Mini-Project
Fractal Analysis using Differential Box-Counting

Under the guidance of Prof. Prabhat Munshi

Introduction

In the domain of medical imaging, the development of robust algorithms is pivotal for enhancing analysis efficiency and accuracy. This project centers on the creation and implementation of a novel algorithm based on fractal dimension analysis for medical image processing.

Fractal geometry, renowned for its ability to characterize complex structures, serves as the theoretical foundation for our algorithm. By quantifying the fractal dimension of images obtained from CT, PET, and MRI scans, we aim to extract meaningful features indicative of underlying tissue characteristics.

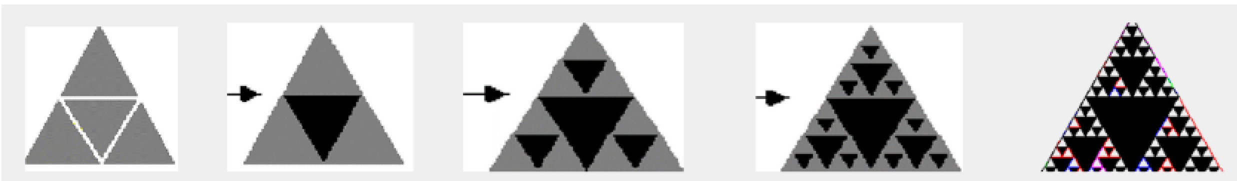
The algorithm's primary objective is to provide a quantitative measure of image complexity, enabling healthcare professionals to assess subtle variations in tissue texture and morphology.

The significance of this algorithm lies in its potential to revolutionize image analysis in medical diagnostics. By automating the extraction of pertinent features from medical images, it offers a valuable tool for clinicians in disease detection, treatment planning, and patient monitoring.

In this report, we elucidate the methodology, implementation details, and experimental results of our fractal dimension-based image analysis algorithm. Furthermore, we discuss the algorithm's implications for the field of medical imaging and its potential for future research and clinical applications.

What is a Fractal?

- Infinite patterns
- Self similar across different scales
- Made by repeating a simple process again and again
- Have fractional dimensions between 2 and 3.



Fractal Dimension

- It is a quantity which measures the complexity of a figure.
- It is a fractional number.
- More FD means a more complex figure.
- Similar FD means similar images.
- Images can be segmented and saved based on fractal dimension thus reducing storage complexity.
- It applies to fluid flow images also
- Many ways to estimate , we used Box Counting.

Algorithm Reference

We have used majorly the paper published by BB Chaudhari and Nirupam Sarkar to devise our algorithm. Our Algorithm is majorly based on the “Differential Box Counting Algorithm” proposed by them.

The Differential Box counting Algorithm:

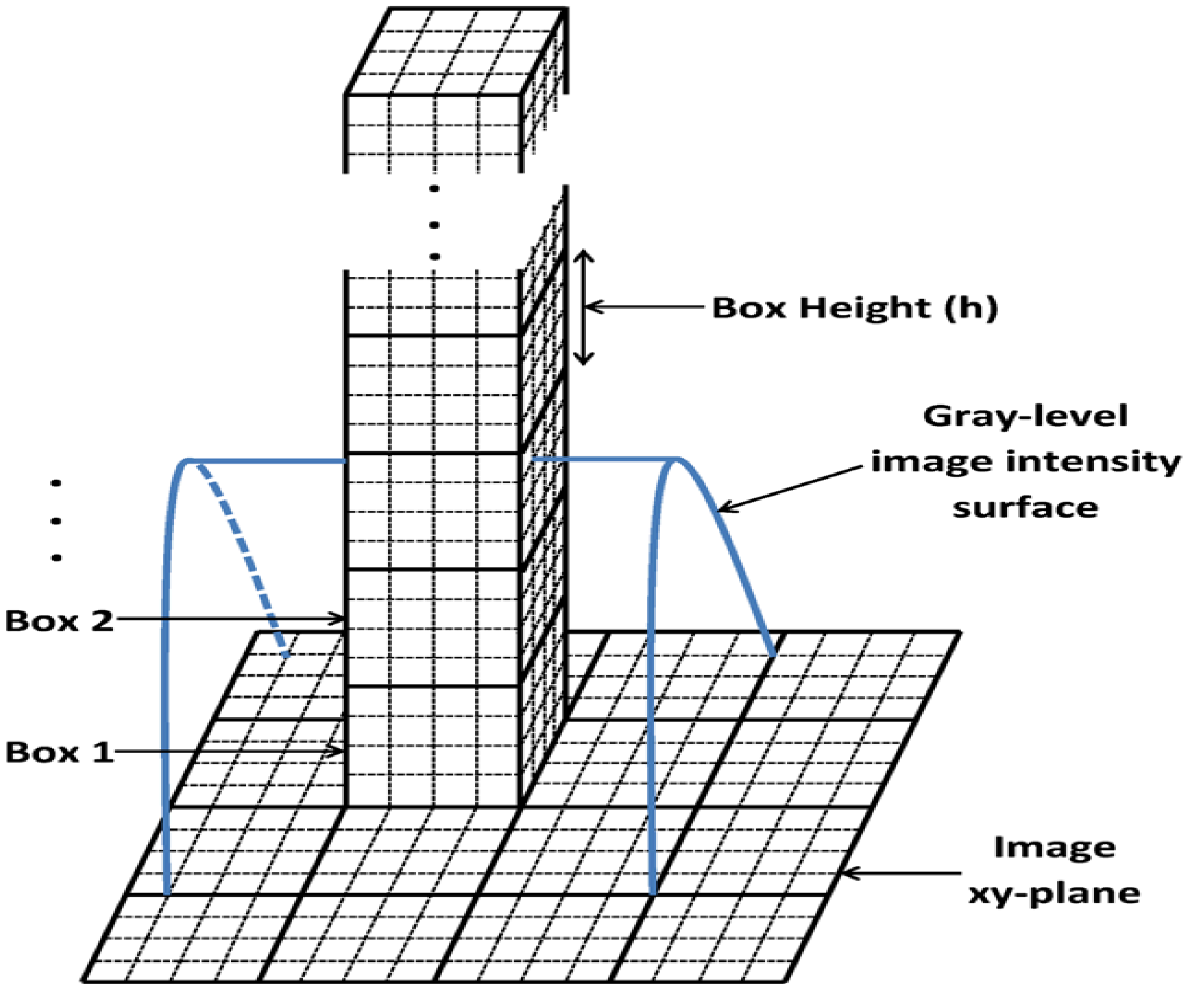
In 1992, Sarkar et al. proposed a straightforward approach for estimating the fractal dimension (FD) of a grayscale image. Their methodology involves partitioning a square image (I) of size $M \times M$ pixels into non-overlapping grids of size $s \times s$ pixels, where s is an integer ranging from 2 to M/2. The scale (r) of a grid with size $s \times s$ pixels relative to the image size M is defined as $r = s/M$. If s is not a divisor of M, then the non-image pixels on the boundary of the grids are treated as zero.

Within each grid, multiple boxes of size $s \times s \times h$ are utilized to cover the rough gray-level image intensity surface. These boxes are assigned numerical values according to a specific scheme. Here, $\lfloor . \rfloor$ denotes the floor function, and it is used to determine the number of boxes needed to cover the intensity surface, denoted by

$$\lfloor G/h \rfloor = \lfloor M/s \rfloor,$$

which implies $h = s \times G/M$.

This approach provides a systematic framework for computing the fractal dimension of an image, enabling the quantification of its complexity and texture. By iteratively adjusting the grid size and analyzing the distribution of intensity values within each grid, Sarkar et al.'s method facilitates the estimation of the fractal dimension, offering valuable insights into the structural properties of the image.



Reference: https://www.mdpi.com/1099-4300/19/10/534

The G is the total number of gray-levels in a gray-scale image i.e., 256. The total number of boxes of size $s \times s \times h$, required to represent the rough surface over a (i, j)th grid of size $s \times s$ pixels, nr (i,j) is

