AI Framework with Computational Box Counting and Integer Programming Removes Quantization Error in Fractal Dimension Analysis of Optical Images

Haoyue Liang,[1] Michael Tsuei,[2] Nicholas Abbott,[2] Fengqi You[1,2,*]

[1] Systems Engineering, College of Engineering, Cornell University, Ithaca, New York, 14853, USA

[2] Robert Frederick Smith School of Chemical and Biomolecular Engineering, Cornell University, Ithaca, New York, 14853, USA

Submitted to *Chemical Engineering Journal* for the *VSI: MaCKiE 2021 Special Issue*

Abstract

Because many Chemical Engineering phenomena and processes involve distinctive shapes and structures, fractal dimension (FD) analysis is of great interest to the Chemical Engineering community due to its utility in providing a statistical index of shape complexity, which not only quantifies image structure but also helps explain functional properties. Past box counting (BC) methods for estimation of FDs are inconsistent due to quantization error (QE) introduced from image rotation and translation. In this work, we propose a systematic and automatic artificial intelligence (AI) framework that consistently estimates FDs of relevance to different fields of Chemical Engineering without QE by integrating image

* Corresponding author. Phone: (607) 255-1162; Fax: (607) 255-9166; E-mail: fengqi.you@cornell.edu

preprocessing, set-covering optimization-based BC, and regression analysis. As a result of the deterministic optimization technique, FD estimations remain consistent for all images regardless of image rotation and translation. The results of image datasets obtained from fields such as interfacial science, biomedical engineering, human anatomy, among others, demonstrate efficient box size specification to maximize regression effectiveness and FD calculation accuracy for a variety of images. Overall, the proposed AI framework offers a new means of estimating FD accurately and efficiently for optical images of interest to the Chemical Engineering community.

Keywords: Fractal Dimension, Box Counting, Integer Linear Program, Regression Analysis, Optical Images

Abbreviations

Sets/Indices

I	Set of the row index
i,i'	Row indices
J	Set of the column index
j,j'	Column indices

Parameters

BS	Integer parameter that represents box sizes in powers of two less than a maximum box size
LBI	Integer parameter that represents the lower bound of row index
LBJ	Integer parameter that represents the lower bound of column index
M	Integer that denotes the maximum of structure height and width
N	Number of distinct copies of the measurement object
PW	Integer parameter that denotes the box size index
R	Scaling ratio whose inverse is proportional to the length of the measurement object
UBI	Integer parameter that represents the upper bound of row index
UBJ	Integer parameter that represents the upper bound of column index
$X_{i,j}$	Integer parameters to indicate pixel luminance at the *i*th row and the *j*th column of the image

Binary 0-1 Variables

$y_{i,j}$	Binary variables that represent the Boolean designating the presence of the upper-left origin of a box at the *i*th row and the *j*th column of the image

Variables

d	Variable that represents the sum of $y_{i,j}$, which is the minimized box count at a certain *BS*
FD	Variable that represents fractal dimension which is derived from (*N* and *R)* or (*BS* and *d)*

1. Introduction

1.1 Motivation

Fractal geometry is ubiquitous in nature, as self-similarity is found in almost every natural setting and scale, from human anatomy to the evolution of earth's geographical features [1]. Past studies have investigated the application of fractals to various fields [1, 2], including neuroscience [3], plant biology [4, 5], and interfacial science [6]. Unlike Euclidean geometry, fractal compositions cannot be defined by an ideal shape primitive, nor do they easily facilitate quantitative comparison to each other. It is important to consistently quantify, characterize, and compare the structures of fractal objects, as slightly different structures can result in tremendously different functions and properties. For instance, a folded protein in its native state can be safe and fully operational, whereas a small change in protein folding may lead to its toxicity or change in functionality [7]. While complex protein conformations can be challenging to compare directly, fractal geometries are a good metric for comparison. One such quantifiable value to compare fractal geometries is their fractal dimension (FD), which provides a statistical index of geometric complexity and helps in shape differentiation.

FD analysis has been widely used in shape analysis and pattern recognition of images in Chemical Engineering-related fields, as many Chemical Engineering products and systems involve distinctive shapes and structures. Creating an AI framework that automatically estimates FD of an input image would be helpful for shape differentiation and provide quantitative insights of shape complexity, which can be applied to various fields such as life science, interfacial phenomena, neuroscience, etc. that are relevant to Chemical Engineering. Past studies have shown that FD can be used to characterize the shapes of aggregated proteins [8], crosslinked polymers [9], activated carbon [10, 11], fluidized nanoparticle agglomerates [12], etc., due to its broad applicability. FD serves as a statistical index to describe the external roughness of objects, quantifying shape complexity as a ratio of the change in detail to the change

in scale. Consequently, external shape complexity can be quantified, shapes can be discriminated, and functional properties can be categorized and better explained [13]. Fig. 1 shows an example of how FD can be used to compare images of interfacial structures formed by proteins and employed to explain and classify experimental samples.

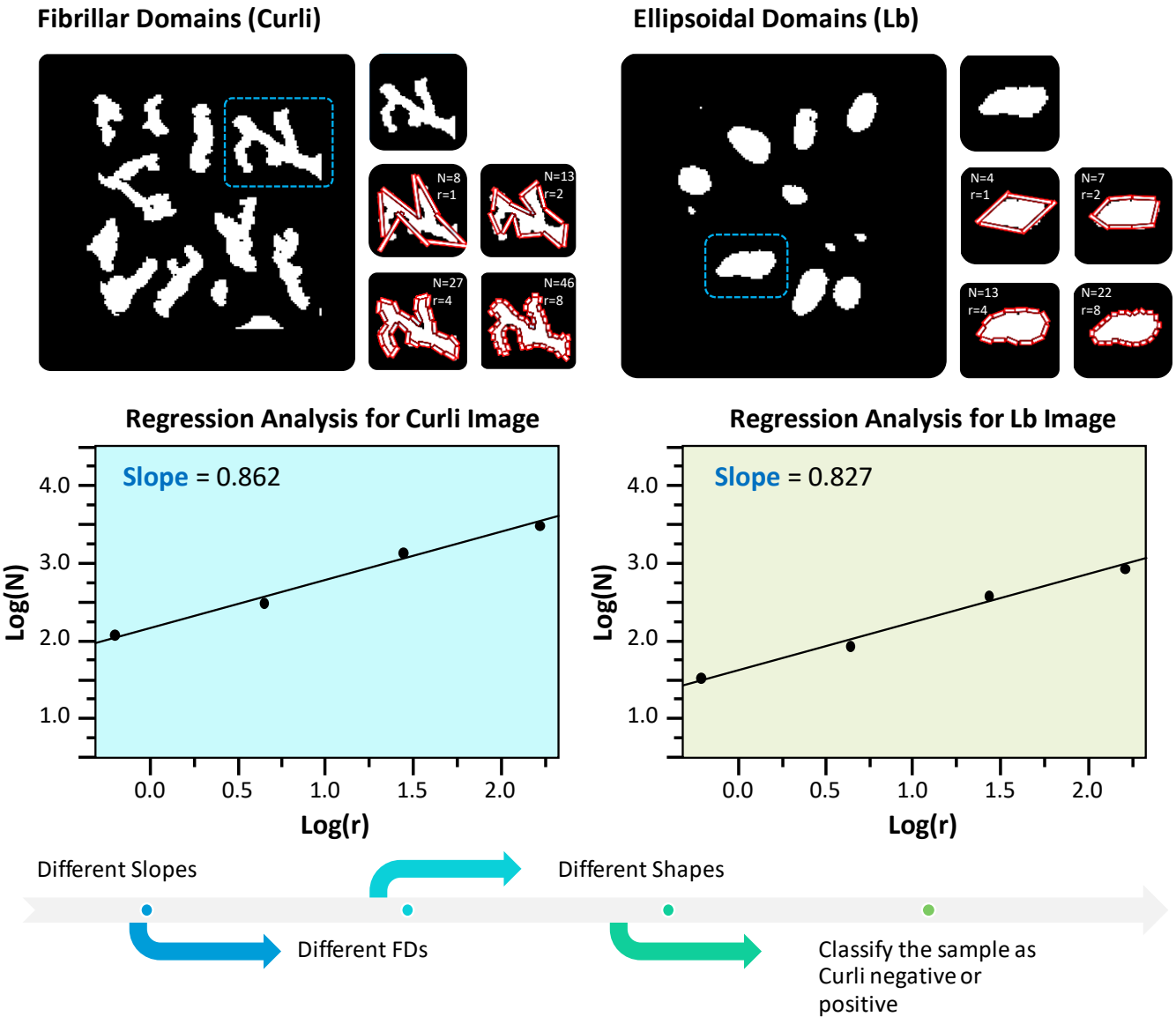

Fig. 1 Fractal dimension quantifies shape complexity as a ratio of the change in detail to the change in scale and can be used as a metric for shape differentiation or classification purposes, as shown by the example of interfacial phenomena (Curli vs. Lb protein) images. Curli images represent optical domains

rich in Curli, while Lb images denote domains rich in a lysogeny broth used for the culture of bacteria. Different regression slopes indicate different fractal dimensions, which further indicates different shape structures to be classified.

1.2 Preliminaries of FD

The basic equation of calculating FD is presented in Eq. (1).

$$N = \frac{1}{R^{FD}} \tag{1}$$

where N is defined as the number of distinct copies of the measurement object, namely, the red line segments in Fig. **1**. R is the scaling ratio, of which the inverse is proportional to the stick length. The number of line segments N is represented by the inverse of the scaling factor R raised to FD [14, 15]. Logarithmic transformations are then applied to both sides of Eq. (1), as shown in Eq. (2).

$$\log(N) = \log\left(\frac{1}{R^{FD}}\right) \tag{2}$$

After Eq. (2) is simplified, FD can be expressed as the ratio of the negative log value of the number of line segments over the log value of the scaling ratio:

$$FD = -\log(N) / \log(R) \tag{3}$$

In Eq. (3), N can only be theoretically determined if the object exhibits self-similarity. In most image analysis applications, the value of N is estimated using methods such as box counting (BC) [16]. The BC method gathers data for N and R, respectively, by subdividing images into unit squares and then quantifying the relationship between the total number of subdivisions and their length scale.

