Acta Geotechnica
https://doi.org/10.1007/s11440-018-0759-x

RESEARCH PAPER

Reconstructing granular particles from X-ray computed tomography using the TWS machine learning tool and the level set method

Zhengshou Lai[1] · **Qiushi Chen**[1]

Received: 21 September 2017 / Accepted: 20 November 2018

Abstract

X-ray computed tomography (CT) has emerged as the most prevalent technique to obtain three-dimensional morphological information of granular geomaterials. A key challenge in using the X-ray CT technique is to faithfully reconstruct particle morphology based on the discretized pixel information of CT images. In this work, a novel framework based on the machine learning technique and the level set method is proposed to segment CT images and reconstruct particles of granular geomaterials. Within this framework, a feature-based machine learning technique termed Trainable Weka Segmentation is utilized for CT image segmentation, i.e., to classify material phases and to segregate particles in contact. This is a fundamentally different approach in that it predicts segmentation results based on a trained classifier model that implicitly includes image features and regression functions. Subsequently, an edge-based level set method is applied to approach an accurate characterization of the particle shape. The proposed framework is applied to reconstruct three-dimensional realistic particle shapes of the Mojave Mars Simulant. Quantitative accuracy analysis shows that the proposed framework exhibits superior performance over the conventional watershed-based method in terms of both the pixel-based classification accuracy and the particle-based segmentation accuracy. Using the reconstructed realistic particles, the particle-size distribution is obtained and validated against experiment sieve analysis. Quantitative morphology analysis is also performed, showing promising potentials of the proposed framework in characterizing granular geomaterials.

Keywords 3D particle morphology · Level set · Machine learning · Shape reconstruction · X-ray computed tomography

1 Introduction

The study of particle morphological features of granular geomaterials, including their form, sphericity, roundness, and roughness, has been a subject of interest in the geotechnical engineering and geomechanics community for decades. These morphological features can help researchers and engineers understand the forming, weathering, and aging process of geomaterials [22, 43]. They are also among the fundamental and the most salient factors that govern the material's macroscopic properties and engineering behavior, such as compressibility, shear strength, and critical state parameters [24, 47, 52, 56, 66]. Moreover, the understanding of the link between particle morphology and its engineering behavior is of importance to design and optimize innovative geomaterials, such as bio-improved and bio-cemented granular soils [10, 14, 15, 23, 55].

Experimental techniques commonly used to obtain the morphological features of granular geomaterials include photography [66], scanning electron microscopy [12], and X-ray computed tomography (CT) [16]. The first two techniques provide two-dimensional (2D) morphological information (i.e., 2D images) that can be used to approximate or infer three-dimensional (3D) morphological features, e.g., through a virtual 3D surface method [41] or through enhancing the illusion of the depth of 2D images [65]. The X-ray CT, on the other hand, can be used to directly obtain 3D morphological information, and therefore, has been the most prevalent technique in recent years [2, 20, 26, 39, 57, 59, 67]. The particle-level morphological information can be integrated into numerical methods, such as the discrete element method (DEM) [13], to develop

---

✉ Qiushi Chen
qiushi@clemson.edu

[1] Glenn Department of Civil Engineering, Clemson University, Clemson, SC 29634, USA

more realistic and predictive numerical models for granular geomaterials [3, 31, 39, 64].

1.1 Motivation of the proposed framework

Reconstructing particle morphology from discretized pixel information of X-ray CT images poses three main challenges. The first challenge is to classify pixel in a raw CT image into different material constituents or phases (e.g., soil particles, voids). Thresholding is a commonly used method for such a purpose in image processing, which has been successfully applied to separate objects from the background [9]. The performance and effectiveness of the thresholding method, however, is greatly complicated by various factors, such as the nonstationary and correlated noise, ambient illumination, busyness of gray levels within the object and its background, inadequate contrast, and object size not commensurate with the scene [49]. Granular geomaterial is typically a composition of various minerals, organic matters, fluids, and internal voids. Each of the constituents has its intrinsic X-ray attenuation, leading to a large variance of pixel intensity in raw CT images [30]. In addition, the perturbation of tomography environment, the practical scanning procedure, and the limitation of CT optical apparatus could bring in significant noises that further aggravate the intensity variance [50]. As such, there is oftentimes no clear demarcation between the pixel intensity of different phases in geomaterials, which makes phase classification using the thresholding method very challenging. Some more advanced energy-based techniques for constituents classification, such as the graph cut and the region-based level set algorithms, are sensitive to the selection of weights on the various terms in the energy functional [1]. These weights are usually tuned beforehand by the developer via trial and error and can only achieve reasonable results for certain types of images.

The second challenge is to accurately identify particle boundaries and segregate particles in contact. Several well-known methods for object segregation include the watershed method [39], the concave curvature segmentation method [28, 61], the edge detection method [19, 59], and the region growing method [26]. Each of these methods has its strengths and limitations, and sometimes, several methods are used in combination to achieve better performance. The watershed method is perhaps the most widely used method but has the over-segmentation issue [40]. Its performance, to a great extent, depends on well-defined markers that are typically difficult to determine. The concave curvature segmentation method examines the curvature of the shape boundary and draws segmenting *split line* through the *split points* with a concave curvature, thus to split the contacting particles at the point of contact [28]. This method requires *a prior* shape information and might not be applicable to particles with concave surfaces as it would mistakenly split a particle if a concave surface is detected. The edge detection-based method utilizes some filters, such as the difference of Gaussian or the Laplacian of Gaussian, to identify object boundaries. These filters alone usually do not produce definitive object boundaries, but may provide useful cues to be used in subsequent algorithms, such as the edge-based level set method [19, 59]. The edge-based level set method depends upon well-defined edge indicator, and its performance is sensitive to the initialization and model parameters.

The last challenge is to approach an accurate characterization of the particle shape based on the discretized pixel matrix of CT images. The marching cubes (MC) method proposed by Lorensen and Cline [36] is a popular isosurface algorithm to reconstruct particle surface. However, in most studies, the MC method operates on binarized images for surface reconstruction and the reconstructed surfaces generally have artificial stair steps [62]. Hence, some smoothing manipulations are necessary to remove the jagged stair steps [21]. A more accurate approach is to utilize the edge-based level set method, which is capable of achieving a sub-pixel accuracy of the object boundary [34]. As aforementioned, the edge-based level set method suffers from the edge indicator, initialization, and model parameter issues when handling image segmentation.

In this paper, a novel framework is developed to identify and reconstruct realistic 3D particle shapes from discretized pixel information of X-ray CT images. The proposed framework takes a fundamentally different approach compared to previous works in that a trainable machine learning technique is innovatively integrated with the edge-based level set method. The machine learning technique is utilized to segment raw CT images of granular geomaterials, i.e., to classify different constituents and to segregate particles in contact. In the machine learning process, the segmentation algorithms are not explicitly programmed. Instead, the method predicts segmentation results based on the weighted combination of various image features at different image scales [29]. The image features and regression functions are implicitly encoded in the classifier model and are determined through a training process. The machine learning-based method, with well-trained classifier model (e.g., the feature weights and regression functions), can provide logical and knowledge-based image segmentation results comparable to experienced engineers or to human's recognition and perception. It has already brought in some successful applications of image segmentation in the biology and medical areas [7, 51]. Once the raw images are segmented, edge indicators can be evaluated from the processed images. A 3D edge-based level set method is then developed to approach an accurate shape representation of real particles.

2 Overview of the framework

Figure 1 shows a schematic illustration of the proposed framework. In this figure, key components of the methodology are shown in the top row, and the expected outputs are listed in the bottom row. The framework starts with a specimen of granular soil scanned using an X-ray CT scanner to produce a set of 3D raw CT images. A machine learning tool termed Trainable Weka Segmentation (TWS) is then utilized to classify image pixels and segment particles in contact. The outputs of the TWS-based segmentation are probability maps showing the probability of the pixels belonging to a specified class (e.g., solid or void). Using the probability maps as inputs, a 3D edge-based level set method is implemented to capture particle boundaries and reconstruct realistic particles. These particles can be used for subsequent analysis such as characterizing shapes or particle-size distributions. In the following sections, the TWS-based segmentation, the level set method and the particle shape analysis will be presented in details.

3 Image segmentation by the TWS machine learning tool

In this section, the Trainable Weka Segmentation (TWS) [4] is introduced for image segmentation. TWS is a pixel-based image classification and segmentation tool, where each pixel in an image is treated as an individual object that possesses a vector of image features. A combination of image features is then used to classify a pixel into different classes. The basic idea of the machine learning-based approach is to regress the weights of different image features through a training process using manual annotations of image features. The result of the training process is a trained classifier model that can be applied to segment other similar image data.

TWS leverages the state-of-the-art machine learning algorithms provided in the data mining and machine learning toolkit Waikato Environment for Knowledge Analysis (Weka) [25]. TWS acts as a bridge between the image processing and machine learning toolkit, where it extracts the features of an image (e.g., an X-ray CT image) and converts them into the format that is expected by the Weka toolkit. By default, TWS uses random forest (RF) as the machine learning algorithm. In a recent study by Fernández-Delgado et al. [18], RF is shown to yield best overall performance and is recommended for new problems involving machine learning.

3.1 Decision tree and random forest

Random forest is built on an ensemble of decision trees. Decision trees are a nonparametric supervised learning method used for classification and regression. The goal is to create a flowchart-like structure that predicts the outcome of a target object (e.g., an X-ray image pixel) by learning simple decision rules inferred from the data features (e.g., various image features).

