

¹ Reconstruction and Generation of 3D Realistic Soil
² Particles with Metaball Descriptor

³ Yifeng Zhao^{a,b,c}, Xiangbo Gao^{a,b,c}, Pei Zhang^{b,c,*}, Liang Lei^{b,c}, Stan Z. Li^{c,*},
⁴ S.A. Galindo-Torres^{b,c,*}

^a*Collage of Environmental and Resources Science, Zhejiang University, 866 Yuhangtang Road, Hangzhou, 310058, Zhejiang Province, China*

^b*Key Laboratory of Coastal Environment and Resources of Zhejiang Province, School of Engineering, Westlake University,, 18 Shilongshan Road, Hangzhou, 310024, Zhejiang Province, China*

^c*School of Engineering, Westlake University,, 18 Shilongshan Road, Hangzhou, 310024, Zhejiang Province, China*

⁵ **Abstract**

⁶ The morphology of soil particles is crucial in determining its granular characteristics and assembly responses. However, how to introduce accurate and various morphologies of realistic particles in modeling can be challenging, as it often requires time-consuming and costly X-ray Computed Tomography (XRCT). This has led to two prevalent problems in modeling: morphological reconstruction and generation. For reconstruction, we develop a geometric-based Metaball-Imaging algorithm. This algorithm is capable of accurately reconstructing the complex morphologies of realistic particles, including those with concave voids, which cannot be easily represented using other shape descriptors such as the spherical harmonic function. It employs a two-step approach, capturing the main contour of the particles using a series of non-overlapping spheres and then refining surface-texture details through gradient search. Four types of soil particles, hundreds of samples, are applied for evaluations. The result shows good matches on key morphological indi-

*Principal corresponding author

Email address: s.torres@westlake.edu.cn (S.A. Galindo-Torres)

20 cators(i.e., volume, surface area, sphericity, circularity, corey-shape factor,
21 nominal diameter and surface-equivalent-sphere diameter), confirming its re-
22 construction precision. For generation, we propose the Metaball Variational
23 Autoencoder. Assisted by deep neural networks, this method can generate
24 new 3D particles in Metaball form, while retaining coessential morphological
25 features with parental particles. Additionally, this method allows for control
26 over the generated shapes through an arithmetic pattern, enabling the gen-
27 eration of particles with specific shapes. Two sets of XRCT images different
28 in sample number and geometric features are chosen as parental data. On
29 each training set, one thousand particles are generated for validations. The
30 generation fidelity is demonstrated through comparisons of morphologies and
31 shape-feature distributions between generated and parental particles. Exam-
32 ples are also provided to demonstrate controllability on the generated shapes.
33 With the Metaball-based simulation framework previously proposed by the
34 authors, these methods can incorporate the real shape of particles into sim-
35 ulations, making it possible to simulate a large number of soil particles with
36 varying shapes and behaviors. Together, they have the potential to provide
37 valuable insights into the properties and behavior of actual soil particles.

38 *Keywords:* The Metaball function, Particle morphology, Particle
39 reconstruction, Optimization method, Particle generation, Variational
40 Autoencoder

41 **1. Introduction**

42 The soil particle (e.g. sands, cobblestones, rocks, etc.) has garnered sig-
43 nificant attention in recent decades within the fields of geology and geotech-

44 nics, due to its significance in a wide range of physical and industrial pro-
45 cesses(Zhang et al., 2020; Golombek et al., 2020; Tolomeo and McDowell,
46 2022; Zheng and Hryciw, 2015). It has been suggested that the macro-
47 behavior of granular media (e.g. soils, minerals etc.) is closely connected
48 to the micro-structural features of contained particles, where the morphol-
49 ogy is one of the most dominant factors(Gong et al., 2019; Santamarina and
50 Cho, 2004; Zhao et al., 2021; Lu et al., 2019; Chen et al., 2020). As illus-
51 trated by various observations and theories, the particle shape governs not
52 only granular characteristics(Shinohara et al., 2000; Xiao et al., 2019; Zhang
53 et al., 2022a) containing friction, interaction and deformation but also assem-
54 bly responses(Zuo et al., 2019; Yin et al., 2020; Zhou et al., 2020) enclosing
55 permeability, strength and failure.

56 On elucidating the impact of particle shape on the granular soil, simula-
57 tion schemes based on micro-mechanical models, especially the discrete ele-
58 ment method (DEM)(Cundall and Strack, 1979), have prevailed in the past
59 few decades. In original DEM, the particle is simplified as circles or spheres,
60 which is hard to reflect the impact of shape, e.g. the resistance to rolling.
61 Under this context, the key issue is how to fully reconstruct the morphology
62 of granular matters in DEM simulations. The initial attempt starts from
63 the rolling resistance model(Iwashita and Oda, 1998; Wensrich and Katter-
64 feld, 2012), which improves the contact model with friction mechanism from
65 the surface roughness to mimic the particle shape. In spite of progresses
66 in dealing with physical phenomenon such as inertial friction(Wensrich and
67 Katterfeld, 2012) and shear banding(Coetzee, 2016), it still suffers from the
68 introduction of a number of free parameters and the property of unrealistic

⁶⁹ fabric(Jiang et al., 2005). The clustering technique is another interesting ap-
⁷⁰ proach, which represents shape characteristics by clustering simple, regular
⁷¹ graphics like spheres(Zheng and Hryciw, 2016), Polyhedrons(Höhner et al.,
⁷² 2012), Ellipsoids(Regueiro et al., 2014) and Sphere-polyhedrons(Galindo-
⁷³ Torres, 2013). This technique is widely used due to its simplicity yet with
⁷⁴ the drawback of impractical surface roughness(Grabowski et al., 2021) and
⁷⁵ computational expensiveness(Garcia et al., 2009).