1.3 Gaps in Knowledge

The conventional BC methods are frequently applied tools for FD estimations, which involve superimposing grids upon the image and counting the number of grids that contain pixels defining the perimeter of the image [17]. However, they suffer from an inconsistency in the number of boxes counted under the same grid size upon rotation and translation of the image, known as quantization error (QE) [18]. As a result, the FD estimations derived from box sizes and box counts are inconsistent. Arbitrary grid placements can cause FD estimation to vary when conventional BC methods are used. Moreover, FD estimations are significantly affected by the range, values, and the number of grid sizes, which can be erratic or lead to invalid regression analysis [19]. It is therefore paramount to not only eliminate QE by counting boxes consistently but also guarantee valid regression analysis by choosing the appropriate box size combination. In order to develop a consistent BC method in this study, the BC approach is designed as a minimization problem, as the fewest box count is obtained each time under a specific box size. In other words, minimization guarantees consistency and eliminates QE despite image rotation and translation. In order to choose the appropriate box size combination, various combinations were tested, and the one with the best regression metrics was selected.

1.4 Contributions of This Work

This work proposes a systematic AI framework for estimating the FD of images in scientific and engineering fields. The framework has three major steps, namely image preprocessing, set-covering optimization-based BC, and regression analysis. Image preprocessing is the first step that involves converting the image from color to binary, formatting it properly, and extracting the perimeter of its largest domain. The image preprocessing step extracts a body perimeter for set-covering optimization-based BC analysis. In the set-covering optimization-based BC step, an Integer Linear Programming (ILP) model is formulated and solved to global optimality to determine the minimum number of boxes fulfilling total

coverage of the perimeter. The ILP model ensures that each perimeter pixel is covered by at least one box and minimizes the total number of boxes in its objective. The range and values of the box sizes used therein are determined by calculating the greatest x-y axis dimension measurement of the body perimeter extracted from image preprocessing, and they are adjusted by comparing the regression analysis metrics corresponding to each box size combination. Finally, the last step, regression analysis, requires plotting regression for logarithmically transformed box sizes and box counts data to obtain FD as the negative slope. To illustrate the applicability of the proposed framework, we put together a corresponding dataset comprising images from human anatomy, life science, interfacial science, etc.

The main contributions of this research are presented as follows:

- A novel AI framework that systematically estimates the FD values of structures serving as a universal metric for shape classification or pattern comparison of images from fields such as materials science, biomedical engineering, and interfacial science.
- The first BC method to eliminate QE arising from image rotation and translation through global optimization of minimum box counts by ILP.
- The first BC method that configures itself to select the most appropriate values, range, and the number of boxes to yield the best regression metrics based upon a dataset.

The rest of this paper is organized as follows. The proposed framework is described in detail in Section 2. The FD estimation results of images from the collected dataset are presented in Section 3. Section 4 concluded the paper.

2. Methods

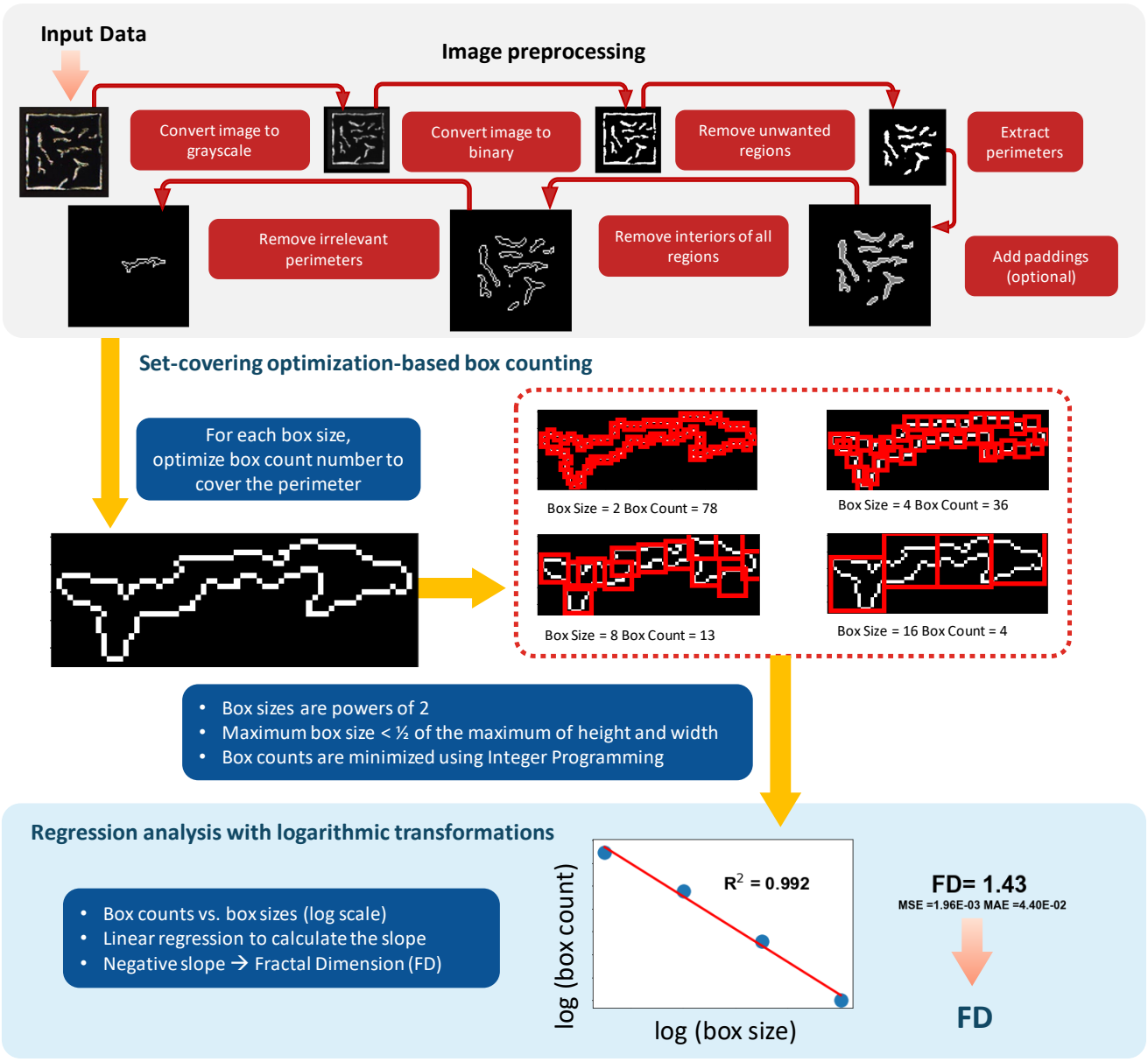

Fig. 2 Flowchart of the three steps in the proposed AI framework for evaluation of the FD of the input image: image preprocessing, set-covering optimization-based box counting, and regression analysis with logarithmic transformations. Box sizes are specified as powers of 2, and FD is calculated as the negative slope of the regression line of the logarithmic value of box counts vs. the logarithmic value of box sizes.

We propose a fully automated AI framework to estimate FD that intelligently connects three major steps, including image preprocessing, set-covering optimization-based box counting, and regression analysis with logarithmic transformations. Important information from the previous step is automatically fed into the next. As shown in Fig. 2, the proposed framework relies on the reproducibility of each major step to estimate the FD automatically and consistently as a whole. This section presents the three major steps of the proposed AI framework, namely image preprocessing, set-covering optimization-based box counting, and regression analysis, in detail.

2.1 Image Preprocessing

The automated image preprocessing step involves three sequentially connected sub-steps: image conversion from color to binary, image formatting, and perimeter extraction of the largest structure with the most extended perimeter. Each of these sub-steps makes the image more suitable for interpretation by the set-covering optimization-based BC step.

2.1.1 Image Conversion from Color to Binary

The first sub-step to make the image algorithmically interpretable is the procedure to convert the image from color to gray-scale by automatically eliminating the hue and saturation information of each pixel while preserving their luminance. After the image is converted to gray-scale, additional sub-steps are automatically implemented to convert the image into strictly binary luminance (black and white, or simply binary), such that the background has a black appearance. This is done by plotting a histogram of pixel count versus luminance (i.e., gray level) from a flattening compression of the image into a one-dimensional array of luminance values (i.e., gray values) inside the framework. For a better understanding of gray values, some gray pixels and their corresponding gray values are presented in Fig. 3. The gray-scale image typically has a dark background, so most background pixels will have gray values near 0,

corresponding to a dark appearance, out of the range of values between 0 and 255. If an input image has a light background to begin with, it would be automatically inverted so that the background has a darker appearance. The high density of dark background pixels appears as a spike in pixel counts in the histogram just above the 0 gray level. To isolate the geometry of interest from the background, the dark background pixels are adjusted to exactly the 0 gray level. In order to ensure geometry isolation works best with image sources of varying brightness and contrast to optimize domain selection protocol under different image brightness, a threshold value t of a scalar factor greater than the mode gray value of the large background peak is automatically generated, and then the threshold t is funneled to adaptably distinguish between pixels that should be categorized as background or as relevant geometry (the selection of t values is discussed in detail in Section 3). All pixels with gray levels less than t are automatically adjusted to a gray level of 0 (pure black), and all other pixels are adjusted to a gray level of 255 (pure white), representing the geometry to be analyzed. An adaptive thresholding method is preferred among all image segmentation methods because of its wide image preprocessing applicability [20], being capable of handling images captured or generated with varying brightness and contrast levels.