Figure 2 shows an example of a decision tree. In this example, there are 100 objects taken as training inputs, where each object possesses 3 features. Forty of the 100 objects are labeled as the foreground objects and 60 of them are labeled as the background objects. At the first level of division, the 100 objects are split into two groups based on the value of their second feature. The selection of a feature for the division is based on the information entropy theory [46], and the feature that achieves the most information gain after the division will be selected. According to the test results on the second feature, the 100 objects are classified into two groups. For instance, 10 of the 40 foreground objects fall into the first group as their second feature is smaller than 0.5, while 30 of them fall into the second group as their second feature is greater than 0.5. The second and third level of division follows the same

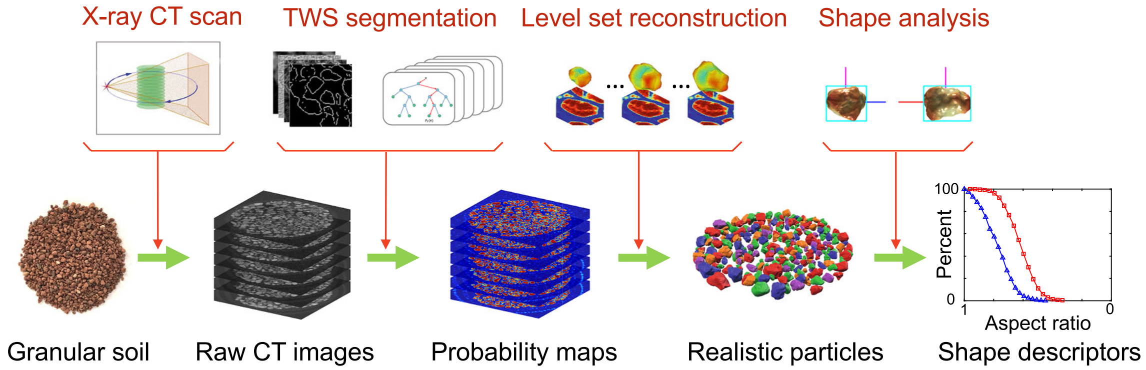


Fig. 1 Schematic illustration of the proposed framework. The top row includes key components of the methodology and the bottom row includes the expected outputs. *TWS* Trainable Weka Segmentation

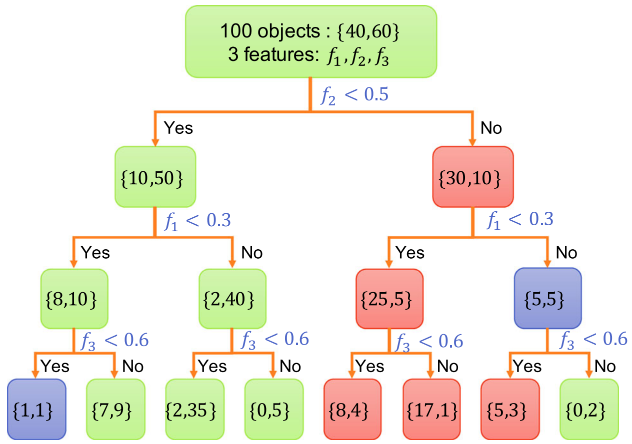


Fig. 2 Example of a decision tree with 3 levels of division. The features and division parameters are determined during the training process

strategy. Eventually, the 100 objects are classified into eight groups. For each of the 8 groups, if the foreground objects are dominating that group, it is considered as a foreground group and vice versa. The decision tree is constructed during the training process to determine the division parameters.

Once the decision tree is constructed, it can be used to predict the outcome of a new object. A new object will go through the branches of the decision tree following values of its features, eventually falling into an end group. The label of that group is the predicted outcome of the new object.

A single decision tree has the issue of over-fitting and is sensitive to the input data. Random forest is an ensemble of decision trees that can mitigate such disadvantages. Assuming that a training sample contains N objects and each object has M features, a decision tree can be constructed using a random subset of the N objects with a random subset of the M features. By repeating this process, one can obtain a set of decision trees. The prediction of a new object is based on the overall votes of all decision trees. Figure 3 shows an example of the random forest and its prediction strategy.

3.2 Image features

When applying the random forest algorithm to segmenting X-ray CT images, the objects are image pixels and the features are image features determined by the user. In this section, selected image features for segmentation are presented. Table 1 summarizes various image features grouped by their purpose that can be included in a classifier model. In this table, the noise reduction aims to mitigate the negative impact of intensity noises, e.g., those resulting from the CT scanning system. The texture filter is used to extract texture information. The edge indicator is used to detect object boundaries. Membrane detector is specialized in identifying membrane-like structures of a certain size and thickness.

A detailed explanation of different image features can be found in [4]. Five of the image features from Table 1 are selected for this study, and they are briefly explained in this section.

Gaussian blur	convolves an image with a Gaussian kernel distribution. It mitigates the noise and smooths the image. The standard deviation controls the shape of the Gaussian kernel, and a larger standard deviation increases the *blur* effect. Performing Gaussian blur with n different standard deviations results in a vector of n Gaussian blur features.
Sobel filter	is to approximate the gradient of an image using a finite difference scheme. In practice, the Sobel filter is usually applied after a prior Gaussian blur. n Gaussian blur gives a vector of n Sobel filter features.
Hessian	calculates the second derivatives of an image and results in a Hessian matrix at each pixel. The module, trace, determinant, first eigenvalue, second eigenvalue, orientation, Gamma-normalized square eigenvalue difference, and the square of Gamma-normalized eigenvalue difference will be used as Hessian features. A prior Gaussian blur is also applied before the Hessian operation. n Gaussian blur gives a vector of $8n$ Hessian features.
Difference of Gaussians	calculates two Gaussian blur images from the original image and subtracts one from the other. Performing difference of Gaussians with n different standard deviations gives a vector of $n(n-1)/2$ difference of Gaussians (DoGs) features.

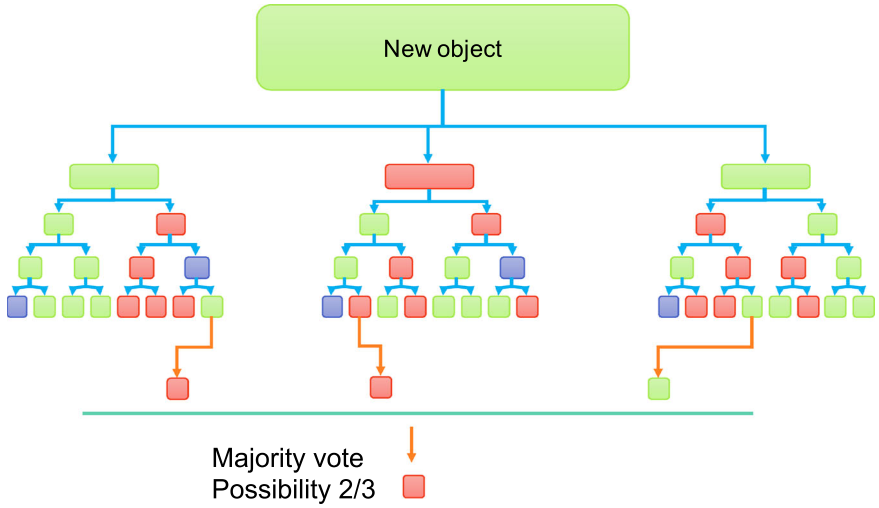


Fig. 3 Example of a random forest consisting of 3 decision trees (modified from Jaccard [29])

Table 1 Various image features and the corresponding purpose in the classifier model

Image features	Purpose
Gaussian blur, bilateral filter, anisotropic diffusion, Kuwahara, Lipschitz	Noise reduction
Minimum, maximum, median, variance, entropy, structure tensor	Texture filter
Laplacian, Hessian, Sobel filter, difference of Gaussian, Gabor	Edge indicator
Membrane projection	Membrane detector

Membrane projection first convolves an image with a set of hardcoded matrix kernel. The original matrix kernel is in $n \times n$ size, with the middle m columns set as 1s, and the remaining elements set as 0s. Then, the matrix kernel is rotated by 6° up to a total rotation of 180°, giving 30 kernels. The sum, mean, standard deviation, median, maximum and minimum of the 30 images are taken as the membrane projections features. The membrane-like structures in an image stand out after membrane projections.

An illustration of combining various images features into a classifier model is shown in Fig. 4. It should be noted that including more image features does not necessarily yield better segmentation results as more features require more training inputs to achieve a desirable classifier. Also, more image features in the classifier mean increased computational expenses. It is found in this work that these five image features yield satisfactory segmentation results.

3.3 Implementation of TWS for X-CT image segmentation

In this work, the process of using TWS to segment X-CT images are implemented in MATLAB. TWS is available as a plugin of Fiji (https://fiji.sc/). To be able to invoke Fiji and TWS methods in MATLAB, it is necessary to first install the Fiji-MATLAB interface *MIJ* (available from http://bigwww.epfl.ch/sage/soft/mij/). After installing the *MIJ* and adding the corresponding path into the MATLAB working path, the methods of TWS can be imported and invoked using MATLAB scripts. The interested reader is referred to the TWS user manual (available as the supplementary document of [4]) for more details of the TWS methods.

The workflow of the machine learning-based image segmentation is summarized in Fig. 5. The workflow consists of two main components: training classifier and applying classifier. In the training process, the raw X-CT images and their corresponding manually labeled images are required as inputs. The output is a trained classifier, where the parameters in each decision tree are obtained. By

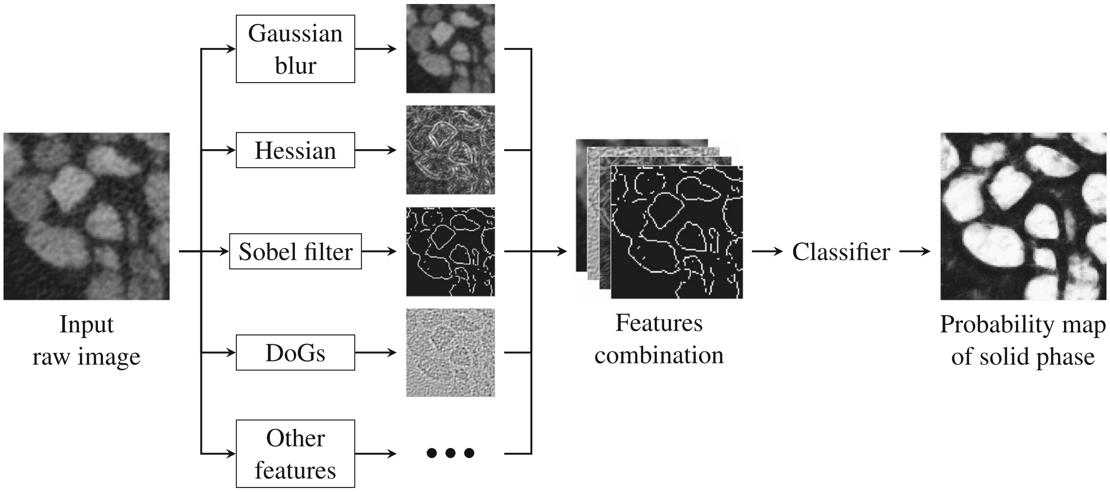


Fig. 4 An illustration of combining various images features into the classifier model in machine learning-based image segmentation process

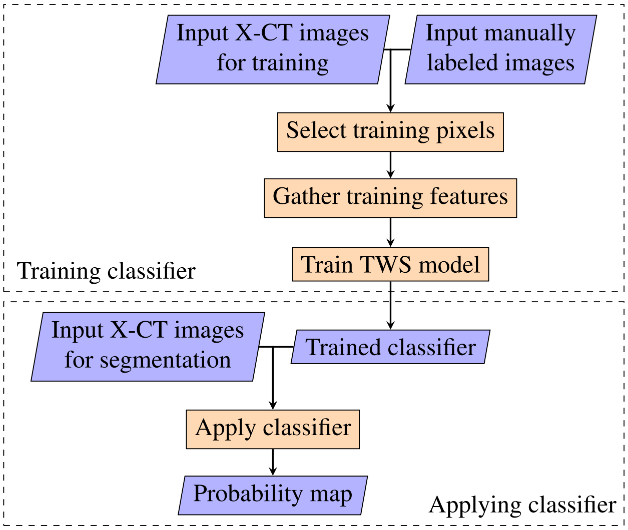


Fig. 5 The workflow of the X-CT image segmentation via TWS machine learning technique

applying the trained classifier to new X-CT images, the labels of each pixel in the new images can be predicted.