⁷⁶ With the inspiring development of X-ray Computed Tomography (XRCT)
⁷⁷ and computer vision techniques, opportunities are provided to bring more
⁷⁸ accurate and sophisticated shape features into DEM simulations. Various
⁷⁹ reconstruction methods, which are called shape descriptors in this paper, are
⁸⁰ developed to compress realistic particle morphologies into a uniform math-
⁸¹ ematical representation for simulations. A good example is the Fourier de-
⁸² scriptor, which is developed to capture particle shapes based on the aver-
⁸³ age normalized Fourier spectrum of main contours from the targeted parti-
⁸⁴ cle(Thomas et al., 1995; Mollon and Zhao, 2013). Through similar pattern,
⁸⁵ the Spherical-Harmonic (SH) descriptor can also be used to capture various
⁸⁶ shapes features(Zhou et al., 2018, 2015; Su and Yan, 2018; Wang et al., 2022).
⁸⁷ However, the above two methods are limited to tackling star-like particles, of
⁸⁸ which all line segments between particle-center and particle-surface points are
⁸⁹ located within the particle body(Su and Yan, 2018). Many particles, such as
⁹⁰ lunar soils and concave sand, do not follow such constraints. Another inter-
⁹¹ esting attempt is that Vlahinić et al(Vlahinić et al., 2017) apply the level-set
⁹² function as the descriptor. However, the shape descriptor has to cooperate
⁹³ with mechanical models in practical simulations. It is necessary to consider

94 its coupling scheme when using a descriptor. Although level-set function has
95 an excellent reconstruction fidelity, it still has a reliance on computational
96 resources due to the look-up table mechanism in coupling mechanical models
97 (mainly in contact detection of DEM)(Kawamoto et al., 2016; Medina and
98 Jerves, 2019; Zhao and Zhao, 2019). To relieve the efficiency problem in
99 coupling, Zhao and Zhao(Zhao and Zhao, 2019; Zhao et al., 2020) developed
100 a poly-superellipsoid based descriptor yet at the cost of certain constraints
101 on the expressed shape(such as smoothness and complexity). Recently, the
102 Metaball descriptor is introduced by the authors(Zhang et al., 2021) to recon-
103 struct non-spherical particle shapes. With proper function form, the contact
104 detection of it can be tackled at a low cost, which enables a more efficient
105 simulation framework(Zhang et al., 2022b). Such framework is further cou-
106 pled with Lattice Boltzmann Method for simulations of more complicated
107 physical processes in fluid-particle systems(Zhao et al., 2023). However, it
108 remains difficult for Metaballs to reconstruct complex-shaped granular par-
109 ticles. Therefore, there is still a pressing need for a shape reconstruction
110 framework with descriptor of better balance between the imaging quality of
111 particle shape and computational efficiency in coupling mechanical models,
112 especially for complex-shaped particles of angular or concave features.

113 Apart from reconstruction, particle generation is another non-negligible
114 challenge. Although XRCT can be used to scan all involved particles in
115 application, particle scanning and image processing can be economy-costly
116 and time-consuming. In practical engineering applications, it is typical that
117 only a small fraction of particles can be analyzed due to limited resources, as
118 reported in (Nie et al., 2020; Shi et al., 2021; Zhou and Wang, 2017). Direct

simulation with them will suffer from repetitive particle morphologies. This makes it necessary to generate realistic particles with coessential morphological features. As the development of the aforementioned shape descriptors, many attempts have been inspired to tackle generation tasks. Among them, the SH-based technique is a popular choice(Liu et al., 2011; Sun and Zheng, 2021). It first incorporates geometric features into specific SH coefficients. Then, different algorithms like random field (Grigoriu et al., 2006), fractal dimension(Wei et al., 2018), Nataf transformation(Nie et al., 2020), and principal component analysis(Xiong and Wang, 2021; Zhou and Wang, 2017) are applied on the distilled SH coefficient to add small variances following the morphological pattern of parental particles for generation. The Fourier-descriptor based approach is similar implementation(Mollon and Zhao, 2013, 2012, 2014; Chen et al., 2022). Another interesting attempts are geometry-based algorithms(Jerves et al., 2017; Medina and Jerves, 2019; Buarque de Macedo et al., 2018; Macedo et al., 2023). It requires some shape feature distributions as "morphological DNA". Then, geometric stochastic cloning algorithms(Jerves et al., 2017; Medina and Jerves, 2019) or the genetic algorithm(Buarque de Macedo et al., 2018; Macedo et al., 2023) are implemented to generate particles based on those distilled DNA. However, the above schemes suffer from some problems, including: (1) underfitting and overfitting problems on shape-feature distributions of the generated particles, e.g. the distributions of surface area and volume(Shi et al., 2021; Xiong and Wang, 2021). Underfitting refers to the situation in which the model fails to learn from the data. In the context of particle generation, it usually happens in methods, which assume the shape-feature distributions of

144 particle to be Gaussian(Grigoriu et al., 2006). This violates the fact that
145 granular particles are complex in morphology and its shape features may not
146 simply follow a Gaussian distribution(Wei et al., 2018). Overfitting, on the
147 other hand, refers to the situation where the model is too focused on specific
148 features of the training data, which can lead to poor performance on un-
149 seen data(Shi et al., 2021). In the context of particle generation, overfitting
150 often occurs when a model only learn from certain geometric features and
151 uses learned distributions to generate particles(Jerves et al., 2017; Medina
152 and Jerves, 2019). However, shape features of granular matter can have dis-
153 tinct distributions. Generation by force sampling from limited distributions
154 can result in particles with incorrect distributions of certain shape-features.
155 (2) Hard to obtain particles with specific morphological features(Buarque de
156 Macedo et al., 2018; Liu et al., 2011), e.g. generating non star-like particles
157 with angled features. (3) Involving complex mixture models, which treat
158 particle generation as a high-dimentional, multi-parameter estimation prob-
159 lem (Medina and Jerves, 2019; Jerves et al., 2017). Such method can achieve
160 advance in performance, but still has an reliance on the computational and
161 human resources, which can not be obtained easily by all individuals or in-
162 stitutions(Shi et al., 2021). (4) Requiring bridging or transformation into
163 other descriptors before practical simulations, which often results in a trade-
164 off between accuracy and efficiency(Wang et al., 2018; Lai et al., 2022). A
165 framework with a more flexible shape descriptor, which can tackle the above
166 problems, is therefore needed.

167 Deep learning techniques, which are able to extract and conclude gen-
168 eral features from high dimensional data, provide an opportunity to tackle

169 generation problems(Bourilkov, 2019). Among them, the variational autoen-
170 coder(VAE) is a popular choice(Kingma and Welling, 2013). It achieves gen-
171 erative modeling through establishing a mapping between some latent vari-
172 ables and the generative targets(Higgins et al., 2016). With neural-network
173 based variational inference, it can interpret and learn the underlying causal
174 relations within the data. And practical generation can then be carried out
175 by controlling those latent variables. Note that those latent variables are reg-
176 ularized, which means variables mapped with similar generated results are
177 located together. This can brings many advantages, including feature ex-
178 traction, generation quality and generation control(Tan et al., 2018). Many
179 successful applications have verified its superiority in generation tasks over
180 traditional methods(Shi et al., 2021; Tan et al., 2018).