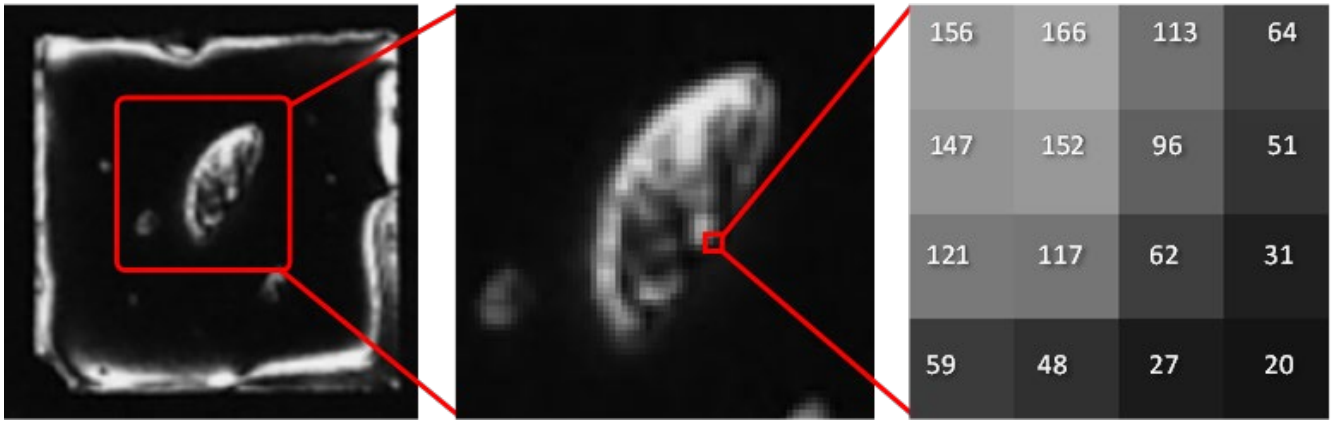

Fig. 3 The left image is a gray-scale micrograph of the optical response of a liquid crystal to the adsorption of protein at its interface, while the zoomed-in image is presented in the middle. Pixels of the optical domain are converted to pixel gray values between 0 to 255, where the lower the pixel intensity, the lower

the pixel gray value. The right image shows the pixel colors of some example boundary pixels with their gray values.

2.1.2 Image Formatting

After being converted to binary from the previous sub-step, the images require further automated processing sub-steps to facilitate the set-covering optimization-based BC. The key requisite for the regression analysis step is that at least three, but more commonly four or five, BC data points are obtained for regression. Therefore, images of at least 100 pixels, which can accommodate at least three or four box sizes, are the minimum valid resolution. Any fewer data points for regression analysis would be invalid and alternative methods to BC should be used to obtain the FD. Image edge paddings are automatically added by introducing extra pixels surrounding the image when the pertinent geometry has inadequate space to the image borders. Without this space, an effective solution to the ILP in the optimization-based BC step will be restricted. The ILP model is formulated in a way such that the upper left corner (top left pixel) of each box might not lie on the image structure directly, but on the background pixel, while the whole box still covers some portion of the perimeter. Therefore, sufficient background pixels are needed, such that box placements are solely restricted by the structure. Padding around the image ensures that all box sizes proposed can be iterated, and indices are not out of range for the set-covering optimization-based BC analysis. This sub-step on perimeter extraction is not conducted if there is adequate space surrounding the structure of interest. The framework would automatically identify if insufficient space is an issue and decide on whether to format the image by image edge padding.

2.1.3 Perimeter Extraction of the Largest Domain

The final sub-step in image preprocessing is the extraction of the perimeter. The set-covering optimization-based BC step and the concept of FD itself are concerned with the perimeter of the geometry involved, so the interior of any structure may be effectively discarded, as is practiced in conventional BC

methods to estimate FD values [2]. Additionally, this perimeter extraction operation can, for the sake of automation, remove all geometries except the largest one. This is particularly useful for isolating important geometries and cleaning up images that may contain artifacts, noise, or small extraneous debris.

For perimeter extraction, pixels are automatically classified into two categories: perimeter pixels and interior pixels. A perimeter pixel is defined as any foreground pixel adjacent to at least one background pixel, and an interior pixel is defined as all other foreground pixels with no adjacent background pixel (grey level of 0). If this is the sole geometry of interest, then interior pixels are all converted to black to effectively remove the interior, and the image is ready for the set-covering optimization-based BC step; otherwise, if deemed necessary, all perimeters except that of the largest structure are removed before proceeding to the next step. Although some other structures are large enough for the analysis as well, when set-covering optimization-based BC is performed on each of these structures separately, the FDs estimated for each shape are not significantly different for the dataset used in this study. Most domains within an image that belong to the same group elicit similar FDs regardless of size. In our FD analysis, the most representative structure is always selected from various datasets. We found that the largest shape is a good representation of the whole image because it is big enough to show details of its shape complexity compared with other smaller shapes. Based on our preliminary observation of domain structures, all of the optical domains first undergo nucleation with only a few pixels in terms of size, and then propagation when the size of the optical domain increases into either well-defined ellipsoidal or fibrillar shapes. During the nucleation step of a structure, when the shape of that structure is not well-developed, we found that the FDs of these structures are nearly identical. Therefore, selecting the largest structure, which would typically propagate over a long period, is a better option due to its more thoroughly developed geometric details. In the Curli and Lb images in **Fig. 1**, for example, the largest domain was selected for its greater detail which enabled us to use at least four box sizes.

To identify the largest structure, the structure sizes are automatically compared simply by the number of non-perimeter pixels within the structures. To do so, non-perimeter inner pixels are automatically set to an intermediate gray value, neither 0 nor 255, while the perimeter pixels are maintained at their luminance of 255 (white), thereby distinguishing interior pixels from perimeter pixels. Each structure is also distinguished from others by a unique index, which is attached to every interior pixel within each structure. Once interiors are indexed and distinguished from their perimeters, area sizes are easily determined by a pixel tally. All structures except the largest one are automatically set to zero gray value, and only the largest perimeter remains to be subject to BC analysis in the next step.

2.2 Set-Covering Optimization-Based BC

2.2.1 Integer Linear Programming (ILP)

With the perimeter of interest extracted and isolated, an automated step of set-covering optimization-based BC is performed to estimate the FD, as shown in the second step in Fig. 2. Compared to the conventional grid counting approach, the proposed method is more systematic and rigorous in minimizing the QE. Specifically, the proposed algorithm is implemented by automatically creating an ensemble of boxes that envelops the entire perimeter with the minimum number of boxes possible. The boxes can overlap if necessary. The minimized number of boxes is obtained by ILP, and the number is independent of irrelevant conditions that may cause QE.

Before an ILP model is created, a few steps are designed in the framework to determine the given parameters. First, the smallest conceivable rectangle that covers the whole perimeter is automatically generated, as shown in the left side image of step 2 in **Fig. 2**. The coordinates of the four vertices of the rectangle are set to (*LBI, LBJ*) at the top left, (*LBI, UBJ*) at the top right, (*UBI, LBJ*) at the bottom left, and (*UBI, UBJ*) at the bottom right. Second, the perimeter width (*RANI*) and height (*RANJ*) are automatically calculated, as shown in Eq. (6) and (7), where the width is the distance between the top left

vertex and top right vertex, and the height is the distance between the top left vertex and the bottom left vertex. Third, the smallest box size is automatically set to two, and the largest box size set to the largest power of two that is smaller than or equal to the maximum of perimeter width or height (M) over two. M is obtained from Eq. (8) , and the range of box sizes (BS) is shown in Eq. (9). The range selection is explained in the next sub-step: Box Size Range and Value Selection in Section 2.2.2. As the range is determined, the framework automatically obtains all the box sizes BS, which are all powers of two and their indices (PW). Each BS is equal to two to the power of each PW, as shown in Eq. (10). Since the constraints vary under different box sizes, a unique ILP problem is solved for each box size correspondingly. The total number of ILP problems created for each image is PW, and we need PW data points for regression analysis.

In each ILP model, box size BS is given. Indices i and j are in the range of LBI to UBI and LBJ to UBJ as shown in Eq. (11) and (12), since the perimeter of interest is within the rectangle. $X_{i,j}$ that designate the pixel gray values, at the ith row and the jth column of the image correspondingly, are given parameters that can be obtained from the image directly. If $X_{i,j}$ is equal to 0, the corresponding pixel is black, and there is no need to cover it; otherwise, at least one box is needed to cover that pixel since it is part of the white perimeter. To indicate if boxes are needed to cover pixels at certain locations, we introduce a binary variable $y_{i,j}$. If a box's upper-left origin is used at the ith row and the jth column of the image, $y_{i,j}$ equals to one; otherwise, it takes the value of 0 if no box is necessary at that coordinate. The values of $y_{i,j}$ are to be determined by the ILP model.

The objective of each model is to minimize d, which is the total number of boxes to cover the perimeter, as represented by the sum of all binary variables $y_{i,j}$. A constraint is only added when $X_{i,j}$ is greater than 0. When $X_{i,j}$ is nonzero at the ith row and the jth column of the image, at least one box is needed to cover the perimeter pixel at that location, which is represented as Eq. (5). Among all the potential boxes that can

cover the perimeter pixel, as indicated by all the $y_{i',j'}$ inside the parenthetical summation in Eq. (5), at least one of them must be used. If a box's top left pixel at the *i'*th row and the *j'*th column lies on the top left of that perimeter pixel, and the box's side length *BS* is big enough, that box can be used to cover the perimeter pixel.

The ILP model is formulated as follows:

$$\min \quad d = \sum_{i=LBI}^{UBI} \sum_{j=LBJ}^{UBJ} y_{i,j} \tag{4}$$

$$\text{s.t.} \quad X_{i,j} \left(\sum_{i'=i-BS+1}^{i} \sum_{j'=j-BS+1}^{j} y_{i',j'} \right) \geq 1, \quad \forall i, i' \in I, j, j' \in J \tag{5}$$

$$RANI = UBI - LBI \tag{6}$$

$$RANJ = UBJ - LBJ \tag{7}$$

$$M = \max\{RANI, RANJ\} \tag{8}$$

$$2 \leq BS \leq \frac{M}{2} \tag{9}$$

$$BS = 2^{PW}, PW \in Z^{+} \tag{10}$$

$$I = [LBI, UBI] \tag{11}$$

$$J = [LBJ, UBJ] \tag{12}$$

$$y_{i,j} \in \{0,1\}, \forall i \in I, j \in J \tag{13}$$

where $X_{i,j}$ is the pixel grey value at the *i*th row and the *j*th column of the image, and a binary variable $y_{i,j}$ indicates if a box whose upper-left origin is at the *i*th row and the *j*th column of the image is used to cover

pixels. We optimize the total number of boxes to cover the perimeter, as represented by the summation of all binary variables $y_{i,j}$. Indices i and j are in the range of *LBI* to *UBI* and *LBJ* to *UBJ*. *BS* are the box sizes, which are all powers of two and *PW* are the indices.