$$nr(i,j) = \lceil gmax/h \rceil - \lceil gmin/h \rceil + 1 \quad (3)$$

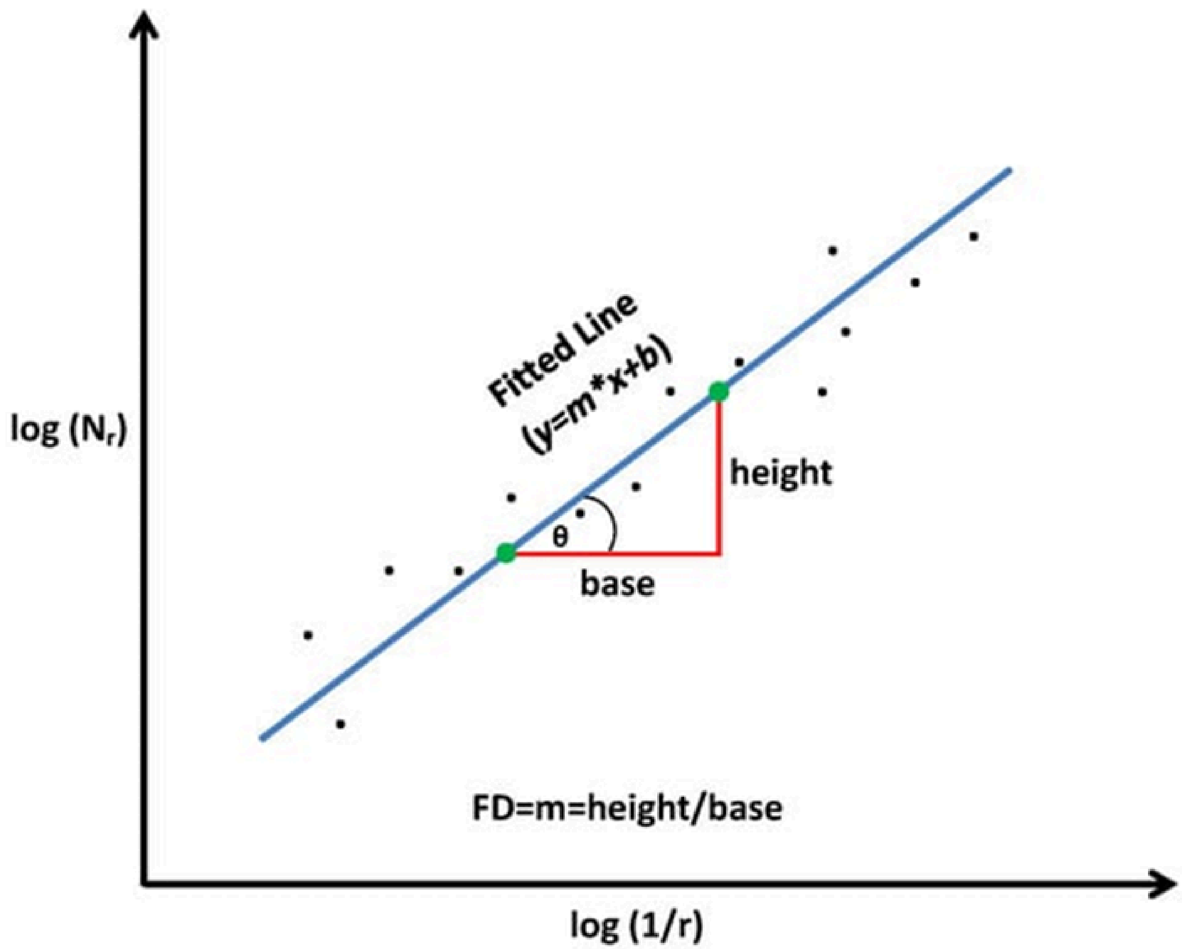
where, ⌈.⌉ denotes the ceiling function.

The $gmax$ and $gmin$ are the maximum and minimum gray-levels present on (i, j)th grid respectively. The $\lceil gmax/h \rceil$ represents the box number, which contains $gmax$. On the other hand, $gmin$ located in a box, which is denoted by box number $\lceil gmin/h \rceil$. The total number of boxes, Nr , at scale r is required to cover the rough intensity surface of an image can be computed by Equation

$$Nr=\sum nr(i,j)$$

where, $2M \le r \le 12$ (4)

Then the FD of an image or the slope of a line is computed by fitting all the points $(1r,Nr)$ using Linear Least Squares as shown



Reference: https://www.mdpi.com/1099-4300/19/10/534

Validation using Brodatz dataset:

We experimented with our Algorithm on some sample images. All the experiments have been performed on one publicly available Brodatz Database. The Brodatz Database contains 112 textured images of size 640×640 pixels. This database is used to propose a novel box height for calculating FD values because this database is very much popular and used by most of the researchers in this context.