181 On the above dilemmas in reconstruction and generation of soil particles,
182 this paper presents Metaball-based solutions. For reconstruction, we propose
183 a geometric-based Metaball-Imaging (MI) algorithm, which can capture com-
184 plex particle morphologies from XRCT images with the Metaball descriptor.
185 The utilization of geometric constraints allows Metaballs to capture more
186 shape-details, even angular edges and concave voids which are hard to express
187 by most of existing shape descriptors. Four types of realistic soil particles in
188 distinct morphologies are selected for validation. The impact of the control
189 point number, is carefully investigated. On this basis, the abilities of MI to
190 reconstruct single particle and particle groups are examined at length on ma-
191 jor morphological features, including the volume, area, sphericity, circularity,
192 Corey shape factor, nominal diameter and volume-equivalent diameter. For
193 generation, we develop the Metaball-based Variational Autoencoder (Meta-

ballVAE). It can learn from the XRCT image of targeted grains and generate random Metaball-based particles retaining major morphological features from a regularized latent space, where complex calculations are converted into one-step solutions. Note that MetaballVAE can learn not only the geometric features of the single particle but also feature distributions of the particle group. The regularized latent space also makes it possible to modify particle morphologies in an arithmetic pattern, allowing for obtaining particles with specific shapes. Examined with two groups of particles different in sample number and geometric characteristics, the proposed generation algorithm has proved to be robust and effective.

2. Characterizing 3D Realistic Granular Particles with Metaball Descriptor

2.1. The Metaball descriptor

The Metaball function is a shape descriptor firstly proposed by Blinn(Blinn, 1982). It is expressed as a function of a distance between the probing point and a number of control points. Herein, we choose the typical inverse-square form:

$$f(\mathbf{x}) = \sum_{i=1}^n \frac{\hat{k}_i}{(\mathbf{x} - \hat{\mathbf{x}}_i)^2} \quad (1)$$

where n is the number of control points. $\hat{\mathbf{x}}_i$ is the position of control point i , representing a skeleton for the parameterized shape. \hat{k}_i is the positive coefficient determining the weight of i th control point. Together, a pair of \hat{k}_i and $\hat{\mathbf{x}}_i$ constitutes a control circle in 2D (black circles in Figure 1, a) or sphere in 3D. By setting this function equal to value 1, these control circles or spheres form the contour of the targeted particle surface, of which the

217 parameter is noted as $\mathbf{M} = \{\hat{k}_i, \hat{\mathbf{x}}_i\}$. A 2D Metaball example is shown in
218 Figure 1 (a), where the particle surface contour is represented by red line.

219 Note that the value range of the Metaball function with random input is
220 from 0 to ∞ . For 2D Metaball as shown in Figure 1 (b), the function value
221 will be exactly 1.0 at the surface (indicated by red), with lower values for
222 points outside the particle (depicted by green) and higher values for points
223 inside (represented by orange). This property offers a straightforward way
224 to determine whether a point is inside or outside the particle, providing
225 possibilities for a more efficient contact-detection scheme(Zhang et al., 2021;
226 Zhao et al., 2023). It also decides the loss function for the gradient search
227 (GS) in Metaball-Imaging as discussed in Section 2.2.3.

228 Moreover, such function form endows the model a strong physical mean-
229 ing. This benefits the treatment of translation, rotation and scaling in cou-
230 pled mechanical models. It also creates convenience for the shape expres-
231 sion. As shown in Figure 2 (a), only a small number of control points are
232 required to represent a complex smooth surface. This reveals that the Meta-
233 ball model can relieve the trade-off between reconstruction fidelity and model
234 complexity, especially avoiding artificial surface roughness of the clustering
235 scheme(Wang et al., 2022). Besides, Figure 2 (b) shows that both protrusion
236 and indentation can be expressed with different settings of \hat{k}_i . This indi-
237 cates that the detailed particle surface texture can be described by adding
238 new control points. Thus, it is possible for the Metaball descriptor to rep-
239 resent highly complex-shaped particles at low cost, even non star-like and
240 concave-void ones as shown in Figure 4. Related studies also indicate that
241 the Metaball descriptor can be used to represent the deformation and failure

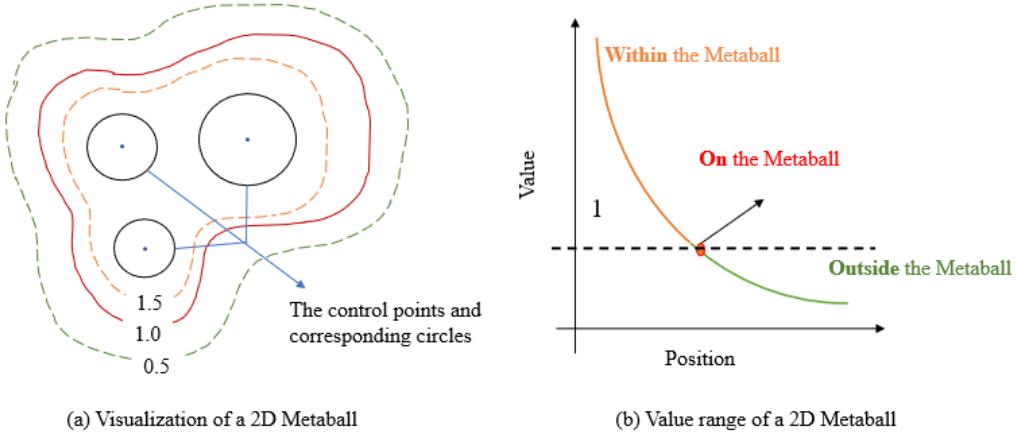


Figure 1: The visualization and value range of a 2D Metaball in the form of Eq. 1. (a) stands for the contour line of metaball, where the values on lines are calculated by Equation 1 and the curve with a value 1.0 is the surface of corresponding metaball. In addition, those black points stand for the control points and the black circles are the corresponding control circles. (b) displays the visualization of the value range of the Metaball.

242 processes effectively(Plankers and Fua, 2001; Jin et al., 2000). The above
 243 advantages together make the Metaball descriptor a proper choice in repre-
 244 senting particle morphologies.