2.2.2 Box Size Range and Value Selection

In conventional box counting approaches, box sizes are typically dictated as (1,2,3,4,6,8) [19], (2,4,8,16) [21], factors of image size [22], (3,9,27,81) [23], or (8,12,16,32) [24], but justifications supporting such selections are often omitted. However, past studies have reported the use of box sizes of powers of two [25, 26], or in the range from two to $M/2$ [27, 28], where M is the image length or the x-y axis dimension of the geometry of interest. We incorporated lots of image preprocessing and optimization-based procedures that are formulated based on pixel colors and properties, so it would be appropriate to have box size as a multiplier of pixel size. For instance, a box with the smallest size two indicates a box whose size is two by two pixels. In the implementation of set-covering optimization-based BC in the proposed AI framework, the box sizes are also calculated as powers of two but ranging from two to $M/2$, where M is defined as either the height or width of the perimeter, whichever is greater. This selected range of box sizes eliminates oversized boxes, which entirely envelop the perimeter with just one box. The box size of one is also eliminated since its box count is just the total number of pixels that constitute the perimeter. Both of the cases of box size equal to one and box count equal to one are eliminated since these extreme values are not of great use to the regression analysis, and no optimization is involved in either case [25, 29]. As FD is defined as the negative log value of the box count over the log value of the box size, both box counts and box sizes are logarithmically transformed in the regression analysis. In any regression analysis where one data point is far away from the rest, the regression result would be greatly affected if an error occurred when obtaining that specific data point. However, when all the data points are equidistant, the impact of statistical errors on regression slope is minimized. Since regression is most

effective when each point is equidistant on a log-scale, powers of integer numbers are selected as box sizes. Moreover, using powers of two can generate more box data than using powers of three or four for small structures and smaller images. The set of box sizes to be counted for the perimeter is thus enumerated as powers of two in the range from two to $M/2$, but if $M/2$ is not a power of two, the largest box size is set to the greatest power of two no larger than $M/2$ [25-28].

2.3 Regression Analysis with Logarithmic Transformations

After BC data are obtained, the FD can be automatically estimated as the negative slope of a regression analysis with logarithmic transformations of both the box size data BS and the minimized box count data d, as shown in the third step of **Fig. 2**. To estimate FD for image analysis using BC, we obtain FD based on the equation below:

$$FD = -\frac{d\log\left(d(BS)\right)}{d\log\left(BS\right)}, \qquad \forall BS \in \left\{ 2^P \,|\, 1 < 2^P \le \frac{M}{2}, P \in Z^+ \right\} \tag{14}$$

where FD is defined as an integer variable, whose negative value equals the slope of box count as a function of the box size. The regression slope indicates the degree of complexity, or negative FD, where FD falls between one and two [14, 30].

3. Results and Discussion

In this section, we applied the proposed AI framework to obtain the FD estimation results of 10 example images, 40 case study images in the field of interfacial science, and all 167 images in our collected dataset.

3.1 FD Estimation Results of 10 Example Images

The FD estimation results of 10 example images in interfacial science, neuroscience, plant science, etc., are presented in this section. FD serves as a statistical index for purposes of pattern comparison and shape characterization, which illustrates the effectiveness of the proposed AI framework.

3.1.1 Preprocessed Image Compatibility

Image preprocessing techniques within the AI framework effectively format and accommodate input images with different resolutions, colorations, and image clutter automatically. During testing, it was found that the majority of fields of study produce images that have enhanced contrast with higher-luminance structures of interest on a low-luminance dark background, as shown in images a1, a2, b1, b2, c1, c2 (interfacial phenomena and neuroscience images) of Fig. 4; such input images can be immediately preprocessed for image conversion from full color (on the left-hand side of Fig. 4) to binary (on the right) automatically. All images which instead have a low-luminance structure of interest or high-luminance background (such as images d1, d2, e1, e2 in Fig. 4) are automatically inverted before being fed into image conversion from color to binary. After the image is converted from full color to binary, the diverse range of image resolutions, as well as image clutter can then be accommodated in the steps detailed in image formatting and perimeter extraction, and as a result, a comprehensive range of images can be correctly converted to a single distinct perimeter. This directly results in the relative success and ease of applying the AI framework across the image dataset. Apart from the need for image inversion to ensure brighter geometry upon a dark background, the AI framework can be made compatible with and automated for nearly any type of image – an invaluable utility for many FD estimation contexts.

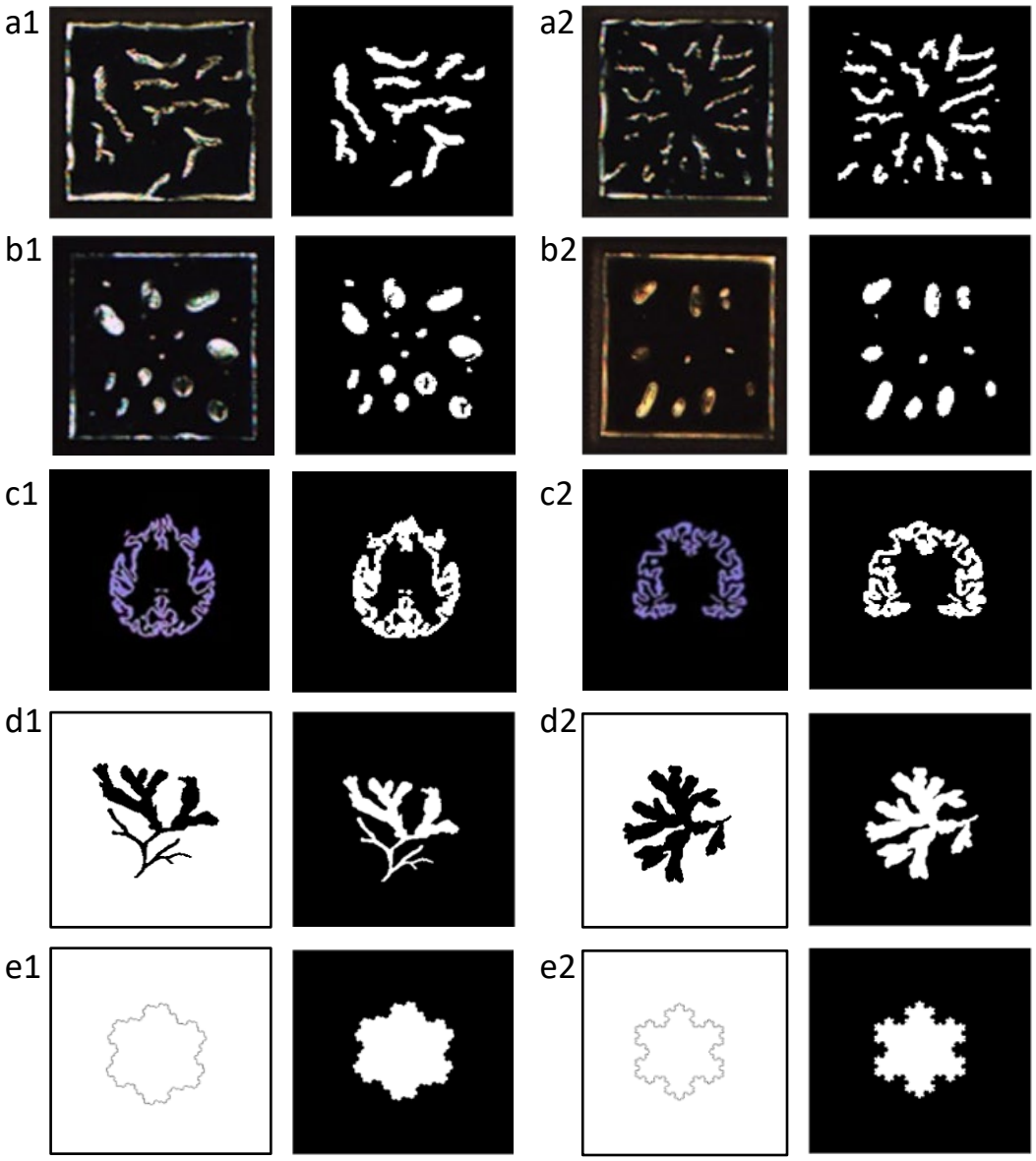

Fig. 4 For each pair of images, the left image corresponds to the original source images, while the right image is the segmented binary image used for testing. a1, a2, b1, and b2 are interfacial phenomena images generated from our high-throughput experiments. c1 and c2 are human anatomy images [31]. d1 and d2 are plant science images [4]. e1 and e2 are classic fractal geometries [18, 32].

3.1.2 Efficient Box Size Specification

The selection of box sizes evaluated in the set-covering optimization-based BC was justified and aptly designated in this study. All box sizes (integers ranging between one to the longest linear dimension measurement) were evaluated on the test data (the 10 example images), while their corresponding minimized box counts were recorded. Sets of box sizes that successfully yield different box counts were chosen, for instance (2,4,8,16) [21], (divisors of image size) [22], (8,12,16,32) [24], etc., and had their resultant regressions plotted to compare the R^2, MSE, and MAE values. We found that when the set of box sizes were in powers of two, they would yield the greatest number of equidistant log-log regression analysis data points below the critical box size, to maintain a high R^2 value and low MSE and MAE**.** Powers of two were chosen based on three criteria. First, logarithmically diverse box sizes minimize regression data error, including QE. Second, well-distributed datasets reduce the likelihood of repeating box counts, which otherwise reduce the accuracy of the regression analysis. Third, powers of two enable more box sizes to be counted before reaching the critical maximum box size, as compared to larger integer values.

3.1.3 Box Counts Closely Follow Linear Behavior

The set-covering optimization-based BC step produces BC data closely following the linear behavior that defines FD based on regression analysis. The computational time of ILP optimization and regression programs implemented in Python 3.8 usually falls between one to four seconds, depending on the image size and largest perimeter size. Below are some example computational times of interfacial science images.

Table 1 Computational times to run the proposed AI framework for the interfacial science images generated from our experiments.