To manually label an image, the values of each pixel are set to either 1 or 0, indicating whether it belongs to the constituent the labeled image represents. An example set of input images is shown in Fig. 6. In this example, the image is intended to be classified into two phases (e.g., the solid phase and void phase as it is the case of this work). Two labeled images are required with one for the solid phase and one for the void phase. For the labeled image of the solid phase, the solid pixels are set to 1s, while the remaining pixels are set to 0s. To facilitate the training process and to reduce the workload of manual labeling, the TWS allows subset pixels of the raw image to be labeled and used as training inputs. After input of raw and labeled images, training features are gathered and the classifier is trained.

After the classifier is trained, the subsequent process is to apply the trained classifier to segment new CT images. TWS provides two types of results, i.e., a label matrix and a probability map. A label matrix consists of integer values for all pixels in the image, i.e., 1s indicating the pixels belong to the constituent the label matrix represents, and 0s indicating otherwise. A probability map is a map of probability values indicating the likelihood of a pixel belonging to a particular constituent and will be adopted in this work. Edge indicators will be evaluated based on the probability map and will be used for the subsequent level set method. Also, it is worth noting that though only two constituents (solid and void) are considered in this work, the implemented TWS technique is applicable to segment multiple classes of constituents.

4 The level set method for shape characterization and reconstruction

The probability maps from machine learning-based image segmentation are used as inputs for particle shape characterization and reconstruction using the level set method [42]. In this work, the edge-based level set method is adopted. The original formulation proposed by Li et al. [32, 33] is applied to 3D particles, and a new scaling coefficient is introduced to the edge indicator function. It should be noted that a 3D version of the Li et al. [33] formulation has also been previously implemented by Sun

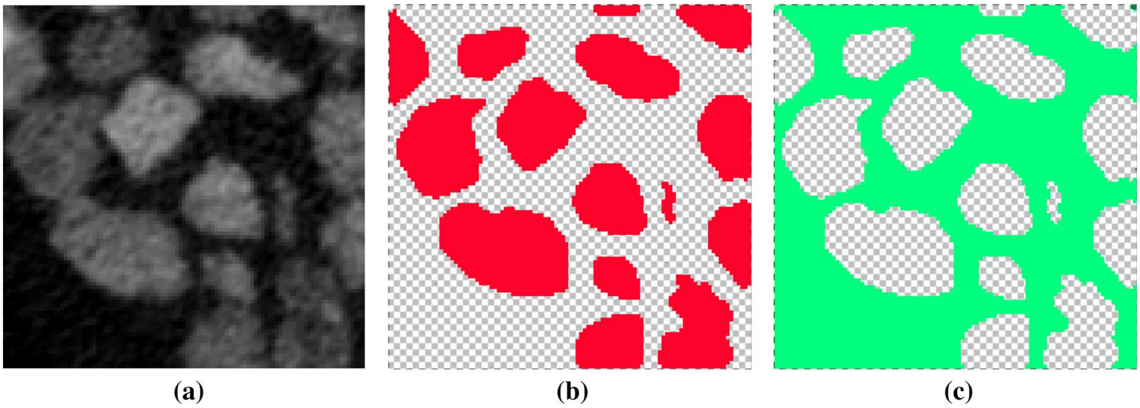


Fig. 6 An example set of input images for training classifier: **a** raw image, **b** the labeled image of the solid phase, and **c** the labeled image of the void phase. The pixels rendered by red and green color in the labeled images (**b**) and (**c**) are set to 1s, whereas the pixels rendered by the black–white pattern are set to 0s (color figure online)

et al. [53, 54], where the authors proposed a semi-implicit integration scheme and used the level sets to determine the 3D medial axes of pore microstructures.

4.1 Energy functional

The shape (or boundary) of a particle can be characterized by a closed surface in 3D (or a contour in 2D). The level set method aims to capture this closed surface using an auxiliary function termed the level set function, where the closed surface is defined as the zero level set, written as

$$\Gamma = \{(x, y, z) \in \Omega | \phi(x, y, z) = 0\} \tag{1}$$

where ϕ is the level set function; (x, y, z) are the spatial coordinates; Ω is the domain of interest and Γ is the closed surface (i.e., the boundary of the particle). Conventionally, the level set function ϕ is assumed to take positive values inside the region delimited by Γ and negative values outside.

When applying the level set method to identify the particle boundary, the level set function ϕ is evolved by minimizing an appropriate energy functional, denoted as $\mathcal{F}(\phi)$. In this work, the following formulation proposed by Li et al. [32] is adopted

$$\mathcal{F}(\phi) = \mu\mathcal{R}(\phi) + \lambda\mathcal{L}(\phi) + \nu\mathcal{A}(\phi) \tag{2}$$

where $\mathcal{R}(\phi)$ is the distance regularization term; $\mathcal{L}(\phi)$ and $\mathcal{A}(\phi)$ are the external energy related to the surface curvature and the inner volume, respectively; $\mu > 0$, $\lambda > 0$ and $\nu \in R$ are the weighting coefficients. A positive ν value drives the surface Γ inwards, while a negative value drives Γ outward. The energy terms are given by

$$\mathcal{R}(\phi) = \frac{1}{2}\int_\Omega (||\nabla\phi|| - 1)^2 \mathrm{d}\Omega \tag{3}$$

$$\mathcal{L}(\phi) = \int_\Omega g(I)\delta(\phi)||\nabla\phi||\mathrm{d}\Omega \tag{4}$$

$$\mathcal{A}(\phi) = \int_\Omega g(I)H(-\phi)\mathrm{d}\Omega \tag{5}$$

where H is the Heaviside function; δ is the Dirac delta function; $||\cdot||$ is the Euclidean norm; $g(I)$ is the edge indicator function defined by

$$g(I) := \frac{1}{1 + c||\nabla I||^2} \tag{6}$$

where I is the image matrix, and ∇ is the gradient operator. In this work, a scaling coefficient c is introduced to the conventional formulation of the edge indicator. In practice, it is found that this coefficient c can singularize the edge indicator (as shown in Fig. 7), thus stabilize and facilitate the convergence of level set evolution.

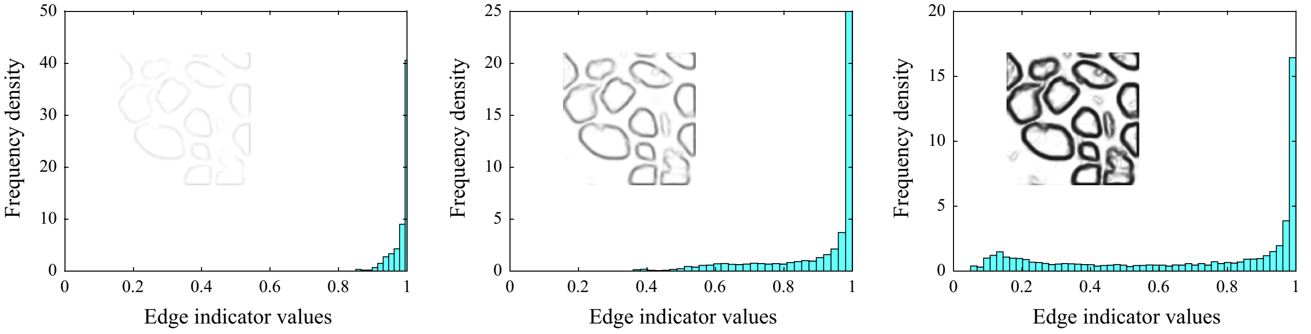
4.2 Solution of the level set function

Solutions of the level set function can be obtained by minimizing the energy functional $\mathcal{F}(\phi)$ in Eq. (2) with respect to ϕ, which leads to the associated Euler–Lagrange equation for ϕ [6]. Parameterizing the descent direction by an artificial time, the solution of the level set function ϕ can be obtained using [32, 33]

$$\frac{\partial\phi}{\partial t} = -\frac{\delta\mathcal{F}}{\delta\phi} = -\left(\mu\frac{\delta\mathcal{R}}{\delta\phi} + \lambda\frac{\delta\mathcal{L}}{\delta\phi} + \nu\frac{\delta\mathcal{A}}{\delta\phi}\right) \tag{7}$$

where $\delta[\cdot]/\delta\phi$ denotes the functional derivative of $[\cdot]$ with respect to ϕ. The functional derivative of each term in $\mathcal{F}$ with respect to ϕ is evaluated as

$$\frac{\delta\mathcal{R}}{\delta\phi} = \nabla\cdot\nabla\phi - \nabla\cdot\left(\frac{\nabla\phi}{||\nabla\phi||}\right) \tag{8}$$



(a) (b) (c)

Fig. 7 A demonstration of the edge indicator with different scaling coefficients: **a** $c = 1$, **b** $c = 10$, and **c** $c = 100$. The histograms indicate the distribution of the edge indicator values, and the inset represents the corresponding edge indicator image (color figure online)

$$\frac{\delta \mathcal{L}}{\delta \phi} = \delta(\phi)\left[\nabla g(I) \cdot \frac{\nabla \phi}{||\nabla \phi||} + g(I) \cdot \nabla\left(\frac{\nabla \phi}{||\nabla \phi||}\right)\right] \tag{9}$$

$$\frac{\delta \mathcal{A}}{\delta \phi} = g(I)\delta(\phi) \tag{10}$$

To solve the minimization problem defined by Eq. (7), a forward discretization scheme in time is implemented such that

$$\phi_{t+1} = \phi_t - \Delta t\left(\mu\frac{\delta \mathcal{R}}{\delta \phi} + \lambda\frac{\delta \mathcal{L}}{\delta \phi} + \nu\frac{\delta \mathcal{A}}{\delta \phi}\right)\Bigg|_t \tag{11}$$

where ϕ_{t+1} and ϕ_t are the level set function ϕ evaluated at the timestep $t+1$ and t, respectively; Δt is the timestep. As illustrated in Fig. 8, the evolution of the level set function ϕ starts with a given initial value ϕ_0 and stops when the energy functional arrives at a stationary state, which corresponds to the surface Γ represented by ϕ matching the particle shape boundary.