245 *2.2. Metaball Imaging*

246 Metaball-Imaging (MI) is developed to transform the XRCT image of
 247 irregular-shaped particle into an explicit, Metaball-function based mathe-
 248 matical representation, which is called MI *avatar* in this paper. It possesses
 249 key necessary morphological information of the given particle for further me-
 250 chanical simulations. MI is first proposed by authors in previous works(Zhao
 251 et al., 2023). Here, its fitting ability is greatly enhanced by considering ge-

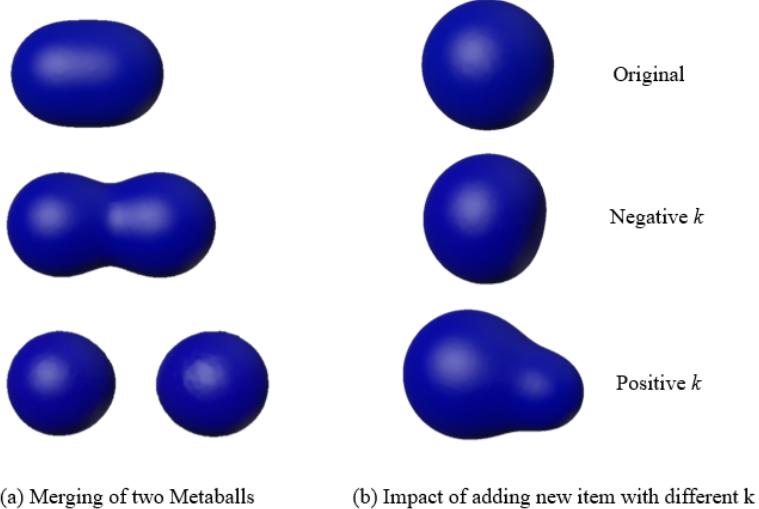


Figure 2: Visualization of 3D Metaballs. (a) stands for two Metaballs getting closer to each other, (b) reveals the influence of k_i value, where adding a positive Metaball can result in protrusion and a negative Metaball for indentation.

ometric information. The upgraded MI consists of three major parts, data pre-processing, capturing principle outer-contour with sphere-clustering (SC) and refining the distilled contour with gradient-search (GS). The general framework is as shown in Figure 3.

2.2.1. Data pre-processing

Two strategies are adopted in pre-processing, transformation of coordinates (TOC) and specification of interested region (SOIR). These techniques involve three distinct objects, namely the XRCT voxels, the centralized voxels \mathbf{V} and the point hull \mathbf{H} . As illustrated in Figure 3, these objects are depicted in black, green and orange, respectively. TOC is implemented to translate the XRCT voxels into the coordinate system centered at the origin.

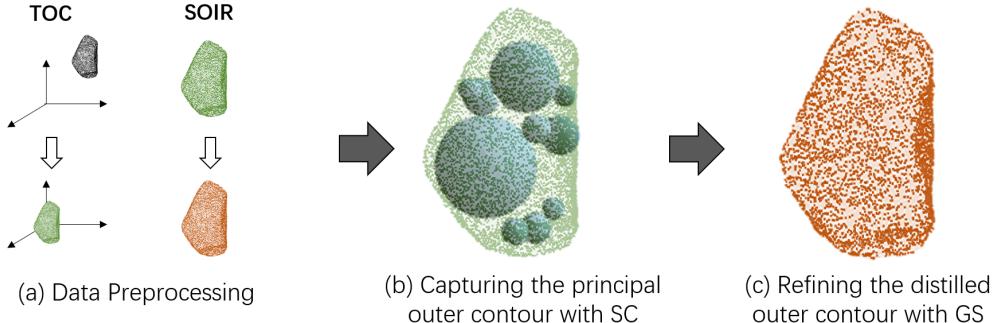


Figure 3: The framework of Metaball Imaging.

263 Such an operation is dedicated to obtain centralized voxelated representation
 264 \mathbf{V} , which can avoid abnormal fitted parameters caused by XRCT coordi-
 265 nates. And SOIR is aimed to distill the point hull \mathbf{H} from the cloud points
 266 of targeted particles.

267 *2.2.2. Capturing principle outer-contour with sphere-clustering*

268 Here we develop an algorithm to capture the principle outer-contour by
 269 searching for a series of non-overlapping inscribed spheres as the control
 270 spheres (Figure 3, b). The flowchart of this algorithm is as shown in Algorithm
 271 1. Initially, the Euclidean distance transform (Bailey, 2004) is carried out
 272 on the pre-processed voxelated representation \mathbf{V} . Then, the radius r_i and
 273 centre \mathbf{c}_i of the maximum inscribed sphere can be found with the largest
 274 value and its corresponding position in the transformed vector. Next, the
 275 voxels belonging to the inscribed sphere are zeroed. The above operations
 276 are repeated until satisfying the termination condition. Here, it is defined
 277 as the number of inscribed spheres, which is equal to the number of control

278 points n . In the end, a set of inscribed spheres S_I can be obtained.

Algorithm 1 The Sphere-Clustering Algorithm for capture of the principal outer contour

Input: The voxelized particle \mathbf{V} , the number of control points n

Output: The distilled set of inscribed spheres \mathbf{S}_I

- 1: **for** $i = 1, 2, \dots, n$ **do**
 - 2: **Transform** - Implementing Euclidean distance transform on \mathbf{V} ;
 - 3: **Search** - Finding the radius r_i and centre c_i of maximum inscribed sphere with the maximum value in the transformed vector;
 - 4: **Reset** - Zeroing those voxels of the searched inscribed circles and updating \mathbf{V} ;
 - 5: **end for**
 - 6: **Return:** The distilled set of inscribed spheres $\mathbf{S}_I = \{r_i, c_i\}, i \in [1, n]$.
-

279 2.2.3. Refining the distill contour with gradient-search

280 In this section, the gradient-search (GS) is applied to refine the distilled
281 outer-contour. Since this technique is already presented in author's previous
282 work(Zhao et al., 2023), only key points are discussed here, as listed in Al-
283 gorithm 2.. With the distilled set of inscribed spheres \mathbf{S}_I , a Metaball model
284 $\mathbf{M}_r = \{r_i, c_i\}$ can be used to represent the principle outer-contour of the
285 target particles. Then, a loss function is defined to calculate the gradient in-
286 formation provided by the distilled point hull \mathbf{H} . Here, we utilize a piecewise

287 function instead of the popular Mean Square Error (MSE):

$$L(\mathbf{M}_r) = \begin{cases} \sum_{i=1}^m (f_{\mathbf{H}_i}^l(\mathbf{M}_r) - 1)^2, & f_{\mathbf{H}_i}^l \in [2, +\infty) \\ \sum_{i=1}^m (f_{\mathbf{H}_i}^l(\mathbf{M}_r) - 1), & f_{\mathbf{H}_i}^l \in [1, 2] \\ \sum_{i=1}^m \left[(f_{\mathbf{H}_i}^l(\mathbf{M}_r) - 1)^2 + \frac{1}{f_{\mathbf{H}_i}^l(\mathbf{M}_r)} - 1 \right], & f_{\mathbf{H}_i}^l \in [0, 1] \end{cases} \quad (2)$$

288 where m is the number of control points and $f_{\mathbf{H}_i}^l(\mathbf{M}_r) = \sum_{j=1}^n \frac{r_j}{(\mathbf{H}_j - \mathbf{c}_j)^2}$.