Example images	Total program running time	Wall time
Interfacial Science 01 (Curli)	3.29 s	4.28 s
Interfacial Science 02 (Curli)	2.62 s	3.53 s
Interfacial Science 03 (Curli)	3.30 s	4.12 s
Interfacial Science 04 (Lb)	2.11 s	2.91 s
Interfacial Science 05 (Lb)	2.23 s	3.06 s
Interfacial Science 06 (Lb)	2.06 s	2.55 s

In most cases where image and perimeter size are large, for each given box size, an ILP problem is formulated with more than 1000 variables and 100 constraints. Each ILP problem generates a box count that passes on for regression analysis with logarithmic transformations. When all box counts are obtained, to quantify the linearity in regression analysis, the R^2, MSE, and MAE metrics are used to evaluate the prediction error rates and model performance in regression analysis. The resulting MSE and MAE are typically small, while R^2 values always remain greater than 0.99. As shown in **Fig. 5**, which contains the regression results of the same segmented binary images presented in **Fig. 4**, the BC data (log box size and log box count) strongly follows a linear relationship in each regression analysis. R^2 values demonstrate nearly ideal fits in regression models for a highly varied set of images. Apart from the R^2 value, which is largely affected by variance, MAE measures the absolute average distance between the real data and the predicted values. The MAEs of all images shown in **Fig. 5** are below 0.05, which indicates the average magnitude of error from every sample in this dataset is minimal. Besides MAE, which sometimes fails to highlight major errors in predictions, MSE indicates the quality of the estimator/predictor of the regression

model. In **Fig. 5**, all MSEs are smaller than 2×10^{-3}, testifying to the accuracy of the regressor in each regression model.

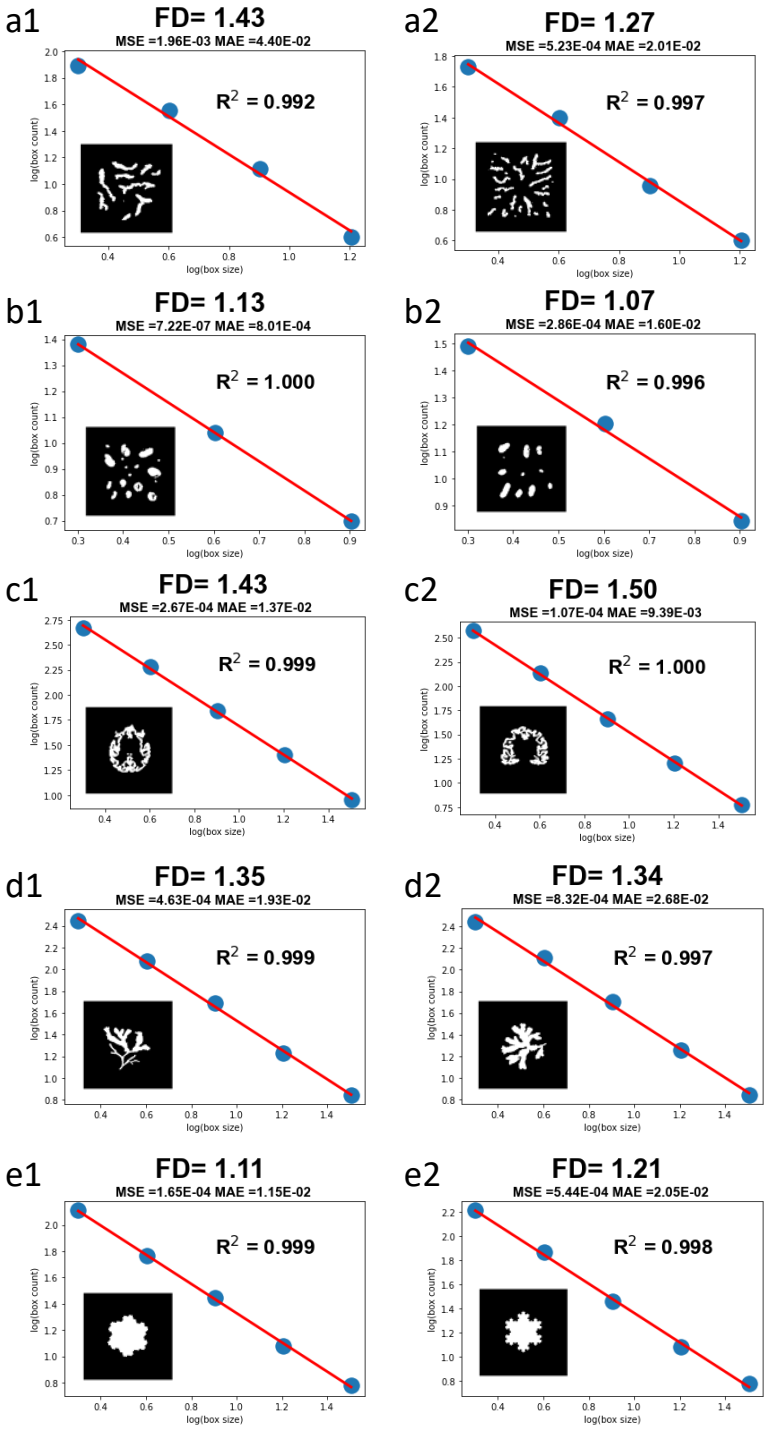

Fig. 5 Regression analysis with logarithmic transformation shows excellent regression metrics of segmented binary images presented in Fig. 4. a1, a2, b1, and b2 are interfacial phenomena images generated from our high-throughput experiments. c1 and c2 are human anatomy images [31]. d1 and d2 are plant science images [4]. e1 and e2 are classic fractal geometries [18, 32].

3.1.4 Quantization Error Elimination for Accurate FD Estimation

The proposed AI framework eliminates QE to precisely and accurately estimate FD. When each image in the dataset is rotated, the set-covering optimization-based BC yields essentially identical box count data regardless of rotation angle, as the box counts are always strictly minimized. The consistency gained from eliminating QE allows for precise prediction of FD, which can then be determined to be both accurate and precise by comparing with the Hausdorff-Besicovitch dimension at less than 4% error.

The Hausdorff dimension is another measure of FD [33], alongside the BC interpretation. The Hausdorff dimension can be used to compare FD values estimated by different methods and is often used on formulaic fractals since their Hausdorff dimensions are exact values. For example, the Hausdorff dimension of the Gosper Island contour (image e1 in Fig. 4) is about 1.13 [34], while the FD estimation from the AI framework yields 1.11, which differs from the Hausdorff dimension by only 1.68%. The Hausdorff dimension of the Koch curve or Koch snowflake (image e2 in Fig. 4) is approximately 1.26 [34], and the framework estimates its FD to be 1.21, a 3.97% difference from the Hausdorff dimension. The FD estimation of the Sierpinski triangle is 1.53 [34], which is 3.16% different from its Hausdorff dimension of 1.58. Similarly, evaluations and comparisons of FD against the Hausdorff dimension for fractals such as the Julia Set, the boundary of the terdragon curve, and the boundary of the tame twindragon curve [35], also result in discrepancies of 4% or lower. The FDs of the dataset tested, including images of formulaic fractals, are presented in Table S1. These benchmark geometries have low FD estimation errors

below 4%, particularly when considering the MAE and MSE of most regressions, representing a strong adherence of the BC method to the mathematically rigorous definition of FD. Based on our AI framework, the results generated provide accurate FD estimations of deterministic fractals compared to their Hausdorff dimensions. The FD results can help us visualize what it means for shapes to high or low fractal dimensions

3.2 FD Estimation Results of 40 Interfacial Science Images

As an example of the framework being able to discriminate FD between similar images, the FD estimation results of Curli vs. Lb images exhibit a significant difference between their FD values.

3.2.1 Image Classification by FD Analysis

The distributions of FD estimations for interfacial science images (Curli vs. Lb) are shown in **Fig. 6**. A parametric method to perform a two-sample *t*-test, which is appropriate for examining the difference in means for two populations, is used in this study, and the test results are shown in Table 2.

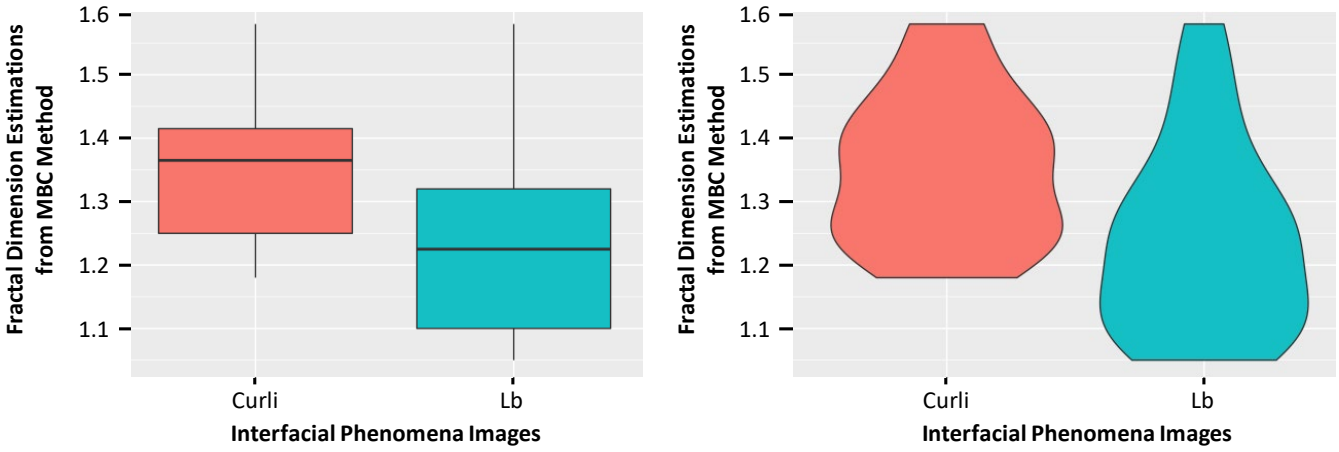

Fig. 6 Fractal dimension estimation comparison of Curli and Lb images. Curli structures typically have fractal dimensions over 1.19 with an average of 1.35, while Lb structures have a broader range of fractal

dimensions with an average of 1.24. Images that have low FD values (below 1.19) could generally be characterized as Lb images rather than Curli images.

Table 2 Two-sample *t*-test for interfacial phenomena Curli and Lb images to determine whether the true difference in means between these two populations is statistically significant.

Test Statistics	**Value**
t	2.67
df	36.46
p-value	0.011
	Alternative Hypothesis:
	The true difference in means is not equal to 0.
95% confidence interval	[0.027, 0.201]
mean of Curli sample estimate	1.35
mean of Lb sample estimate	1.24

The statistical analysis results demonstrate that we are 95% confident that the true difference in average FD values between Curli and Lb images is greater than 0.027 while smaller than 0.201, and the true difference in means is over 98% likely to be nonzero. Experimental protocols can be optimized to maximize the differences in the FD values between Lb and Curli datasets. In the Curli/Lb example, optical

signatures of the LC in response to protein aggregates rich in alpha-helices yield elliptical structures with fractal dimensions close to 1, whereas the LC response to protein aggregates rich in beta sheets yield branching structures with fractal dimensions around 1.6. Thus, if an LC film is exposed to an unknown solution comprising either LB or curli, fractal dimension estimations of the LC optical domains generated by the AI framework can help us determine which solution the LC was exposed to.