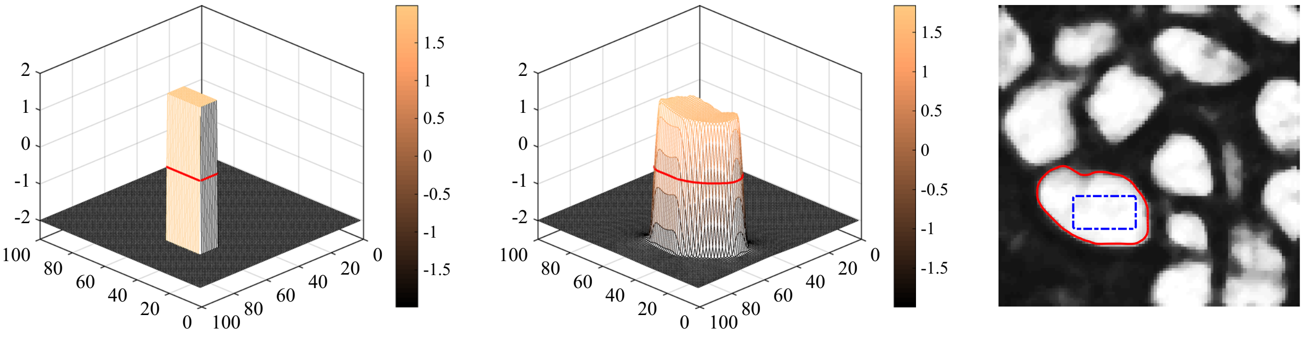
Solving the level set function also requires the evaluation of spatial derivatives in Eqs. (8–10), in particular, the evaluation of $\nabla \phi = [\nabla^x \phi, \nabla^y \phi, \nabla^z \phi]$. In this work, a finite difference scheme is implemented to discretize the spatial derivatives, which yields

$$\nabla^x \phi(i,j,k) = \begin{cases} \phi(2,j,k) - \phi(1,j,k), & \text{if } i = 1 \\ [\phi(i+1,j,k) - \phi(i-1,j,k)]/2, & \text{if } i = 2,3,\ldots,N_i - 1 \\ \phi(N_i,j,k) - \phi(N_i - 1,j,k), & \text{if } i = N_i \end{cases} \tag{12}$$

$$\nabla^y \phi(i,j,k) = \begin{cases} \phi(i,2,k) - \phi(i,1,k), & \text{if } j = 1 \\ [\phi(i,j+1,k) - \phi(i,j-1,k)]/2, & \text{if } j = 2,3,\ldots,N_j - 1 \\ \phi(i,N_j,k) - \phi(i,N_j - 1,k), & \text{if } j = N_j \end{cases} \tag{13}$$

$$\nabla^z \phi(i,j,k) = \begin{cases} \phi(i,j,2) - \phi(i,j,1), & \text{if } k = 1 \\ [\phi(i,j,k+1) - \phi(i,j,k-1)]/2, & \text{if } k = 2,3,\ldots,N_k - 1 \\ \phi(i,j,N_k) - \phi(i,j,N_k - 1), & \text{if } k = N_k \end{cases} \tag{14}$$

where i, j, and k indicate the pixel index for each dimension in the image matrix; N_i, N_j and N_k, are the dimension sizes, i.e., the total number of pixels in that dimension.



(a) (b) (c)

Fig. 8 An illustration of the level set function evolution: **a** the initial level set values, **b** the final level set values, and **c** the particle shape captured by the zero level set. In (**c**), the blue dashed line indicates the initial shape, and the red solid line indicates the final captured particle shape (color figure online)

The Dirac delta function $\delta(\phi)$ in Eqs. (9) and (10) is approximated by the following regularized function $\delta_\epsilon(\phi)$ as

$$\delta_\epsilon(\phi) = \begin{cases} \frac{1}{2\epsilon}\left[1+\cos\left(\frac{\pi\phi}{\epsilon}\right)\right], & \text{if } |\phi| \le \epsilon \\ 0, & \text{if } |\phi| > \epsilon \end{cases} \quad (15)$$

where ϵ is the regularization coefficient; $|\cdot|$ indicates the absolution value. The regularized $\delta_\epsilon(\phi)$ will converge to $\delta(\phi)$ as ϵ approaches 0.

By minimizing the energy functional, the zeroth isosurface of the level set function eventually approximates the surface of a particle. The marching cubes method can be used to reconstruct the zeroth isosurface from the level set matrix, which returns a triangle mesh representing the particle surface.

5 Particle morphology descriptors

Once particles are reconstructed using the level set method, various quantitative descriptors can be calculated to characterize the particle morphology. The three commonly used descriptors are the aspect ratio, sphericity, and roundness. The definition and calculation of these descriptors are presented in this section, while results of morphology analysis on the granular soil of interest are presented in Sect. 6.

5.1 Aspect ratio

The aspect ratio of a particle is the ratio of its sizes in different dimensions. It characterizes the elongation extent of a particle. A 3D particle possesses two independent aspect ratios. Standard practice assigns L to the longest dimension, with I being the longest dimension perpendicular to L and S being perpendicular to both L and I. The aspect ratios are calculated as the ratios between L, I and S. However, as pointed out by Blott and Pye [8], such definitions cannot give an appropriate description of certain shapes (e.g., cubes). An alternative is to use the Feret diameters to characterize the size of a particle along specific dimensions [17, 27]. There are infinite sets of Feret diameters for a given geometric shape. In this work, the three Feret diameters corresponding to the three principal axes are adopted. With the three Feret diameters denoted as D_1, D_2, and D_3, the aspect ratios are calculated as

$$a_{21} = \frac{D_2}{D_1} \quad (16)$$

$$a_{31} = \frac{D_3}{D_1} \quad (17)$$

where a_{21} is the aspect ratio of the medium Feret diameter (D_2) to the major Feret diameter (D_1); a_{31} is the aspect ratio of the minor Feret diameter (D_3) to the major Feret diameter (D_1). To evaluate the Feret diameters, the moment of inertia tensor is firstly calculated. A rotation matrix is obtained by converting the moment of inertia tensor to a diagonal matrix. The rotation matrix defines the directions of the principal axes of the particle. Then, the particle is rotated by multiplying the rotation matrix, so that the principal axes of the particle are aligned with the Cartesian coordinate axes. Finally, the range of the coordinates of all the vertexes on the particle surface would be the Feret diameters in each direction, respectively.

5.2 Sphericity

The sphericity describes the degree to which a 3D particle shape resembles a mathematically perfect sphere. This work adopts the 3D sphericity definition proposed by Wadell [60]

$$S = \frac{\left(36\pi V_\mathrm{p}^2\right)^{1/3}}{A_\mathrm{p}} \quad (18)$$

where S is the sphericity; V_p and A_p are the volume and surface area of the given particle, respectively.

5.3 Roundness

The definition or calculation of roundness is subjected to the most controversy. Most authors have accepted that roundness should refer to the relative sharpness of corners and edges of a particle rather than to the degree to which the overall outline of a particle approaches circularity (or sphericity in three dimensions). Wadell [60] first defined the 2D roundness as the average ratio of the curvature radius of all corners to the radius of the largest inscribed circle of a particle. Following the same logic, Zhou et al. [67] defined the 3D roundness, which is adopted in this work.

$$R = \frac{\Sigma g(k)|k_\mathrm{max}|^{-1}}{NR_\mathrm{ins}} \quad (19)$$

where R denotes the roundness; N is the total number of acceptable corners; R_ins is the radius of the maximum inscribed sphere of the particle; k_max is the maximum curvature at a corner; and $g(k)$ is a function indicates whether a corner is acceptable or not, defined as

$$g(k) = \begin{cases} 1 & \text{if } |k_\mathrm{max}|^{-1} < R_\mathrm{ins} \\ 0 & \text{if } |k_\mathrm{max}|^{-1} \ge R_\mathrm{ins} \end{cases} \quad (20)$$

In this work, the radius of the maximum inscribed sphere R_{ins} is approximated by the minimum radial distance of all vertexes on the particle surface to the particle centroid. The practical approach proposed by Colombo et al. [11] is adopted to evaluate the local principal curvatures at the corners of a particle. The interested reader is referred to [11, 67] for more details about the calculation.

6 Results and discussion

In this section, the proposed framework is applied to characterize and reconstruct 3D irregularly shaped particles from X-ray CT images of a particular granular material, the Mojave Mars Simulant (MMS). MMS is a Martian regolith simulant developed using a basalt mined in the western Mojave Desert and is among the suite of test rocks and soils that were used in the development of the 2007–2008 Phoenix Scout and the 2009 Mars Science Laboratory missions [44]. The type of MMS used in this study has particle sizes mainly ranging from 1 to 2 mm. The MMS sample is placed in a cylindrical container of 30 mm in diameter and 114 mm in length. An MILabs U-CT system with a resolution of 60 microns is used to obtain the raw CT image data, which yields a total of 500 × 500 × 1900 3D pixels data. To visualize the raw CT image data, the pixels are grouped into 1900 slices, each being an image of 500 × 500 pixels. Selected images from the middle are used to demonstrate the proposed framework.

6.1 Results by the proposed framework

To begin with, the *classifier* in the machine learning method is determined through a training process. In this process, TWS takes a raw X-ray CT image and two training images with labeled pixels as inputs, shown in Fig. 9. It should be pointed out that the performance of TWS depends on the quality of the training images. One can repeatedly amend the training images until satisfied with the classification results. The settings for TWS used in this study are summarized in Table 2.

The next process is to apply the trained *classifier* to the remaining raw X-ray CT images. For batch processing, this process is implemented in MATLAB through the Fiji-MATLAB interface *MIJ*, as described in Sect. 3. Figure 10 shows an example of the raw and TWS-processed images. As shown in Fig. 10b, the TWS-processed images are valued by the probability indicating the likelihood of a pixel belonging to a designated phase (i.e., the solid phase in this study). The probability map will then be used as inputs for level set-based shape reconstruction. The performance of level set method relies heavily on well-defined edge indicators. In practice, it is found that, by applying a prior TWS process, the resultant probability map can provide better edge indicators than the raw X-ray CT image. The corresponding edge indicator map of the example CT image is shown in Fig. 10c.