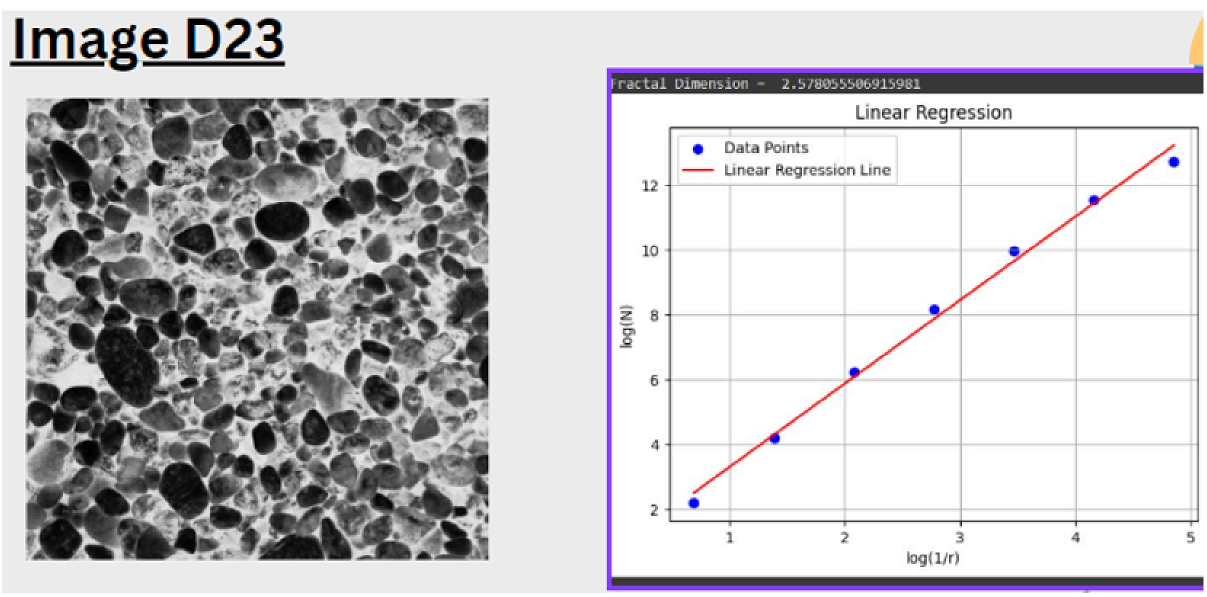
So far, We have experimented our algorithm on 25 images from Broadtz dataset + 4 images of multiphase flow.

This is the link for the colab notebook for reference:

Box_counting_BBC.ipynb

Results

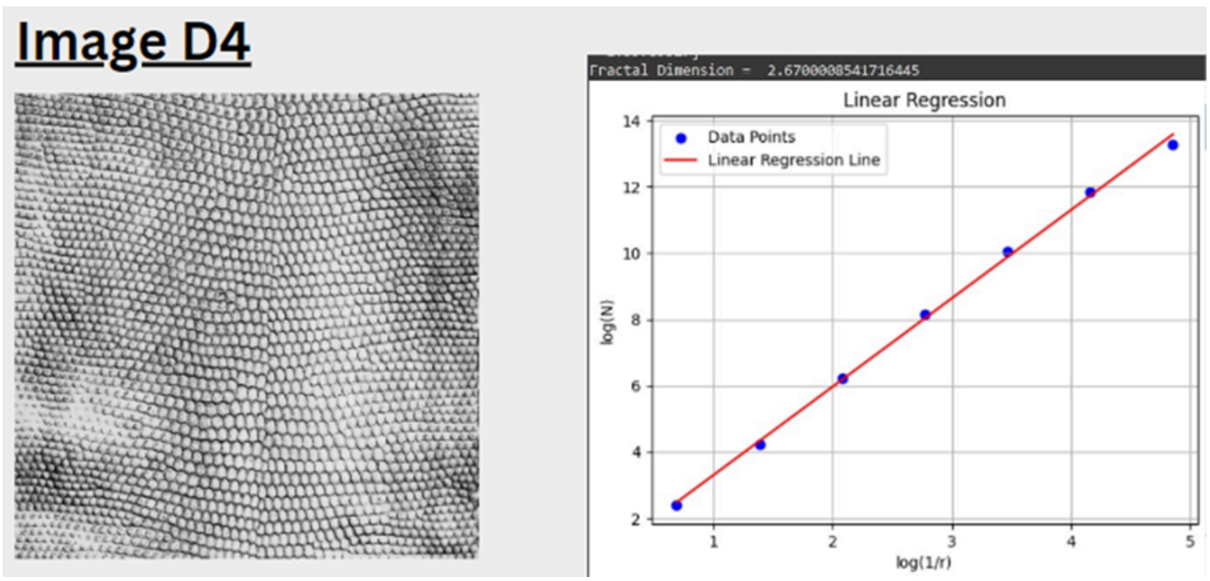
Here are a few images and the graph generated by our algorithm for your reference.