289 The reasons for implementing the above function are two-fold. On the
290 one hand, it can avoid obtaining distorted Metaball models. As mentioned
291 in Section 2.1, when the study point is internally close to or externally far
292 from the Metaball surface, its Metaball function value will all be close to 1.
293 This can make the corresponding MSE very small, resulting in local opti-
294 mal GS solutions, distorted models with control points outside the targeted
295 surface(Zhao et al., 2023). The proposed function effectively remedies this
296 problem, by amplifying the loss value of these points. On the other hand, it
297 can also improve the adaptability to complex geometry, since the amplifica-
298 tion of loss values allows GS to capture more gradient information, especially
299 on the surface texture.

300 In the end, the model parameters are optimized through gradient descent:

$$\mathbf{M}_r \leftarrow \mathbf{M}_r - \eta \cdot \nabla_{\mathbf{M}_r} L(\mathbf{M}_r) \quad (3)$$

301 where η is the learning rate; $\nabla_{\mathbf{M}_r} L(\mathbf{M}_r)$ is the gradient of the $L(\mathbf{M}_r)$ to
302 the parameter \mathbf{M}_r . As for gradient update, Adaptive Moment Estimation
303 (Adam) and Stochastic Gradient Descent (SGD) (Zhang, 2018) are com-
304 bined to achieve a balance in computational efficiency and convergence abil-
305 ity(Wang et al., 2018). The above function is repeated until satisfying the
306 termination condition, which is the number of generations E^{gs} .

Algorithm 2 The Gradient Search for refinement of outer contour

Input: the particle point hull \mathbf{H} , the number of generations E^{gs} , the learning rate η , the distilled set of inscribed spheres \mathbf{S}_I .

Output: the metaball model of the refined outer contour \mathbf{M}_f .

- 1: \mathbf{S}_I is taken as the Metaball model of principle outer-contour, the initial value \mathbf{M}_r ;
 - 2: **for** $i = 1, 2, \dots$, to E^{gs} **do**
 - 3: $\mathbf{M}_r \leftarrow \mathbf{M}_r - \eta \cdot \nabla_{\mathbf{M}_r} L(\mathbf{M}_r)$;
 - 4: **end for**
 - 5: **Return:** The searched parameter \mathbf{M}_f .
-

307 **3. Evaluation of Metaball-Imaging**

308 In this section, we discuss the reconstruction performance of the proposed
309 MI by examining four types of grains with different geometric characteristics.

310 *3.1. Dataset and XRCT setting*

311 Four types of particles, which are popular in the research field, are chosen
312 for evaluation. They are 20 cobblestones, 290 Ottawa sands, 25 angular sands
313 and 5 concave sands, as shown in Figure 4 (a). The cobblestone is a common
314 granite or basalt in the riverbed. It possesses obvious round features, which
315 are the results of water flow. The Ottawa sand is a typical grain of pure
316 quartz exploited from Ottawa in Canada. Due to geological transportation,
317 it features smooth and angled characteristics. As for the angular sand and the
318 concave sand, they are sand for common building materials mainly composed
319 of quartz and feldspar. They have angular and concave characteristics caused
320 by mechanical disruption and chemical weathering.

321 In XRCT imaging, the ZEISS Xradia 610 Versa is utilized. The voltage
 322 of X-ray source is set to 140kV for cobblestone and 80kV for the others. The
 323 0.4X detector is chosen in the scan recipes, which means the corresponding
 324 optical magnification is 0.4. The voxel size is $45.48\mu m$ for the cobblestone,
 325 $18.56\mu m$ for the Ottawa sand, $51.64\mu m$ for the angular sand, $19.59\mu m$ for the
 326 concave sand. On average, the particle in XRCT images contains more than
 327 7.9×10^6 voxels to represent a real grain geometry. The particle segmentation
 328 is done by "ilastik", a machine-learning driven edge-detection algorithm for
 329 XRCT images (Sommer et al., 2011). The "ilastik" algorithm can detect
 330 particle edges by considering both voxel intensity and brightness gradient,
 331 which effectively avoids omitting particles with concave spots on the surface
 332 or broken spots on the edge. The processed particles are shown in Figure4
 333 (b).

334 *3.2. Metrics*

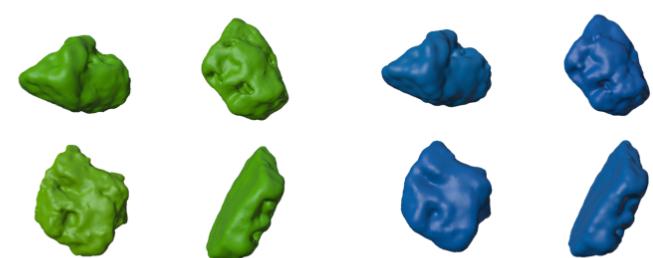
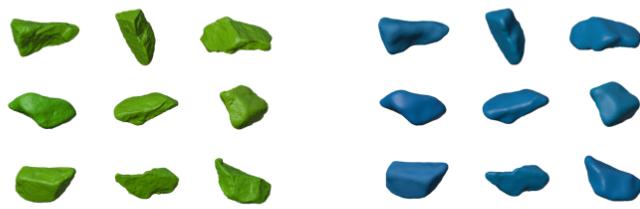
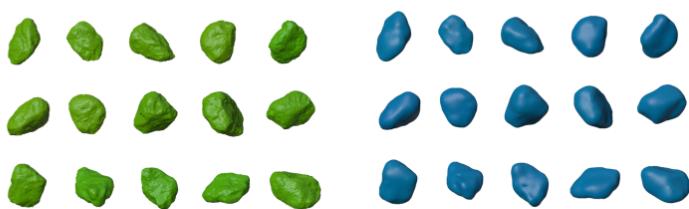
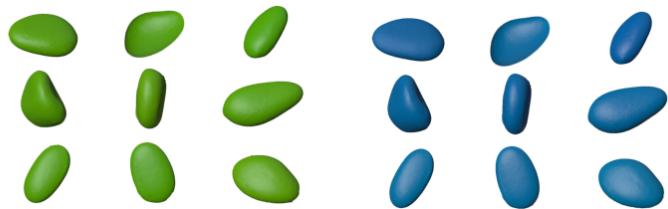
335 We select seven shape factors for evaluation: surface area A , volume V ,
 336 Corey Shape Factor CSF , nominal diameter D_n , surface-equivalent-sphere
 337 diameter D_s , sphericity ϕ and circularity C .