Table 3 summarizes the level set parameters used in this study. The energy coefficients μ, λ and ν, and the regularization coefficient ϵ are selected following the previous application and recommendation by Li et al. [32]. The scaling coefficient c in the edge indicator equation Eq. (6) is chosen such that the edge indicator values are scaled between 0 and 1. To obtain the initial level set for a particle to be reconstructed, one option is to adopt the "erosion and labeling" morphological image processing method, where pixels with the same label after the erosion process are considered to belong to the same particle and could be used as the initial level set. However, it is found that this option cannot provide good results for images with complex features as those of MMS considered in this work. Thus, the other option is to label a cuboid inside a particle manually. The cuboid is taken as the initial level set and gradually

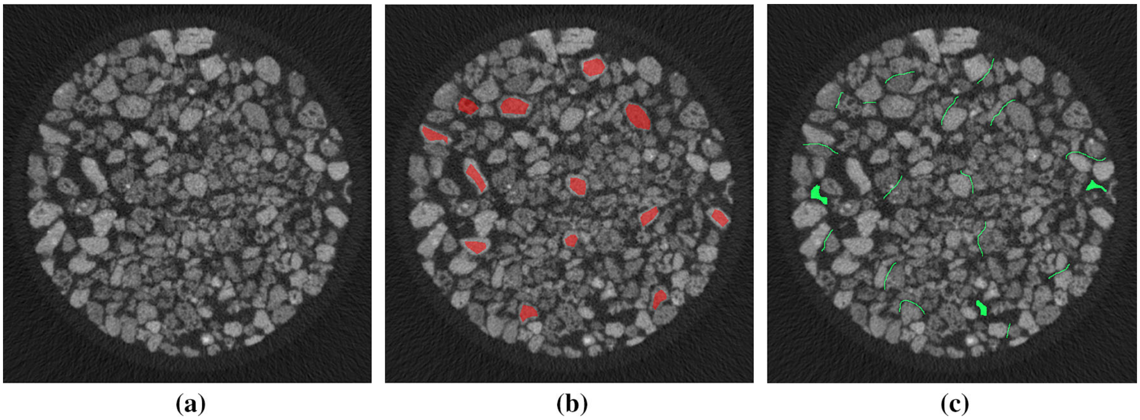
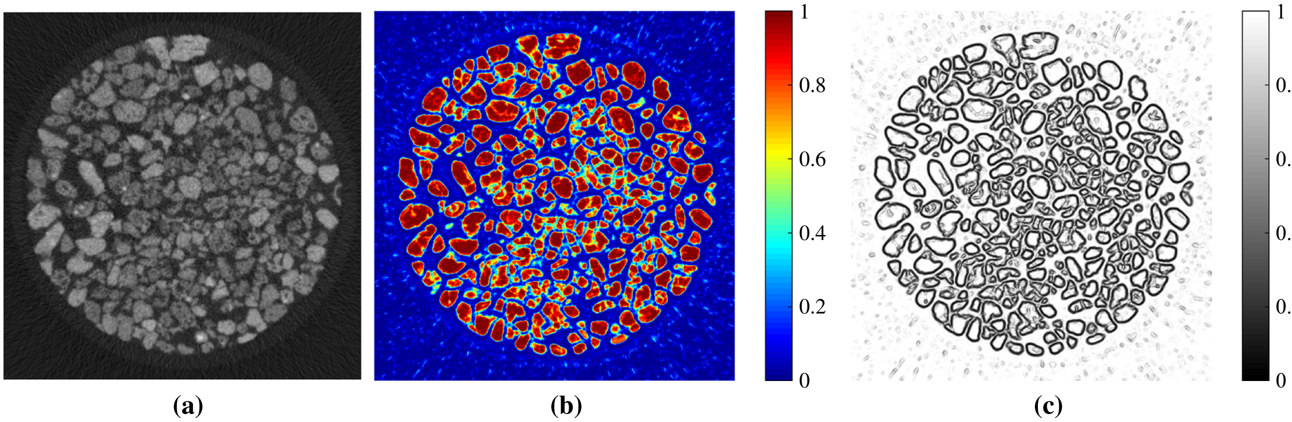

(a) (b) (c)

Fig. 9 Raw and training X-ray CT images: **a** the raw image, **b** the first training image with labeled solid phase (solid pixels are highlighted in red), and **c** the second training image with labeled void phase (void pixels are highlighted in green) (color figure online)

Table 2 Settings for TWS used in the present study

Entry/feature	Parameter	Value
Gaussian blur	Standard deviation	1, 2, 4, 8, 16
Sobel filter	Standard deviation	1, 2, 4, 8, 16
Hessian	Standard deviation	1, 2, 4, 8, 16
Difference of Gaussian	Standard deviation	1, 2, 4, 8, 16
Membrane parameters	Thickness	1
	Patch size	19
Random forest	No. of decision trees	200
	No. of random features per tree	2



(a)

(b)

(c)

Fig. 10 Example of **a** the raw X-ray CT image, **b** the probability map, and **c** the corresponding edge indicator calculated based on the processed image. The scaling coefficient for the edge indicator is $c = 100$

Table 3 Level set parameters used in the present study

Δt	μ	λ	v	ϵ	c
1	0.2	5.0	1.5	0.5	100

expands toward the particle boundary as level set evolves. An illustration of the surface evolution of one particle by the edge-based level set method is shown in Fig. 11. The final 3D surface indicates the surface of the reconstructed particle.

The level set process is applied to all particles in the processed CT images. In this work, the particle shape is represented and visualized using the surface triangle mesh, which is built upon the level set matrices using the marching cubes method [36]. There are other methods to

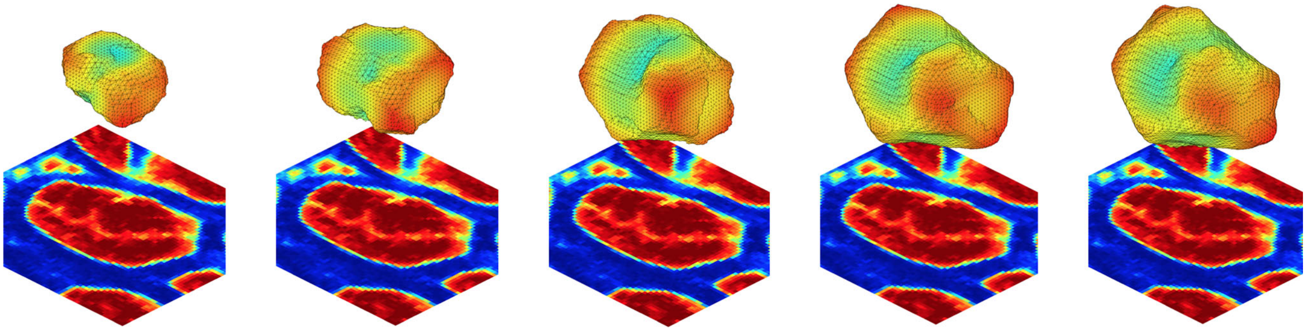


Fig. 11 An illustration of the surface evolution of one particle by the edge-based level set method at the 100th, 200th, 300th, 400th, and 500th iterations. The 2D image, colored by the probability values evaluated by the TWS method, shows the part of the CT image containing the target particle to be constructed. The final 3D surface indicates the surface of the reconstructed particle (color figure online)

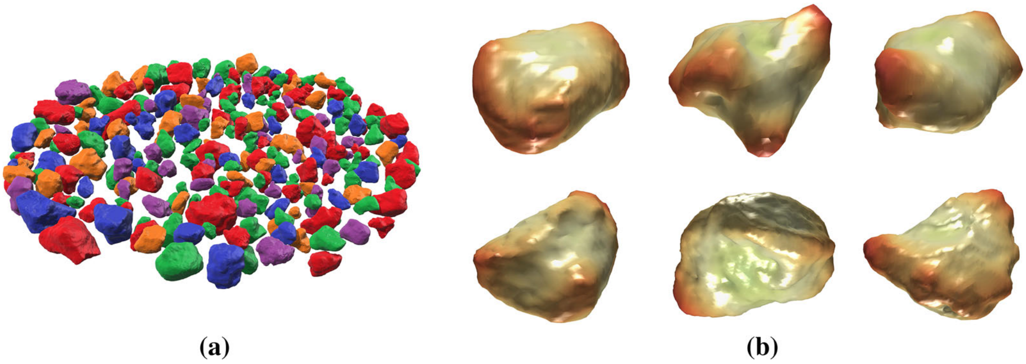


Fig. 12 Showcases of **a** one layer of the reconstructed 3D particles, and **b** zoom-in views of several particles. Particles in (**a**) are rendered with different colors for better visualization

represent a complex geometry, such as the orthogonal decomposition, Fourier descriptor [37], and the non-uniform rational basis spline [3]. Figure 12 presents one layer of the reconstructed 3D particles and the zoom-in view of several particles. It can be seen that particles of different sizes are successfully reconstructed. Some of the particles are relatively round, whereas others exhibit pretty sharp edges. The roughness of the particle surface is also well preserved, and some concave pits are captured on the particle surface. Quantitative shape and size analysis on the reconstructed particles will be presented in the following sections.

6.2 Accuracy analysis and validation

To evaluate the performance of the proposed framework, accuracy analysis is conducted in this section. The same analysis is also conducted for a watershed-based method as a comparison. As an additional validation, the size distribution of the reconstructed particles is compared to that obtained from a laboratory sieve analysis.

6.2.1 Definitions of accuracy

Before defining the accuracy, a reference solution of image segmentation is required. In this work, a manually labeled image is taken as the reference solution (i.e., the ground truth). To manually label the image, the particle boundary is carefully traced using the magnetic lasso tool in *Adobe Photoshop*. After that, the region inside the boundary is colored using the paint bucket tool. This process is repeated until all particles are identified and painted in different colors. It should be pointed out that the manually labeled particle shape might still deviate slightly from its *real* shape due to human error. Such deviation, however, is only a couple of pixels and would be acceptable considering that the particle sizes are of 25 pixels on average in each dimension.