As shown in both the boxplot and the violin plot of **Fig. 6**, the maximum FD estimations in the sample dataset are approximately the same for Curli and Lb images. However, their minimum FD values and first quantiles (Q1s), which are the values under which 25% of data points are found when they are arranged in increasing order, are not the same. The minimum FD estimation for Curli images is 1.19, while a great number of Lb images have FD estimations between 1.05 and 1.19. Since the sample dataset shows the absence of Curli images whose FD values fall below 1.19, it is likely that images that do have low FD values (below 1.19) could be characterized as Lb images rather than Curli images. This is an example of the capability of the BC method within the AI framework to differentiate between otherwise indistinguishable images.

3.2.2 Threshold t Selection Affects FD Result

The threshold t used in the image segmentation thresholding step can change the FD result. As mentioned earlier, in the image preprocessing step where color conversion was performed, the threshold selection t is automatically generated and not fixed. The value of t is automatically derived from a histogram of the pixel count at each gray-scale value, where the mode value is used as a benchmark to gauge the overall luminance of the imaged bodies and for t to be set relative to. Too high a threshold value breaks the structure domain into multiple pieces, and too low of a threshold fails to eliminate noise from the image. This effect is shown in the bottom left of **Fig. 7**. Neither breaking a structure into two pieces nor keeping debris and noise is advised, so t is chosen to be exactly the mode in the example image.

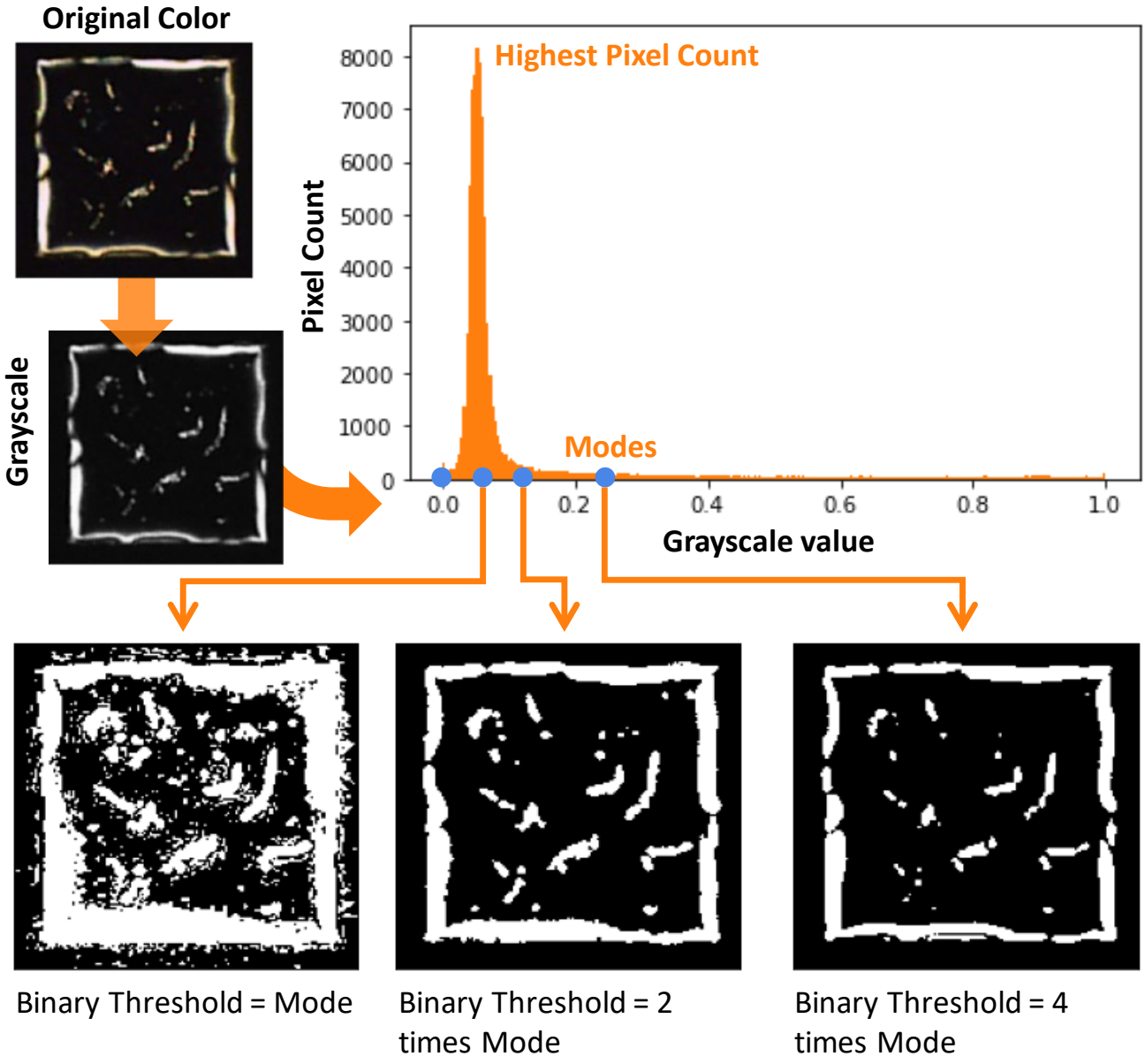

Fig. 7 The effect of threshold value *t* on image preprocessing outputs. Too high a threshold value breaks a structure into multiple pieces, and too low of a threshold fails to eliminate noise from the image.

Certain categories of images can prove to be more sensitive to this issue than others, such as the interfacial science images in the case study. In these circumstances, it will likely be necessary to work with domain experts who are familiar with the true structure of the imaged body, and if such a body had ought to be interpreted as monolithic or fractured. The importance of background knowledge, such as our

insights on these Curli and Lb protein aggregates and knowledge of the experimental conditions under which the images were obtained, cannot be understated. Ideally, for any study of fractal geometry, may it be image thresholding, clustering, or some other image segmentation method, some familiarity of the actual geometry represented by the image will always be needed to determine the threshold value *t* appropriately, being key to the interpretation of the roughness of the perimeter and therefore key to truly measuring the FD.

3.3 FD Estimation Results of the Entire Dataset of 167 Images

3.3.1 Dataset Description

The AI framework we proposed is data-independent, and it was tested on the image data from both our own experimental studies and literature relevant to the field of Chemical Engineering that raises the concept of fractal dimension. The framework is not specifically tailored for the image dataset but is rather designed independently to be universally applicable to images from multiple disciplines. To show the general applicability of the AI framework, from as much literature involving the FD concept as we know, we collected 167 meaningful images for testing, as shown in Table S1. We filtered out the unusable images from previous publications and made the dataset as diverse as possible to the best of our knowledge. Based on our FD estimation results of the image dataset, the framework is compatible with many different fractal geometry representations and can be used for a variety of applications. The framework was validated with a dataset with 167 images, which are collected from reputable datasets in seven different fields of study from 12 sources [4, 18, 31, 32, 36-42], demonstrating its robustness and versatility, as shown in **Fig. 8**.

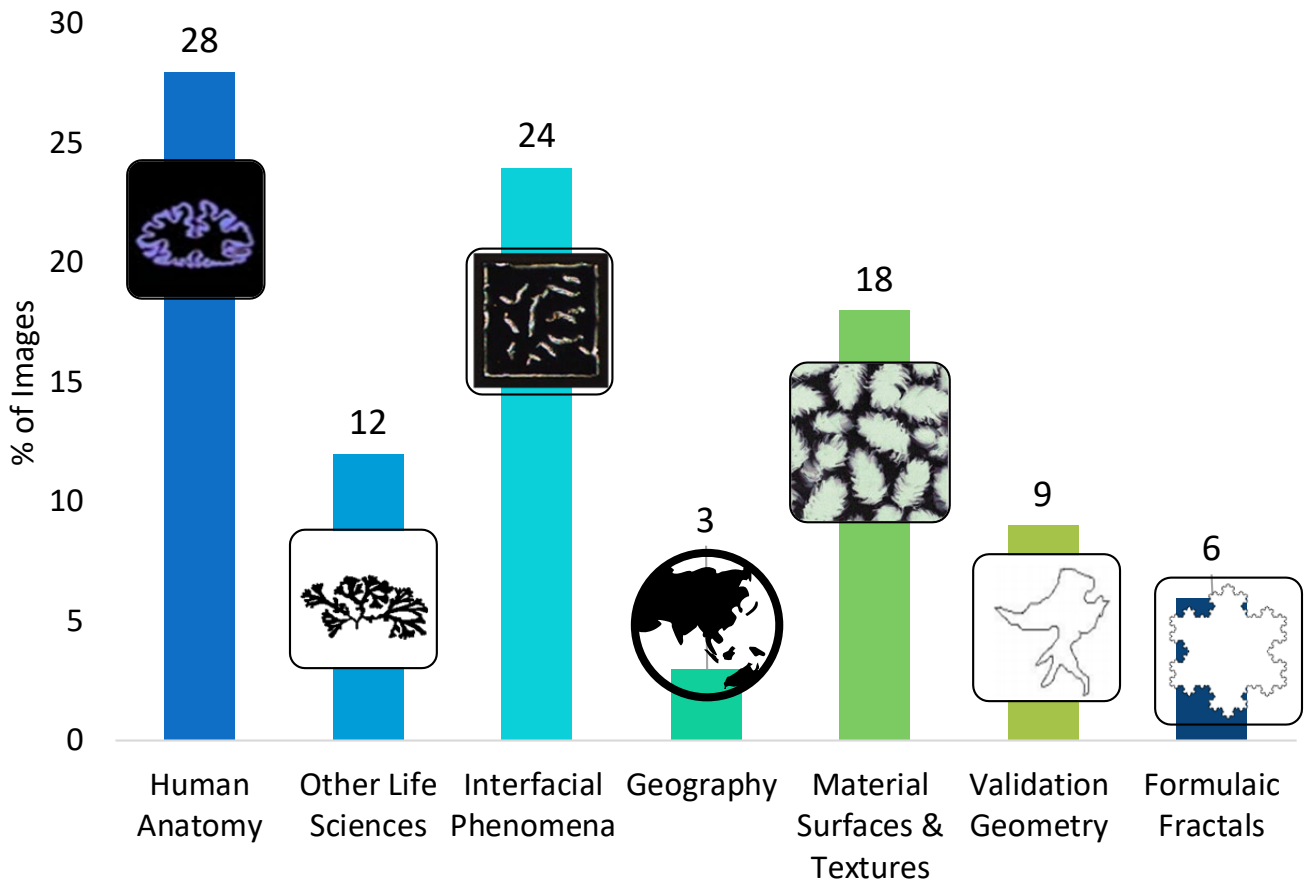

Fig. 8 A dataset of 167 images from diverse disciplines [4, 18, 31, 32, 36-42], including human anatomy, life science, interfacial phenomena, etc., and the percentage of images from each field used in the analysis presented in this paper.