338 The Corey Shape Factor CSF (Dietrich, 1982) reveals the dimension fea-
 339 ture of the studied particle, as given by:

$$CSF = \frac{L_s}{\sqrt{L_i L_l}} \quad (4)$$

340 where L_s , L_i and L_l are the shortest, intermediate and longest axis lengths
 341 of particles.

342 The nominal diameter D_n and surface-equivalent-sphere diameter D_s are
 343 two widely used parameters(Bouwman et al., 2004; Zhang et al., 2016). The



(a) Microscopic Image

(b) XRCT Images

(c) MI Avatar

Figure 4: Representative particles, XRCT images and corresponding MI avatars of the selected four types of particles. Numbers of control points to reconstruct the cobblestone, Ottawa sand, angular sand and concave sand are 40, 40, 100 and 120 respectively.

344 D_n is defined as the diameter of the volume-equivalent sphere. And the D_s
345 takes the following form:

$$D_s = \sqrt{\frac{4A_p}{\pi}} \quad (5)$$

346 where A_p = the maximum projected area of the particle. Here, they are
347 combined as $D_{ns} = D_n/D_s$ to form a dimensionless quantity.

348 The sphericity ϕ (Mora and Kwan, 2000) is the measure of similarity
349 between the studied particle and the sphere, which is defined as:

$$\phi = \frac{A_{ve}}{A} \quad (6)$$

350 where A_{ve} = the surface area of the volume-equivalent sphere to the studied
351 particle; A = the surface area of the studied particle.

352 Another frequently used metric is the circularity C (Bouwman et al.,
353 2004), which evaluates the roundness of non-spherical particle:

$$C = \frac{\pi D_s}{P_p} \quad (7)$$

354 where P_p is the the perimeter of the particle's projected-area.

355 3.3. Impact of the control-point number on reconstruction quality

356 In Metaball-Imaging, a vital variable is the number of control points
357 n , which is closely related to fidelity. For applications, it is necessary to
358 understand the relationship between control point number and reconstruction
359 quality. Evaluations are carried out on multiple, randomly-selected particles
360 of each type.

361 The converged value of the loss function (Eq. 2) in GS is a critical index
362 for reconstruction. As shown in Figure 5, the log loss values of studies par-
363 ticles decrease rapidly as the increase of n . When n equal to 40, all of them

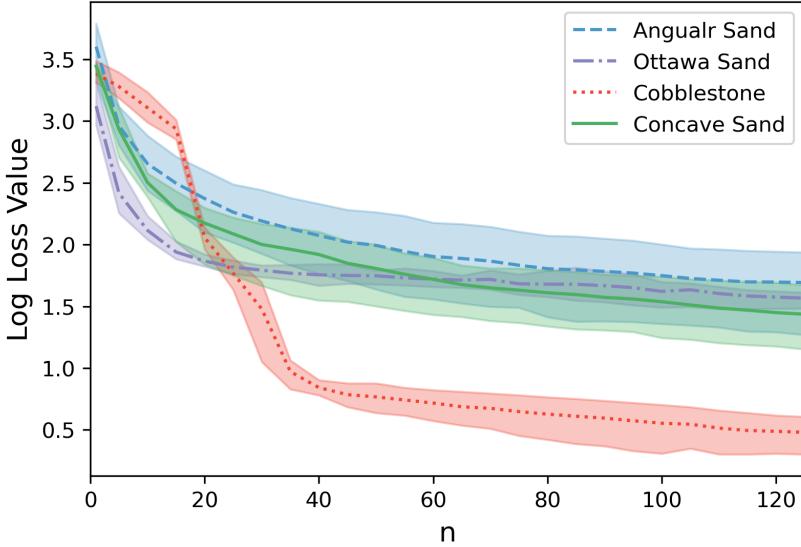


Figure 5: Correlations of the log loss value with the number of control points for four types of particles, where the log loss value is base-10 logarithm of the final value of Eq. 2 in gradient search.

364 reached valleys, where further increasing the control points will not decrease
 365 the loss function significantly. Compared with other particles, the cobble-
 366 stone has a very different loss curve. Its loss value first decreases slowly,
 367 where limited control points are hard to represent cobblestones with flat fea-
 368 tures. Then, the loss value starts to drop rapidly as more control points
 369 are involved until converges at n equal 40, following a similar pattern of the
 370 other particles. Note that it takes much less control points (about 20) for
 371 the cobblestone to reach the converged loss value of the rest particles and
 372 its converged value is also lower. These together indicate the superiority of
 373 Metaball function in representing particles with smooth, round features.

374 Figure 6 illustrates the morphology of typical particles reconstructed by

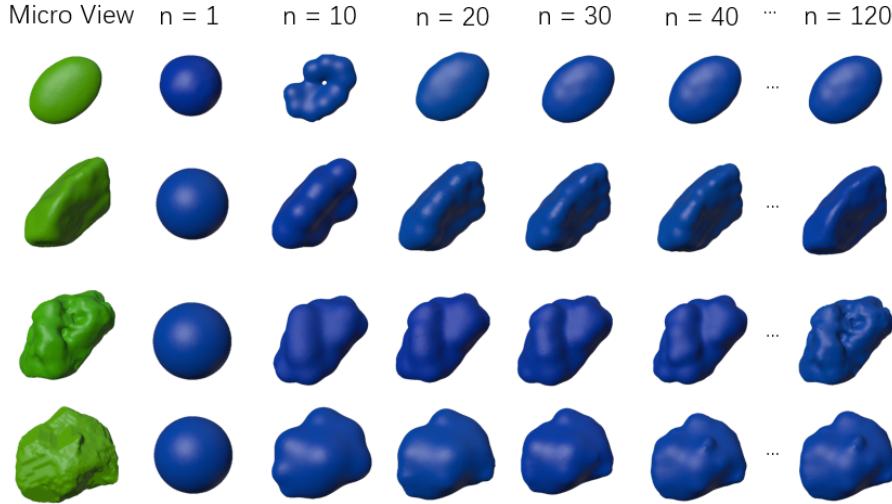


Figure 6: Impact of the control point number on the reconstructed particle morphologies