This study considers two types of accuracy: pixel-based classification accuracy and particle-based segmentation accuracy. The pixel-based classification accuracy indicates the possibility that a pixel is classified into the correct phase (i.e., solid phase and void phase in this work). To quantify it, the pixel accuracy index and random accuracy index proposed by Arganda-Carreras et al. [5] are adopted. In particular, the joint probability p_{ij} is first defined as the probability that a randomly chosen pixel belongs to phase i in S and phase j in T, where S denotes the predicted classification and T denotes the reference solution. Then, the pixel accuracy index V_{pixel} and the random accuracy index V_{rand} are calculated as

$$V_{\text{pixel}} = p_{ii} \tag{21}$$

$$V_{\text{rand}} = \frac{\Sigma_{ij} p_{ij}^2}{\alpha \Sigma_k s_k^2 + (1-\alpha)\Sigma_k t_k^2} \tag{22}$$

where $s_i = \Sigma_j p_{ij}$ and $t_j = \Sigma_i p_{ij}$, which is the probability of a randomly selected pixel belonging to phase i or j, respectively; α is a parameter indicating the weights of split and merge errors. By definition, V_{pixel} indicates how many pixels are correctly classified. V_{rand} quantifies the probability that two randomly chosen pixels belong to the same phase in both predicted classification S and reference T. V_{rand} should be close to one if S and T are similar. It should be noted that other indices have also been proposed and used for accuracy or error analysis. For instance, Semnani and Borja [48] used three measures to compare a simulated image with the true image: variograms, pixel-wise error histograms and visual comparisons, where continuous variables (i.e., gray values) are used as opposed to categorical variables (i.e., individual phases) adopted in the current study.

To evaluate the particle-based segmentation accuracy, the number of correctly identified particles are counted. As a particle consists of hundreds of pixels, in this study, a particle is considered to be correctly identified if more than

90% of its pixels are correctly predicted and grouped to the same particle. The 10% tolerance is set to account for the slight errors in manually labeling a particle and the acceptable tolerance in the reconstructed particle shapes. It is worth noting that, for the accuracy analysis, we are focusing on just one slice of the CT images as a demonstration. The particle shape reconstructed using the proposed framework is indeed three-dimensional, but only a cross-section of the shape is extracted and compared to the reference solution.

6.2.2 Compare with the watershed-based method

Watershed [58] is a common tool for image segmentation. Segmentation by the watershed method requires meaningful local extrema to initiate the watershed and an appropriate threshold to stop the watershed. Otherwise, the watershed process can result in serious over-segmentation. Some variants of the original watershed method have been proposed, aiming to mitigate the over-segmentation issue [39, 63]. The particular implementation of watershed method used in this work for the comparison is the interactive H-watershed [35], which is available as a plugin in Fiji. The interactive H-watershed consists of a threshold process to classify phases and a watershed process to segment objects. In the watershed process, it employs the so-called H-extrema as the local extrema to initiate the watershed, where the H-extrema are known to be robust to noise. The interactive H-watershed also provides an interactive way to explore local extrema and threshold, updating the resulting watershed on the fly.

The pixel-based classification accuracy of the interactive H-watershed is controlled by the threshold parameter. Histograms of pixel values for the raw CT image and the corresponding TWS-processed image are plotted in Fig. 13. Two peaks are observed in the histogram of the raw CT image (Fig. 13a), corresponding to the void phase and the solid phase, respectively. The distribution of pixel intensity can be approximated by two normal distributions, and it can be seen in Fig. 13a that the two distributions representing the void phase and the solid phase would have large areas of overlap. In this case, it becomes critical to determine a single threshold to classify the void phase and solid phase. A small change in the threshold value would result in significant changes in the resulting fraction of the void phase and the solid phase.

In the proposed TWS-based segmentation, the material phases are determined using both the pixel intensity values and various image features (e.g., those listed in Table 1). After the TWS process, each pixel is given a probability value indicating its likelihood to be a designated phase (e.g., the solid phase). The histogram of probability values of the corresponding TWS-processed image is shown in Fig. 13b. It is observed that the two peaks representing the void and solid phases are more separated apart. Such a profile indicates that most of the pixels can be classified with great confidence, leaving only a small portion of the pixels (i.e., those with probability values around 0.5) with lower confidence.

To illustrate the influence of the assumed threshold values, different threshold values are applied and the accuracy of the interactive H-watershed is summarized in Table 4. In TWS, pixels with a probability value greater than 0.5 (i.e., voted as the solid phase in the random forest process) are classified as the solid phase. With a reasonable choice of a threshold value, the interactive H-watershed method can achieve a good performance that is comparable with the TWS. However, as it will be shown later, a reasonable choice of threshold values does not assure good particle segmentation. A significant limitation with the thresholding is that it cannot resolve the noises existed in the void phase or solid phase. TWS, on the other hand, has shown to be much more effective when applied to low-resolution images with noise [38].

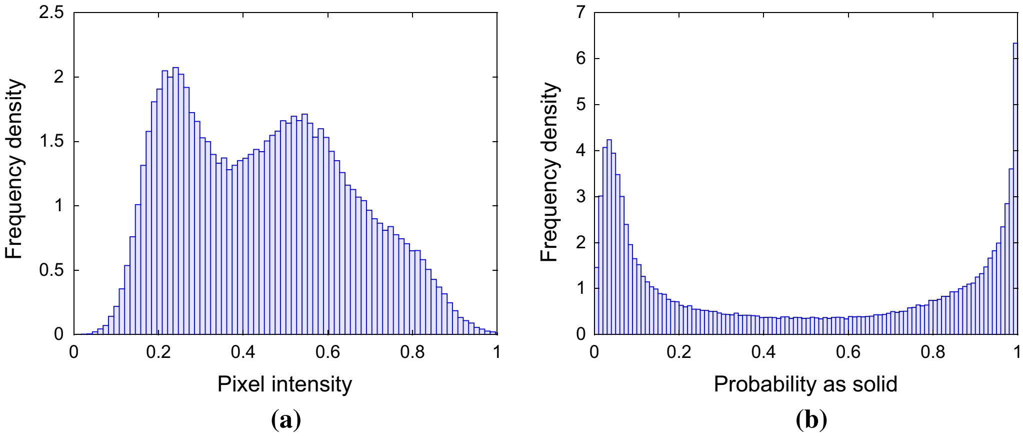


Fig. 13 Histograms of **a** a raw CT image, and **b** the corresponding TWS-processed image

Table 4 Pixel-based classification accuracy: interactive H-watershed versus TWS-LS

Method	V_{pixel}	V_{rand}
TWS-LS	0.86	0.76
Interactive H-watershed		
Threshold = 0.3	0.85	0.75
Threshold = 0.4	0.89	0.81
Threshold = 0.5	0.84	0.74

For the second type of accuracy, i.e., the particle-based segmentation accuracy, the interactive H-watershed mainly relies on the parameter called the H-extrema seed. Large seed values can help suppress local extreme used to initialize the watershed, thus minimizing the over-segmentation issue. However, an optimal choice of the seed parameter is problem dependent. The performances of the interactive H-watershed with different H-extrema seeds are investigated in this study. Table 5 summarizes the number of correctly segmented particles using the interactive H-watershed and the proposed TWS-LS method. For the interactive H-watershed, the number of correct particles significantly depends on the choice of the seed parameter. With regard to the particle-based segmentation accuracy, the TWS-LS shows superior performance. It is worth noting that, even if the TWS-LS cannot correctly reconstruct all the particles (i.e., reaching 100% accuracy), there will be statistically enough particles for the purpose of morphology characterization and particle-size distribution calculations.

To visualize the particle-based segmentation accuracy, the correctly and incorrectly segmented particles are shown in Fig. 14. In these figures, the incorrect particles are plotted as hollow shapes. For TWS-LS, most of the incorrect particles are the particles of smaller sizes. As aforementioned, this is partially due to the relatively insufficient CT resolution to resolve fine particles. For the interactive H-watershed, the incorrect particles have a wide size range, most of which are results from the issue of over-segmentation.

Table 5 Particle-based segmentation accuracy: interactive H-watershed versus TWS-LS

Method	Correct particles	Percent
TWS-LS	234	77.7
Interactive H-watershed		
Seed = 20	188	62.5
Seed = 25	163	54.2
Seed = 30	144	47.8

6.2.3 Validation with laboratory sieve analysis

In the laboratory sieve analysis, the following sieves in the Unified Soil Classification System are used: No. 4, 10, 20, 40, 60, 100, and 200. Based on the extracted particle shapes from CT image analysis, the particle sizes can be estimated as the diameter of their equivalent sphere (e.g., a sphere of the same volume). The particle-size distribution (PSD) curves obtained from the laboratory and the proposed TWS-LS framework are shown in Fig. 15. A reasonably good agreement is shown between the two curves. The PSD obtained from the proposed framework is slightly shifted to the left. Also, there is some discrepancy observed in particle sizes smaller than 0.5 mm. Those smaller particles are missing in the PSD obtained from the X-ray CT images as the resolution of the X-ray CT scanner cannot capture those very fine particles.

6.3 Particle morphology analysis

Having the particles fully reconstructed in 3D, quantitative morphology analyses are conducted using the morphology descriptor defined in Sect. 5. Figure 16 showcases the calculated morphology of several example 3D particles. Among these examples, the top-left particle is more elongated in one direction, corresponding to smaller values of aspect ratio (a_{21} and a_{31}). By comparison, the bottom-right particle has similar size in each direction, which results in larger values aspect ratio. It is also observed that the particles with more *angular corners* have smaller values of sphericity (S) and roundness (R), as is the case for the three particles in the bottom row.

To provide a statistical description of the particle morphology, the histogram of the aspect ratio, sphericity, and roundness from the reconstructed particles are plotted in Fig. 17. It is found that these descriptors can be roughly approximated by normal distributions. In this regard, their mean values and standard deviations are summarized in Table 6. As a comparison, the mean sphericity and roundness of the Leighton Buzzard sand are reported to be about 0.92 and 0.65 [67]. Leighton Buzzard sand is an English sand commonly used for academic research. It can be seen that the MMS particles, made from crashes of Saddleback Basalt, is more irregular than the Leighton Buzzard sand.