We used 40 optical images of liquid crystal films with interfaces decorated by proteins, which we obtained as optical domains with characteristic geometries representative of interfacial science. Specifically, 20 images were obtained from LC films exhibiting optical domains rich in Curli, amyloid that forms the main structural unit of bacterial biofilms, and 20 images were obtained from LC films exhibiting domains rich in a lysogeny broth used for the culture of bacteria (Lb). The ability to distinguish the LC response to Curli and Lb has the potential to enable early detection of the presence of pathogenic bacteria.

Aside from the interfacial science images, 46 human anatomical images, a popular field for FD estimation applications, were also included in the dataset, such as brain images whose FDs can be correlated with stages of disease or aging. Among these 46 human anatomical images, 14 were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) [31]. Among these 14 images in the field of neuroscience within human anatomy, the patient group's coronal and axial plane images, as well as the control group's coronal and axial plane images, were selected from an FD analysis paper in neuroscience [31]. An additional eight human anatomical images in a similar fashion were obtained from an FD study on neuroscience-related diseases [36]. These eight images are axial plane brain images of the left and right white matter in one human subject with a right-hemisphere lesion [36]. An additional 14 human anatomical images were obtained from an FD analysis by Wu et al. [37], of which seven of them show cerebellum white matter structures and the other seven show cerebellum gray matter structures. A supplementary five human anatomical images depicting brain visualizations of various patients were also procured [38]. Finally, the remaining six human anatomical images are of benign and malignant cases of breast cancer [39].

Another major component of the image dataset is a collection of other life science images, mainly consisting of those from plant biology, 20 of which are related to plant development [4]. Of these 20 images, 16 are specimens of fucus vesiculosus plants arranged in order of complexity, and four images are outlines of specimens of fucus serratus.

We also included some inorganic matter images in the dataset for this study. A total of 20 images of material surfaces and textures are sourced from the Brodatz Texture Dataset [40], from which these 20 images were deemed suitable (of the adequate resolution, size, fractal geometry, etc.) to use in this study. Another inorganic but still extensively naturally occurring fractal image type is landscape geography. Six

landscape images, in the form of generated environmental topographies from one of the randomized-phase modes in an inverse Fourier generation method, were also used for BC validation [41].

Moreover, 15 images of simply constructed sample contours of increasing perimeter irregularity were included in the dataset to validate the AI framework. In a similar vein, to ensure that the AI framework is still congruent with strictly formulaic fractals of mathematically defined FD, ten images of classical fractal figures are included, such as contours of the Gosper Island (Limit of the Gosper Curve) [18], Koch Curve (Koch Snowflake) [32], the boundary of the terdragon curve [42], etc. The diversity of this dataset provides concrete evidence that the proposed AI framework is a robust yet well-grounded tool that remains in agreement with the true benchmarks of the FD.

3.3.2 FD Distributions of the Entire Dataset

As shown in **Fig. 9** below, all FD values of 167 images fall between one and two, as mentioned earlier in Section 2. The distributions of FD values between different categories tend to overlap extensively, indicating that comparison within each category is more meaningful than between different categories. Given that all FDs fall between one and two, the difference between the maximum FD and the minimum FD within each group is at least 0.25 and at most 0.55. Most distributions seem to be relatively symmetric, such as those of interfacial science, life science, material surfaces and textures, and validation geometries. The distribution of formulaic fractals and geography are asymmetric, most likely since fewer images were included in the dataset. The input images selected were diverse in sources, yet the analysis methodology reported in this study suitably handled them all. Compared with methods mentioned in previous studies [4, 31, 36-39], which did not globally minimize the number of boxes used to cover the structure perimeter, our method eliminates QE to precisely and accurately estimate FDs, which are not affected by image rotation or translation. With our set-covering optimization-based BC approach in the AI framework, we eliminate subjective or user-varying protocols that would affect FD estimations.

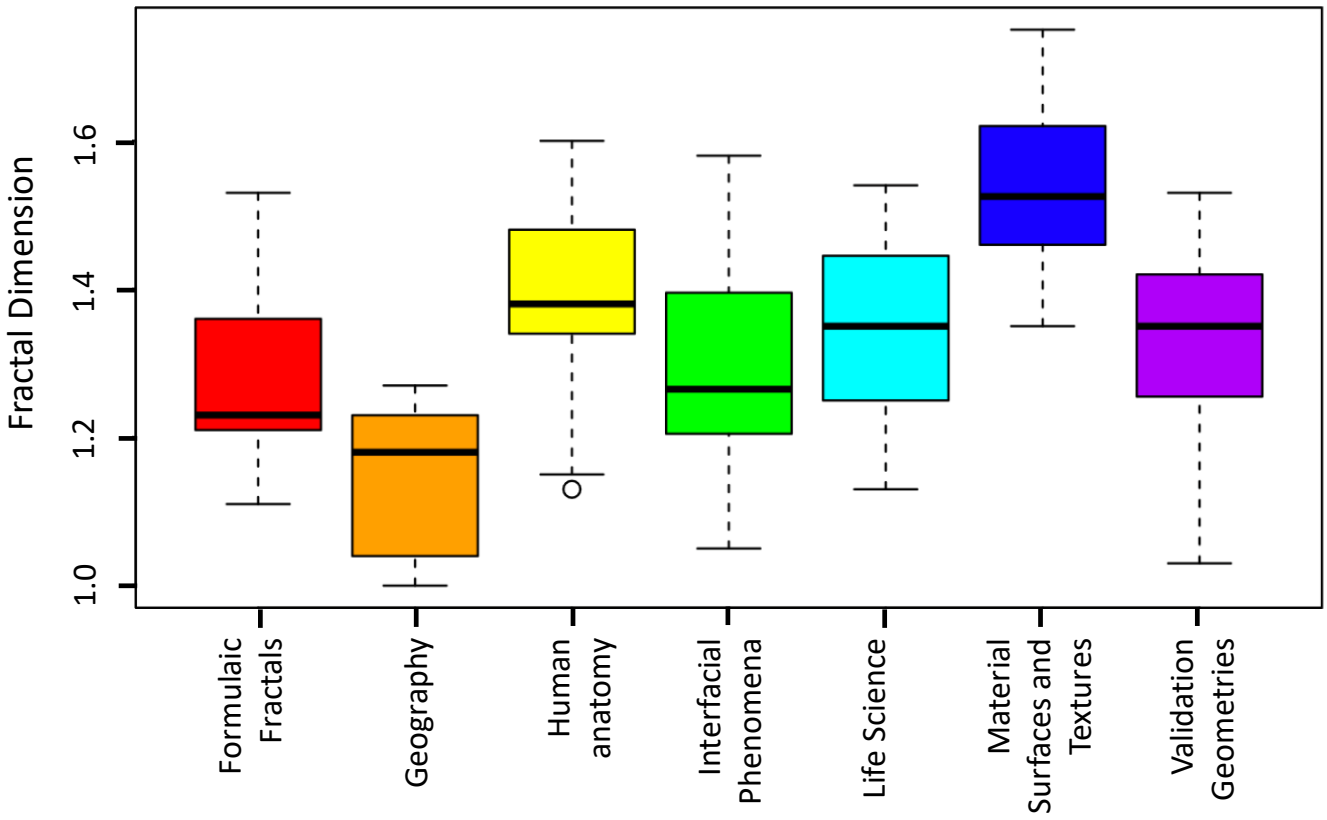

Fig. 9 Fractal dimension results of all seven fields of 167 images. Geography images are shown to have the lowest average fractal dimension, while material surfaces and textures have the highest average fractal dimension. All fractal dimension values fall between one and two.

3.3.3 Applicability to Large Dataset Applications

The AI framework lends itself to being fully automated so it can handle large batches of data. As shown in **Fig. 2**, the three step framework (image preprocessing, set-covering optimization-based BC, and regression analysis with logarithmic transformations) for evaluation of FD of input images is visualized, regarding how the AI framework automatically feeds subsequent information output from one major algorithmic step to the next. Within each major step, data (images and numerical results) are also processed and fed from the previous sub-step to the next. Since all procedures in the AI framework are conveniently connected, the framework can handle large batches of data simply by running the program for each image.

This may become a particularly valuable trait for any application wherein large sets of diverse image data must be handled simultaneously.

4. Conclusions

This paper proposed an automatic and systematic AI framework to estimate FD accurately without QE. The framework was proposed to integrate three major steps, including image preprocessing, set-covering optimization-based BC, and regression analysis. Each step was also divided into individual sub-steps, including image conversion from color to binary, image formatting, and perimeter extraction under image preprocessing. Additional substeps include ILP and box size range and value selection under set-covering optimization-based BC, and regression analysis with logarithmic transformations to obtain the final FD value. The framework was constructed to integrate each of the three major steps (and each sub-step within the corresponding step) systematically such that the resultant data/images/plots can be automatically formatted for appropriate output/input to the next step. Compared to the conventional methodologies, the AI framework guarantees consistency in the number of boxes minimized and counted for a given box size, so FD estimations remain uniform despite any arbitrary image rotation and translation. In the proposed AI framework, the set-covering optimization-based BC problem was formulated as an ILP problem to minimize the total number of box counts given each box size, and the box size combination selected in the end was improved by comparing and finding the ideal regression analysis metrics. Only one box count value is obtained, via the tailored optimization algorithm, for each box size of a given feature to guarantee consistency.