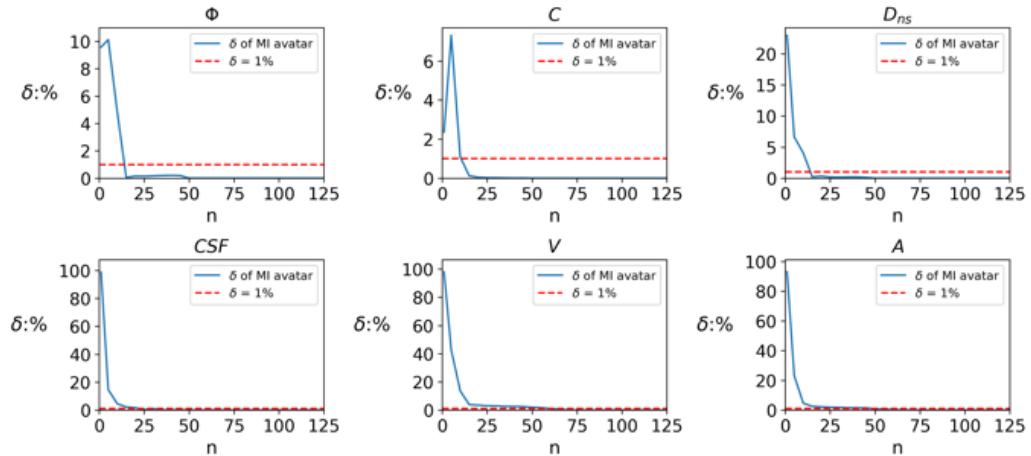
375 MI with different settings of n . It is shown that the increasing control points
 376 allow MI avatar to capture more morphological details. The reconstruction
 377 fidelity first improves with the increase of control points and then become
 378 steady, which indicates a proper n can avoid too many avatar parameters
 379 while guaranteeing reconstructed fidelity. It is worth noting that the number
 380 of control points required for different types of particles is different. For
 381 the cobblestone and Ottawa sand, dozens of control points are enough for
 382 reconstruction. In contrast, more points are needed for the capture of angular
 383 and concave sand, which possess more complex geometrical features.

384 To select n properly, we carry on further evaluations on more precise
 385 geometric metrics of typical particles (See Figure 7). The relative error δ
 386 between the real particle and MI avatar is selected for comparison. It is
 387 found that all δ of chosen metrics converged rapidly to low stable values
 388 after n greater than 40, which indicates a good match between parents and

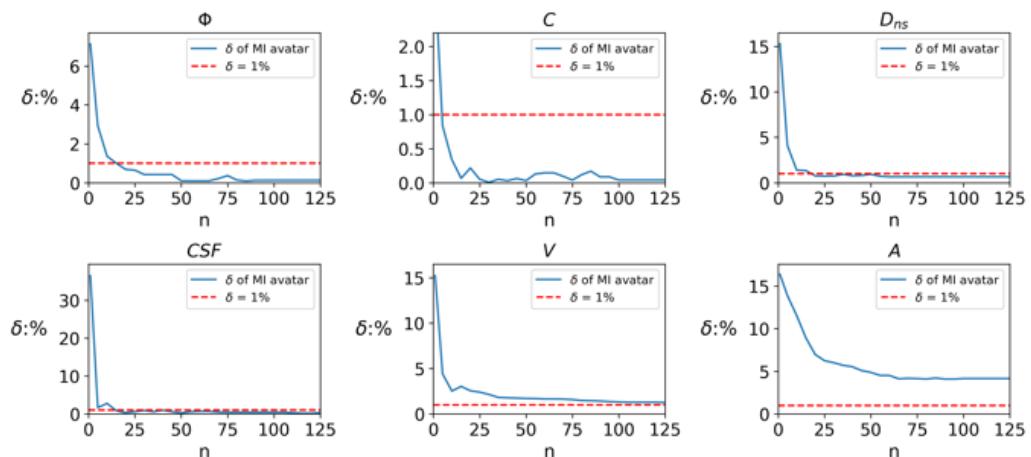
389 MI avatars. However, the reconstruction quality of each indicator varies
390 considerably. The sphericity and circularity are much easier to reconstruct
391 compared to the rest. And small deviations can be observed in some metrics,
392 like volume and surface area of angular and concave sands.

393 Such phenomena are mainly attributed to those minor details of the
394 XRCT images, such as the surface roughness and concave holes. These
395 details will interfere the imaging quality of real particles. Theoretically, a
396 better agreement can be achieved with MI avatars of more control points. In
397 fact, those error curves are not converged yet and are still slowly decreasing.
398 However, a refined MI avatar may not be idealized in application due to pos-
399 sible errors in image measured value of the real particle. Taking an example
400 on the image measured result of surface area, the source of error can either
401 be from the selection of XRCT optical magnification or treatment of voxel
402 connectivity, which might make the scanned "real" particle possess limited
403 fidelity and show unrealistic wrinkle characteristics. On the contrary, the
404 MI avatar is distilled directly from the cloud points, of which the results are
405 more convincing. A similar problem is also reported in (Zhou et al., 2015),
406 where a 10% relative error is observed on the image-measured surface area
407 compared to the reconstructed SH value.

408 In conclusion, the MI avatar with n of 40 is found to be sufficient to recon-
409 struct the major morphological features and surface texture for all four types
410 of particles involved in the current study. For higher fidelity reconstruction of
411 angular and concave features, a MI avatar with more control points, offering
412 more geometric details, can be a better choice.