To classify the roundness grades of the MMS sample, the cumulative distribution of the roundness values by volume is displayed in Fig. 18. The Powers classification [45] of roundness is also shown. Results indicate that there are about 4% well-rounded particles, 95% rounded particles, and less than 1% subrounded particles, following the roundness grade table of Powers [45]. The detailed morphology analysis presented in this section is not possible

(a) (b) (c)

Fig. 14 Visual comparison of segmentation accuracy of the interactive H-watershed and the proposed TWS-LS method. Incorrectly identified particles are plotted as hollow shapes. **a** Ground truth, **b** TWS-LS, **c** interactive H-watershed

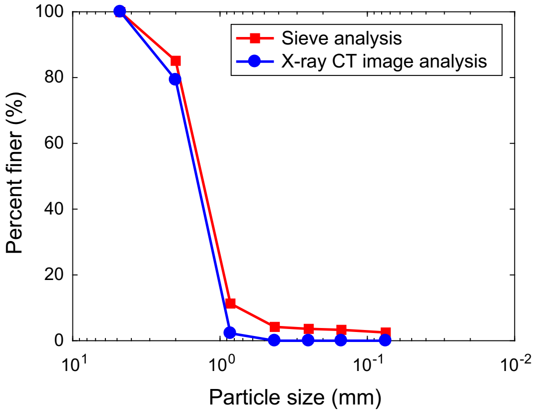


Fig. 15 Particle-size distribution

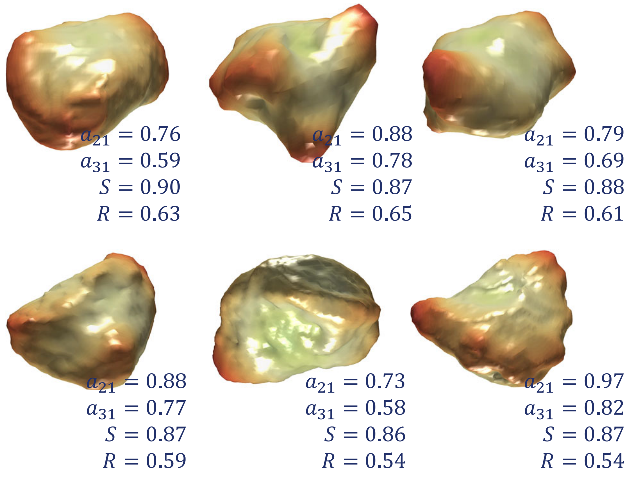


Fig. 16 Morphology of several example 3D particles

without the reconstructed particles using the proposed framework and the analysis can provide important insights into the microscopic features of the granular material.

7 Summary

In this work, a novel machine learning and level set-based framework is proposed to segment X-ray CT images of granular geomaterials and to reconstruct realistic 3D particle shapes. X-ray CT images of the Mojave Martian Simulant (MMS) show that the intensity of various constituents exhibit significant variance, and there is no clear demarcation between the solid and the void phases, making it particularly challenging for conventional binarization watershed-based methods.

To address this challenge, a feature-based machine learning technique termed the Trainable Weka Segmentation (TWS) is implemented and utilized to segment X-ray CT images. This is a fundamentally different approach in that it predicts segmentation results based on a trained classifier model that implicitly includes image features and regression functions. Probability maps indicating the likelihood of pixels belonging to a particular phase are obtained from the segmentation process, in which the original intensity contrast feature is to the most extent preserved. Compared to the intensity values of a raw CT image, it is found that the probability values exhibit much less variance and have a more distinct demarcation between different material phases. The probability map provides excellent edge indicators that can be used as the basis for the subsequent edge-based level set method.

Using the segmented X-ray CT images (i.e., probability maps), a 3D edge-based level set method is implemented to approach an accurate shape representation of real particles. Realistic 3D particles of the MMS are successfully reconstructed from raw CT images. Quantitative accuracy analyses are performed for the proposed framework and a conventional watershed method. The analyses show that the proposed framework has superior performance in both pixel-based classification accuracy and particle-based segmentation accuracy. The particle-size distribution using the

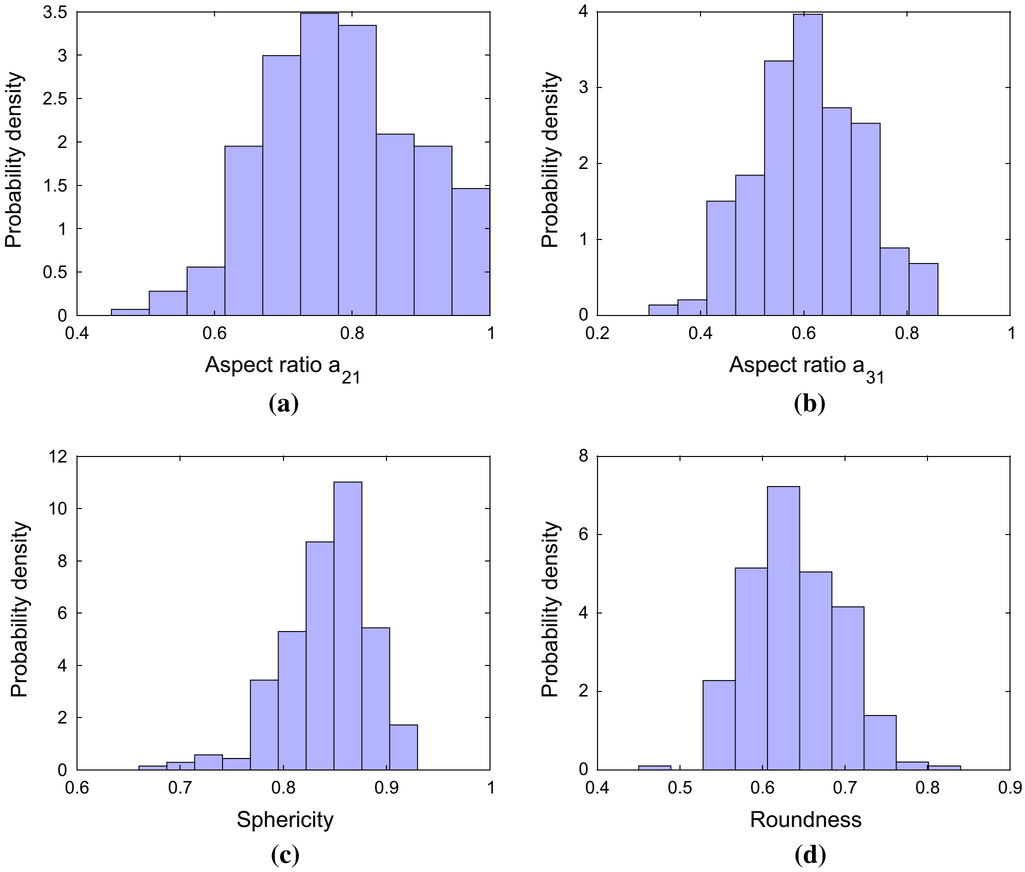


Fig. 17 Histogram and fitted distribution of the morphology descriptors of MMS soil particles: **a** aspect ratio a_{21}, **b** aspect ratio a_{31}, **c** sphericity and **d** roundness

Table 6 Statistics of the morphology descriptors of MMS soil particles

Descriptors	Mean	Standard
Aspect ratio a_{21}	0.78	0.11
Aspect ratio a_{31}	0.61	0.11
Sphericity S	0.84	0.04
Roundness R	0.64	0.05

reconstructed particles are also validated and compared well with results from a lab sieve analysis. In addition, the morphological features, e.g., sphericity, roundness, and roughness, of real particle shapes are well captured. Quantitative particle morphology analyses are performed to provide more insights into the morphological features of the granular material.

The machine learning technique has shown great potentials in segmenting CT images of geomaterials with complex constituents. Future work will expand the capability and application of machine learning-based particle identification. An immediate step would be to promote the current 2D TWS into the 3D regime, where image features from the third dimension are included in the segmentation process synchronously. Another interesting and important issue would be the effect of the quantity and quality of the training sets on the accuracy of the segmentation results. It

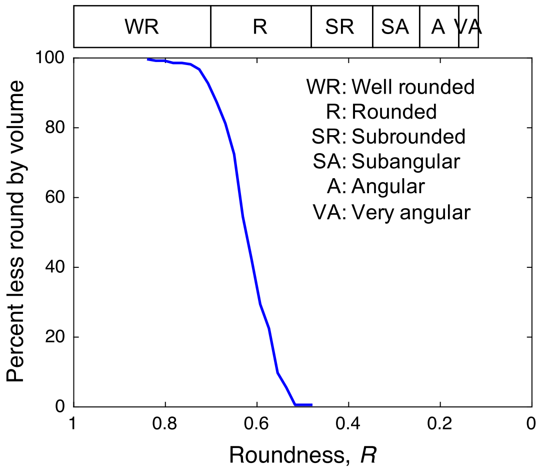


Fig. 18 Particle roundness distribution of the MMS sample with the Powers [45] classification

is also noticed that there are several competing machine learning-based image segmentation algorithms. It would be interesting to explore alternative methods and compare their performances when applied to granular geomaterials. Finally, the closed surface reconstructed by the edge-based level set method can be readily used for shape analysis and in a subsequent numerical model (e.g., the discrete element model) development.

Acknowledgements The authors would like to acknowledge the financial support provided by the NASA SC Space Consortium Grant (No. NNX15AL49H).

