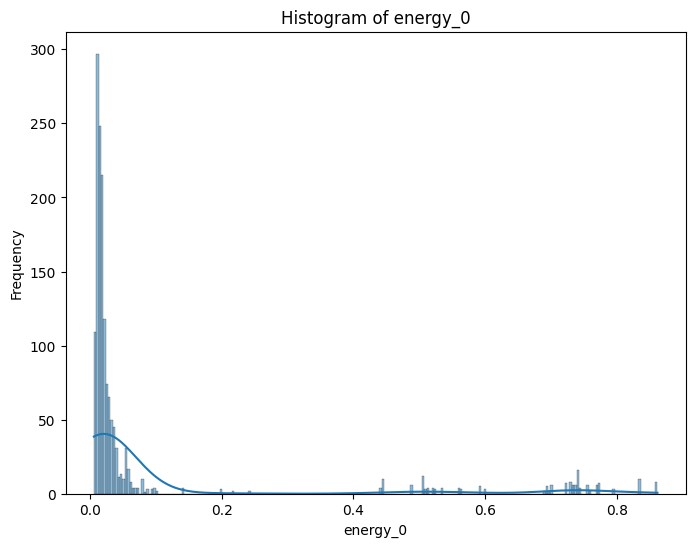












# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# ------------SVM MODEL-------------------------------------

Implemented SVM Model on the processed dataset and got the following findings :