Original FD: 2.621

Calculated FD: 2.578

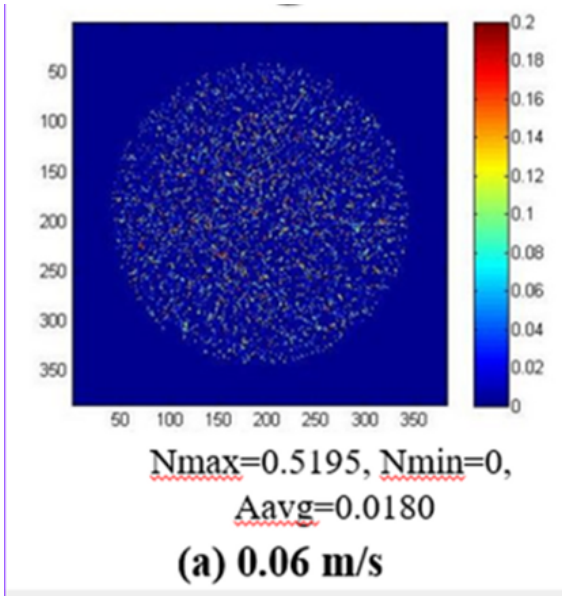
Abs error: 1.6406 %

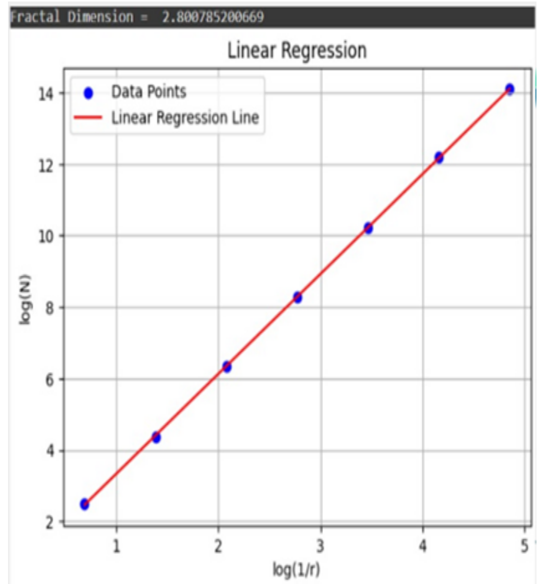


Original FD: 2.67

Calculated FD: 2.670001

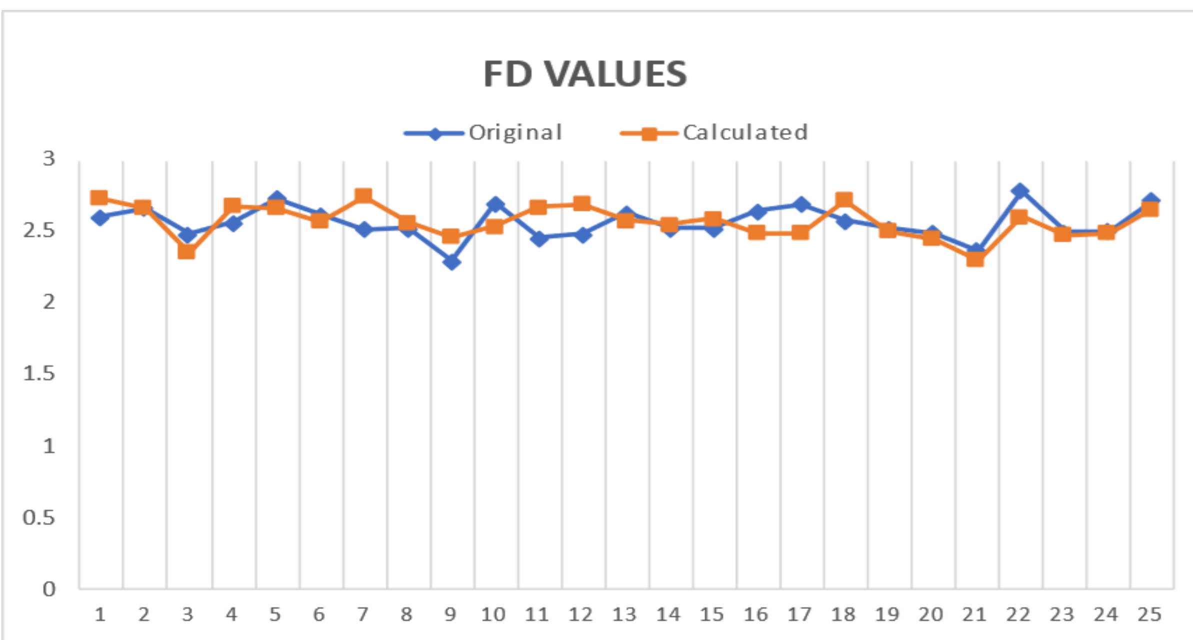
Abs error: 0.0000375 %





Calculated FD: 2.8007

Here is the graph between original values and values of FD calculated by our algorithm. It is based on 25 images of the Broadtz dataset.



Here is the CSV of our results :

https://docs.google.com/spreadsheets/d/1MYCr4n0EcmILWbQOUcD5jKj_nC8A9qnDLgJvZmwuRrw/edit#gid=0

Progress of our Project

- We have devised an algorithm to convert an image into a matrix.
- We have also implemented the algorithm used in phase flow diagram to calculate fractal dimension
- We are able to segregate the images to some extent based on our algorithm but it still needs to be reviewed for precision.
- We are yet to train the AI model for concluding our project which we have decided to do in the summer.

Possible Improvements:

- In our calculation we resize all images to 256*256 so some information could be lost (for larger images) and some incorrect info could be considered (for smaller images).
- Some figures gave error > 10% , we could try changing parameters to possibly reduce this.
- Our method works only for 3d images.
- Over the years different formulas have been proposed to calculate h so we could try different variations.
- In our calculation , we are taking G=M=256 which gives h=s always. We could try varying the values of G and M.
- And some more..

Other Applications : Multiphase Flow

It is the simultaneous flow of materials with two or more thermodynamic phases

Concurrent movement of oil, water, gas, and solids within pipelines and production facilities.

Oil: This is the primary substance being pumped. It's the valuable resource extracted from underground reservoirs.

Gas: Natural gas can often be found alongside oil deposits. It's composed primarily of methane but can contain other hydrocarbons as well.