References

1. Al-Kofahi Y, Lassoued W, Lee W, Roysam B (2010) Improved automatic detection and segmentation of cell nuclei in histopathology images. IEEE Trans Biomed Eng 57(4):841–852
2. Andò E, Viggiani G, Hall S, Desrues J (2013) Experimental micro-mechanics of granular media studied by X-ray tomography: recent results and challenges. Géotech Lett 3(3):142–146
3. Andrade J, Vlahinić I, Lim K, Jerves A (2012) Multiscale 'tomography-to-simulation' framework for granular matter: the road ahead. Géotech Lett 2(3):135–139
4. Arganda-Carreras I, Kaynig V, Rueden C, Eliceiri K, Schindelin J, Cardona A, Sebastian Seung H (2017) Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification. Bioinformatics 33(15):2424–2426
5. Arganda-Carreras I, Turaga S, Berger D, Cireşan D, Giusti A, Gambardella L, Schmidhuber J, Laptev D, Dwivedi S, Buhmann J et al (2015) Crowdsourcing the creation of image segmentation algorithms for connectomics. Front Neuroanat 9(14):1–13
6. Aubert G, Kornprobst P (2006) Mathematical problems in image processing: partial differential equations and the calculus of variations, vol 147, 2nd edn. Springer, Berlin
7. Avendi M, Kheradvar A, Jafarkhani H (2016) A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. Med Image Anal 30:108–119
8. Blott S, Pye K (2008) Particle shape: a review and new methods of characterization and classification. Sedimentology 55(1):31–63
9. Bruchon J, Pereira J, Vandamme M, Lenoir N, Delage P, Bornert M (2013) Full 3D investigation and characterisation of capillary collapse of a loose unsaturated sand using X-ray CT. Granul Matter 15(6):783–800
10. Cheng L, Cord-Ruwisch R, Shahin M (2013) Cementation of sand soil by microbially induced calcite precipitation at various degrees of saturation. Can Geotech J 50(1):81–90
11. Colombo A, Cusano C, Schettini R (2006) 3d face detection using curvature analysis. Pattern Recognit 39(3):444–455
12. Cox M, Budhu M (2008) A practical approach to grain shape quantification. Eng Geol 96(1):1–16
13. Cundall P, Strack O (1979) A discrete numerical model for granular assemblies. Géotechnique 29(1):47–65
14. Dadda A, Geindreau C, Emeriault F, du Roscoat S, Garandet A, Sapin L, Filet A (2017) Characterization of microstructural and physical properties changes in biocemented sand using 3D X-ray microtomography. Acta Geotech 12(5):955–970
15. DeJong J, Soga K, Kavazanjian E, Burns S, Van Paassen L, Al Qabany A, Aydilek A, Bang S, Burbank M, Caslake L et al (2013) Biogeochemical processes and geotechnical applications: progress, opportunities and challenges. Géotechnique 63(4):287–301
16. Desrues J, Viggiani G, Besuelle P (2010) Advances in X-ray tomography for geomaterials, vol 118. Wiley, London
17. Ersoy A, Waller M (1995) Textural characterisation of rocks. Eng Geol 39(3–4):123–136
18. Fernández-Delgado M, Cernadas E, Barro S, Amorim D (2014) Do we need hundreds of classifiers to solve real world classification problems. J Mach Learn Res 15(1):3133–3181
19. Gao H, Chae O (2010) Individual tooth segmentation from CT images using level set method with shape and intensity prior. Pattern Recognit 43(7):2406–2417
20. Garboczi E (2011) Three dimensional shape analysis of JSC-1A simulated Lunar regolith particles. Powder Technol 207(1):96–103
21. Gibson S (1998) Constrained elastic surface nets: generating smooth surfaces from binary segmented data. In: Wells W, Colchester A, Delp S (eds) Medical image computing and computer-assisted intervention-MICCAI'98, vol 1496. Springer, Berlin, pp 888–898
22. Gilkes R, Suddhiprakarn A (1979) Biotite alteration in deeply weathered granite. I. Morphological, mineralogical, and chemical properties. Clays Clay Miner 27(5):349–360
23. Gleaton J, Xiao R, Lai Z, McDaniel N, Johnstone C, Burden B, Chen Q, Zheng Y (2018) Biocementation of martian regolith simulant with in-situ resources. In: Proceedings of the 2018 ASCE Earth and Space: engineering for extreme environments conference. ASCE
24. Guo P, Su X (2007) Shear strength, interparticle locking, and dilatancy of granular materials. Can Geotech J 44(5):579–591
25. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH (2009) The weka data mining software: an update. ACM SIGKDD Explor Newsl 11(1):10–18
26. Hashemi M, Khaddour G, François B, Massart T, Salager S (2014) A tomographic imagery segmentation methodology for three-phase geomaterials based on simultaneous region growing. Acta Geotech 9(5):831–846
27. Hentschel M, Page N (2003) Selection of descriptors for particle shape characterization. Particle Particle Syst Charact 20(1):25–38
28. Hobson D, Carter R, Yan Y (2009) Rule based concave curvature segmentation for touching rice grains in binary digital images. In: 2009 IEEE instrumentation and measurement technology conference, pp 1685–1689. IEEE
29. Jaccard N (2015) Development of an image processing method for automated, non-invasive and scale-independent monitoring of adherent cell cultures. PhD thesis, University College London
30. Ketcham R, Carlson W (2001) Acquisition, optimization and interpretation of X-ray computed tomographic imagery: applications to the geosciences. Comput Geosci 27(4):381–400
31. Lai Z, Chen Q (2017) Characterization and discrete element simulation of grading and shape-dependent behavior of JSC-1A Martian regolith simulant. Granul Matter 19(4):69
32. Li C, Xu C, Gui C, Fox M (2010) Distance regularized level set evolution and its application to image segmentation. IEEE Trans Image Process 19(12):3243–3254
33. Li C, Xu C, Gui C, Fox MD (2005) Level set evolution without re-initialization: a new variational formulation. In: IEEE Computer Society conference on computer vision and pattern recognition, 2005. CVPR 2005, vol 1, pp. 430–436. IEEE
34. Liu Y, Captur G, Moon J, Guo S, Yang X, Zhang S, Li C (2016) Distance regularized two level sets for segmentation of left and right ventricles from cine-MRI. Magn Reson Imaging 34(5):699–706
35. Lombardot B (2017) Interactive H-Watershed. https://imagej.net/Interactive_Watershed. Accessed 30 Apr 2018

36. Lorensen W, Cline H (1987) Marching cubes: a high resolution 3D surface construction algorithm. In: Proceedings of the 14th annual conference on computer graphics and interactive techniques, SIGGRAPH '87, New York, USA. ACM, pp 163–169
37. Luerkens D, Beddow J, Vetter A (1982) Morphological fourier descriptors. Powder Technol 31(2):209–215
38. Madra A, El Hajj N, Benzeggagh M (2014) X-ray microtomography applications for quantitative and qualitative analysis of porosity in woven glass fiber reinforced thermoplastic. Compos Sci Technol 95:50–58
39. Matsushima T, Katagiri J, Uesugi K, Tsuchiyama A, Nakano T (2009) 3D shape characterization and image-based DEM simulation of the lunar soil simulant FJS-1. J Aerosp Eng 22(1):15–23
40. Meijering E (2012) Cell segmentation: 50 years down the road [life sciences]. IEEE Signal Process Mag 29(5):140–145
41. Mollon G, Zhao J (2013) Generating realistic 3D sand particles using Fourier descriptors. Granul Matter 15(1):95–108
42. Osher S, Sethian J (1988) Fronts propagating with curvature-dependent speed: algorithms based on Hamilton–Jacobi formulations. J Comput Phys 79(1):12–49
43. Papoulis D, Tsolis-Katagas P, Katagas C (2004) Progressive stages in the formation of kaolin minerals of different morphologies in the weathering of plagioclase. Clays Clay Miner 52(3):275–286
44. Peters G, Abbey W, Bearman G, Mungas G, Smith J, Anderson R, Douglas S, Beegle L (2008) Mojave Mars simulant—characterization of a new geologic Mars analog. Icarus 197(2):470–479
45. Powers M (1953) A new roundness scale for sedimentary particles. J Sediment Res 23(2):117–119
46. Quinlan J (1986) Induction of decision trees. Mach Learn 1(1):81–106
47. Santamarina J, Cho G (2004) Soil behaviour: the role of particle shape. In: Advances in geotechnical engineering: the Skempton conference, vol 1. Thomas Telford, London, pp 604–617
48. Semnani SJ, Borja RI (2017) Quantifying the heterogeneity of shale through statistical combination of imaging across scales. Acta Geotech 12(6):1193–1205
49. Sezgin M, Sankur B (2004) Survey over image thresholding techniques and quantitative performance evaluation. J Electron Imaging 13(1):146–168
50. Sleutel S, Cnudde V, Masschaele B, Vlassenbroek J, Dierick M, Van Hoorebeke L, Jacobs P, De Neve S (2008) Comparison of different nano-and micro-focus X-ray computed tomography set-ups for the visualization of the soil microstructure and soil organic matter. Comput Geosci 34(8):931–938
51. Sommer C, Gerlich D (2013) Machine learning in cell biology-teaching computers to recognize phenotypes. J Cell Sci 126(24):5529–5539
52. Stark N, Hay A, Cheel R, Lake C (2014) The impact of particle shape on the angle of internal friction and the implications for sediment dynamics at a steep, mixed sand–gravel beach. Earth Surf Dyn 2(2):469–480
53. Sun W, Andrade JE, Rudnicki JW (2011a) Multiscale method for characterization of porous microstructures and their impact on macroscopic effective permeability. Int J Numer Methods Eng 88(12):1260–1279
54. Sun W, Andrade JE, Rudnicki JW, Eichhubl P (2011b) Connecting microstructural attributes and permeability from 3d tomographic images of in situ shear-enhanced compaction bands using multiscale computations. Geophys Res Lett. https://doi.org/10.1029/2011GL047683
55. Tagliaferri F, Waller J, Andò E, Hall S, Viggiani G, Bésuelle P, DeJong J (2011) Observing strain localisation processes in bio-cemented sand using X-ray imaging. Granul Matter 13(3):247–250
56. Tsomokos A, Georgiannou V (2010) Effect of grain shape and angularity on the undrained response of fine sands. Can Geotech J 47(5):539–551
57. Viggiani G, Andò E, Takano D, Santamarina J (2015) Laboratory X-ray tomography: a valuable experimental tool for revealing processes in soils. Geotech Test J 38(1):61–71
58. Vincent L, Soille P (1991) Watersheds in digital spaces: an efficient algorithm based on immersion simulations. IEEE Trans Pattern Anal Mach Intell 6:583–598
59. Vlahinić I, Andò E, Viggiani G, Andrade J (2014) Towards a more accurate characterization of granular media: extracting quantitative descriptors from tomographic images. Granul Matter 16(1):9–21
60. Wadell H (1933) Sphericity and roundness of rock particles. J Geol 41(3):310–331
61. Wang H, Zhang H, Ray N (2012) Clump splitting via bottleneck detection and shape classification. Pattern Recognit 45(7):2780–2787
62. Zhao B, Wang J (2016) 3D quantitative shape analysis on form, roundness, and compactness with μCT. Powder Technol 291:262–275
63. Zheng J, Hryciw R (2016) Segmentation of contacting soil particles in images by modified watershed analysis. Comput Geotech 73:142–152
64. Zheng J, Hryciw R (2017a) An image based clump library for DEM simulations. Granul Matter 2(19):26
65. Zheng J, Hryciw R (2017b) Soil particle size and shape distributions by stereophotography and image analysis. Geotech Test J 40(2):317–328
66. Zheng J, Hryciw R, Ventola A (2017) Compressibility of sands of various geologic origins at pre-crushing stress levels. Geotech Geol Eng 35(5):2037–2051
67. Zhou B, Wang J, Wang H (2018) Three-dimensional sphericity, roundness and fractal dimension of sand particles. Géotechnique 68(1):18–30