The validation dataset included images originating from seven different fields of study, wherein the FD of each image is estimated accurately enough to provide quantitative evaluation and comparison

between otherwise indiscernible image sets, attesting to both the framework's broad compatibility and discriminating quantification. Furthermore, the AI framework produces box data that very closely follows the linear behavior that defines FD, facilitating regression analysis, in that the MSE and MAE are typically negligible while R^2 values often remain >0.99. This is in part due to how the framework eliminates QE to precisely and accurately calculate FD, as was verified by the minimal difference (<4%) between the Hausdorff dimension and FD values of enclosed, formulaically generated true fractals. As a testament to the framework's discerning analytical capability, a sample study to discriminate FD between similar images demonstrated that the FD estimation results of Curli-spiked vs. Lb images exhibit a considerable difference between their average FD values. Ultimately, the framework can be effectively used on a wide variety of images to obtain FD estimations in a consistent and uniform way. The proposed AI framework demonstrates high computational efficiency, robustness in application, flexibility, and versatility for processing diverse types of images. The method shows a warranted potential to be developed into a 3-dimensional BC method in which the differential gray-scale can be incorporated. Most of the FD estimations we obtained fall in the range of 1.0 to 1.8, similar with previous studies' FD estimations generated by 2D box counting methods [6, 43, 44]. A future extension of this study can be conducted to theoretically investigate the reason for an upper bound of 1.8 of FD values of Chemical Engineering-related structures, but that is beyond the scope of this current paper.

References

[1] K. Falconer, Fractal geometry: mathematical foundations and applications, John Wiley & Sons2004.

[2] M.F. Barnsley, Fractals everywhere, Academic press2014.

[3] E. Fernandez, H.F. Jelinek, Use of fractal theory in neuroscience: methods, advantages, and potential problems, Methods 24(4) (2001) 309-321.

[4] J.D. Corbit, D.J. Garbary, Fractal dimension as a quantitative measure of complexity in plant development, Proceedings of the Royal Society of London. Series B: Biological Sciences 262(1363) (1995) 1-6.

[5] O.M. Bruno, R. de Oliveira Plotze, M. Falvo, M. de Castro, Fractal dimension applied to plant identification, Information Sciences 178(12) (2008) 2722-2733.

[6] X. Wang, P. Yang, F. Mondiot, Y. Li, D.S. Miller, Z. Chen, N.L. Abbott, Interfacial ordering of thermotropic liquid crystals triggered by the secondary structures of oligopeptides, Chemical Communications 51(94) (2015) 16844-16847.

[7] R.N. Rambaran, L.C. Serpell, Amyloid fibrils: abnormal protein assembly, Prion 2(3) (2008) 112-117.

[8] A. Maltais, G.E. Remondetto, M. Subirade, Mechanisms involved in the formation and structure of soya protein cold-set gels: A molecular and supramolecular investigation, Food Hydrocolloids 22(4) (2008) 550-559.

[9] M. Muthukumar, H.H. Winter, Fractal dimension of a crosslinking polymer at the gel point, Macromolecules 19(4) (1986) 1284-1285.

[10] A. Macías-García, M. Diaz-Diez, E. Cuerda-Correa, M. Olivares-Marin, J. Ganan-Gomez, Study of the pore size distribution and fractal dimension of HNO3-treated activated carbons, Applied surface science 252(17) (2006) 5972-5975.

[11] V. Gómez-Serrano, E. Cuerda-Correa, M. Fernandez-Gonzalez, M. Alexandre-Franco, A. Macias-Garcia, Preparation of activated carbons from chestnut wood by phosphoric acid-chemical activation. Study of microporosity and fractal dimension, Materials Letters 59(7) (2005) 846-853.

[12] L. de Martín, A. Fabre, J.R. Van Ommen, The fractal scaling of fluidized nanoparticle agglomerates, Chemical Engineering Science 112 (2014) 79-86.

[13] Y.-H. Lee, J. Carr, D. Barr, C. Haas, The fractal dimension as a measure of the roughness of rock discontinuity profiles, International journal of rock mechanics and mining sciences & geomechanics abstracts, Elsevier, 1990, pp. 453-464.

[14] B.B. Mandelbrot, B.B. Mandelbrot, The fractal geometry of nature, WH freeman New York1982.

[15] B.J. West, A.L. Goldberger, Physiology in fractal dimensions, American Scientist 75(4) (1987) 354-365.

[16] A.K. Bisoi, J. Mishra, On calculation of fractal dimension of images, Pattern Recognition Letters 22(6-7) (2001) 631-637.

[17] J. Gagnepain, C. Roques-Carmes, Fractal approach to two-dimensional and three-dimensional surface roughness, wear 109(1-4) (1986) 119-126.

[18] M. Bouda, J.S. Caplan, J.E. Saiers, Box-counting dimension revisited: presenting an efficient method of minimizing quantization error and an assessment of the self-similarity of structural root systems, Frontiers in plant science 7 (2016) 149.

[19] G.-B. So, H.-R. So, G.-G. Jin, Enhancement of the box-counting algorithm for fractal dimension estimation, Pattern Recognition Letters 98 (2017) 53-58.

[20] M.H. Chowdhury, W.D. Little, Image thresholding techniques, IEEE pacific Rim conference on communications, computers, and signal processing. Proceedings, IEEE, 1995, pp. 585-589.

[21] A. Conci, C.B. Proença, A fractal image analysis system for fabric inspection based on a box-counting method, Computer Networks and ISDN Systems 30(20-21) (1998) 1887-1895.

[22] C. Panigrahy, A. Seal, N.K. Mahato, Image texture surface analysis using an improved differential box counting based fractal dimension, Powder Technology 364 (2020) 276-299.

[23] S. Buczkowski, P. Hildgen, L. Cartilier, Measurements of fractal dimension by box-counting: a critical analysis of data scatter, Physica A: Statistical Mechanics and its Applications 252(1-2) (1998) 23-34.

[24] P. Dey, S.K. Mohanty, Fractal dimensions of breast lesions on cytology smears, Diagnostic cytopathology 29(2) (2003) 85-86.

[25] S. Buczkowski, S. Kyriacos, F. Nekka, L. Cartilier, The modified box-counting method: analysis of some characteristic parameters, Pattern Recognition 31(4) (1998) 411-418.

[26] D.P. Popescu, C. Flueraru, Y. Mao, S. Chang, M.G. Sowa, Signal attenuation and box-counting fractal analysis of optical coherence tomography images of arterial tissue, Biomedical optics express 1(1) (2010) 268-277.

[27] M. Long, F. Peng, A box-counting method with adaptable box height for measuring the fractal feature of images, Radioengineering 22(1) (2013) 208-213.

[28] C. Panigrahy, A. Seal, N.K. Mahato, Quantitative texture measurement of gray-scale images: Fractal dimension using an improved differential box counting method, Measurement 147 (2019) 106859.

[29] B. Klinkenberg, A review of methods used to determine the fractal dimension of linear features, Mathematical Geology 26(1) (1994) 23-46.

[30] K. Foroutan-pour, P. Dutilleul, D.L. Smith, Advances in the implementation of the box-counting method of fractal dimension estimation, Applied mathematics and computation 105(2-3) (1999) 195-210.

[31] R.D. King, A.T. George, T. Jeon, L.S. Hynan, T.S. Youn, D.N. Kennedy, B. Dickerson, Characterization of atrophic changes in the cerebral cortex using fractal dimensional analysis, Brain imaging and behavior 3(2) (2009) 154-166.

[32] M. Alizadeh, K. Hosseinzadeh, H. Mehrzadi, D. Ganji, Investigation of LHTESS filled by Hybrid nano-enhanced PCM with Koch snowflake fractal cross section in the presence of thermal radiation, Journal of Molecular Liquids 273 (2019) 414-424.

[33] K.J. Falconer, The Hausdorff dimension of self-affine fractals, Mathematical Proceedings of the Cambridge Philosophical Society, Cambridge University Press, 1988, pp. 339-350.

[34] A. Napolitano, S. Ungania, V. Cannata, Fractal dimension estimation methods for biomedical images, MATLAB–a fundamental tool for scientific computing and engineering applications 3 (2012) 161-178.

[35] C. Bandt, N.V. Hung, Fractal n-gons and their Mandelbrot sets, Nonlinearity 21(11) (2008) 2653.

[36] L. Zhang, G.H. Yue, Fractal dimension studies of the brain shape in aging and neurodegenerative diseases, The fractal geometry of the brain (2016) 213-232.

[37] Y.-T. Wu, K.-K. Shyu, C.-W. Jao, Z.-Y. Wang, B.-W. Soong, H.-M. Wu, P.-S. Wang, Fractal dimension analysis for quantifying cerebellar morphological change of multiple system atrophy of the cerebellar type (MSA-C), Neuroimage 49(1) (2010) 539-551.

[38] G. Reishofer, K. Koschutnig, C. Enzinger, F. Ebner, H. Ahammer, Fractal dimension and vessel complexity in patients with cerebral arteriovenous malformations, PloS one 7(7) (2012) e41148.

[39] M. Etehad Tavakol, C. Lucas, S. Sadri, E. Ng, Analysis of breast thermography using fractal dimension to establish possible difference between malignant and benign patterns, Journal of Healthcare Engineering 1(1) (2010) 27-43.

[40] I. Hersey, Textures: A Photographic Album for Artists and Designers by Phil Brodatz, Leonardo 1(1) (1968) 91-92.

[41] A.W. Juliani, A.J. Bies, C.R. Boydston, R.P. Taylor, M.E. Sereno, Navigation performance in virtual environments varies with fractal dimension of landscape, Journal of environmental psychology 47 (2016) 155-165.

[42] D.H. Werner, W. Kuhirun, P.L. Werner, Fractile arrays: A new class of tiled arrays with fractal boundaries, IEEE Transactions on Antennas and Propagation 52(8) (2004) 2008-2018.

[43] R.B. Moruzzi, A.L. de Oliveira, F.T. da Conceição, J. Gregory, L.C. Campos, Fractal dimension of large aggregates under different flocculation conditions, Science of the Total Environment 609 (2017) 807-814.

[44] T. Li, Z. Zhu, D. Wang, C. Yao, H. Tang, The strength and fractal dimension characteristics of alum–kaolin flocs, International Journal of Mineral Processing 82(1) (2007) 23-29.