Water: Water is also present in many oil wells, either naturally occurring alongside the oil or injected into the well to maintain pressure or enhance oil recovery.

- **Efficiency:** Efficient separation of oil, gas, and water is essential for optimizing production and minimizing waste.
- **Safety:** Proper distribution ensures that equipment operates safely and effectively, reducing the risk of accidents or equipment failures.
- **Environmental Impact:** Accurate phase distribution helps minimize environmental impact by preventing spills or leaks of oil or other substances.
- X-ray tomography is used to visualize and analyze the distribution of these phases within the flow.

Future Prospects

The significance of this project lies in its potential to revolutionize cancer detection by introducing a novel approach that combines fractal geometry and machine learning. By providing automated and objective analysis of medical images, this approach has the potential to enhance diagnostic accuracy, improve patient outcomes, and streamline clinical workflows.

Further, we aim at training an AI model for the same after validating our algorithm on a greater dataset.

References

- An_Improved_Method_to_Estimate_the_Fractal_Dimension_of_Color_Images?
- https://en.wikipedia.org/wiki/Minkowski%E2%80%93Bouligand_dimension
- Fractal_Algo_BoxCounting by Nirupam Sarkar and B. B. Chaudhuri.
- https://www.mdpi.com/1099-4300/19/10/534
- https://multibandtexture.recherche.usherbrooke.ca/normalized_brodatz.html
- https://www.researchgate.net/figure/Computational-fractal-dimensions-of-Brodatz-texture-images_tbl1_224057945
- https://www.intechopen.com/chapters/6684BB Chaudhari and Nirupam Sarkar.

- Special thanks to Prof. Prabhat Munshi, Pridhi Athe, Nehlata Shakhya.

Report by-

Asmi Srivastava (2022ucs0084)

Purvanshi Nijhawan (2022ucs0103)

Shubham Gupta (2022ucs0110)

Sohit Dhawan (2022ucs0112)