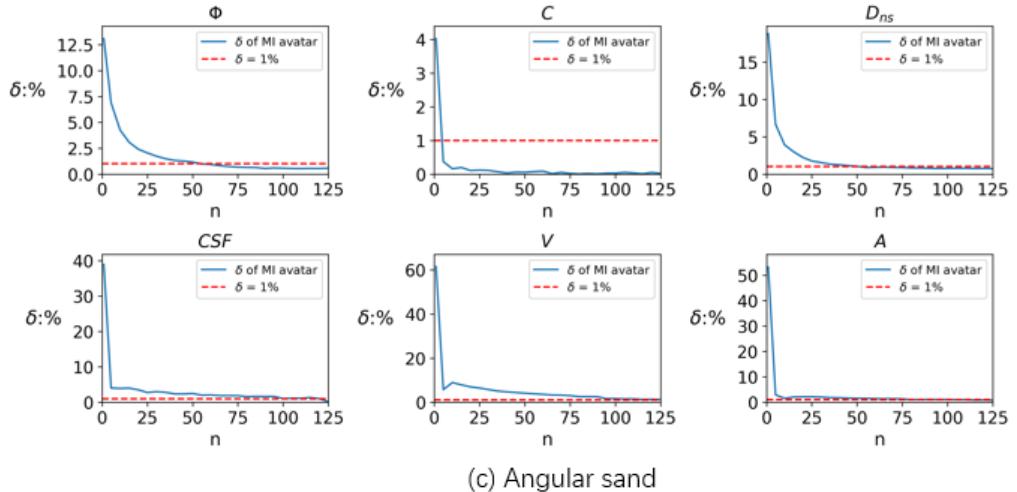


(a) Cobblestone

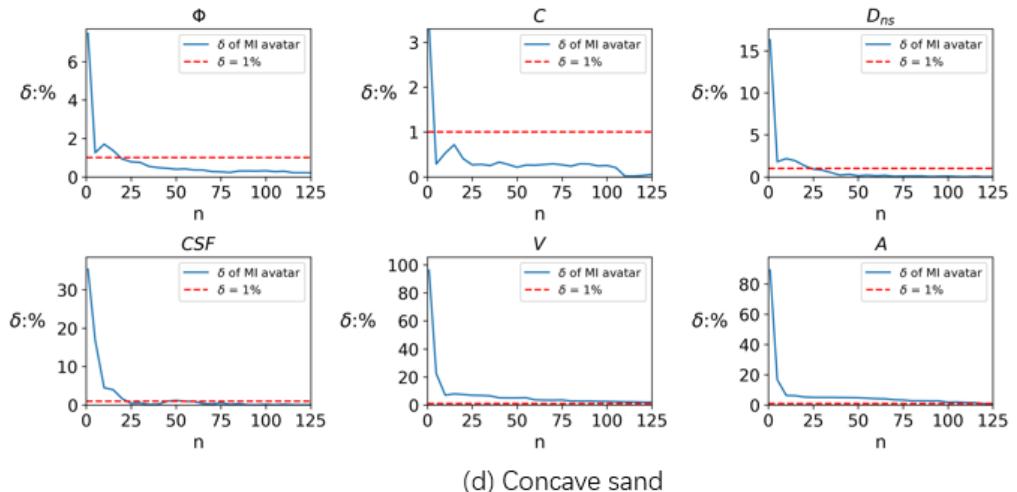


(b) Ottawa sand

Figure 7 I: Impact of the control point number on reconstructed geometric metrics



(c) Angular sand



(d) Concave sand

Figure 7 II: Impact of the control point number on reconstructed geometric metrics

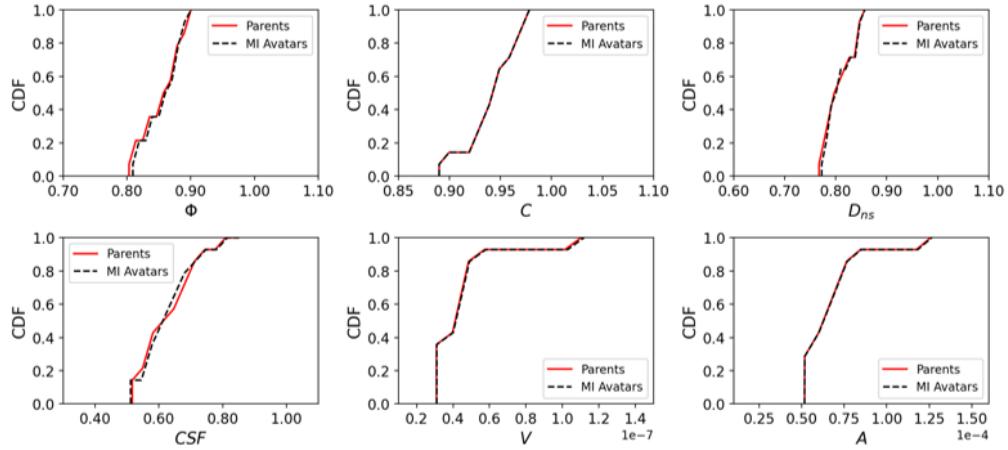
Processor	Intel Xeon W-2133 3.60GHz
RAM	32GB 2400MHz

413 3.4. Reconstruction performance on particle groups

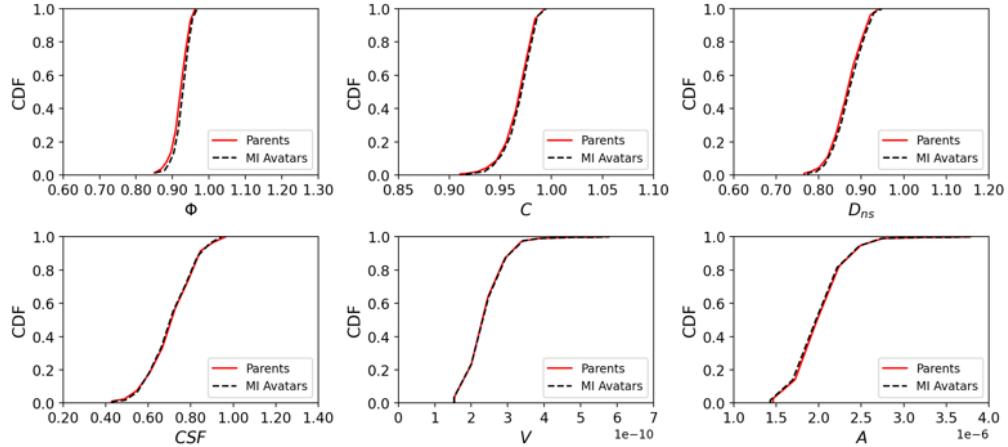
*414 To further evaluate the performance of MI, experiments are implemented
 415 on particle groups introduced in Section 3.1. MI avatars of 40 control points
 416 are adopted to reconstruct the cobblestone and Ottawa sand. As for angular
 417 sand and concave sand, 100 and 120 control points are applied separately.
 418 The learning rate of gradient search is chosen to be 0.001.*

*419 Visualizations of randomly-selected parental particles and MI avatars of
 420 each type can be found in (b) and (c) of Figure 4. Such a good match can also
 421 be observed in cumulative distribution functions (CDFs) of shape indicators
 422 as shown in Figure 8. For example, Figure 8 (b) compares the CDF of the
 423 six shape factors between the parents and MI avatars of the Ottawa sand.
 424 A good agreement is found. Noting that small deviations can be found on
 425 some metrics of the rest particle groups. This is mainly caused by the limited
 426 particle number rather than the reconstruction error.*

*427 The above results are calculated with Python in the parallel mode on
 428 a personal workstation (Technical details listed in Table 3.4). A total of
 429 340 avatars are characterized from their XRCT images and the average time
 430 taken to reconstruct each particle is 1.03 minutes.*

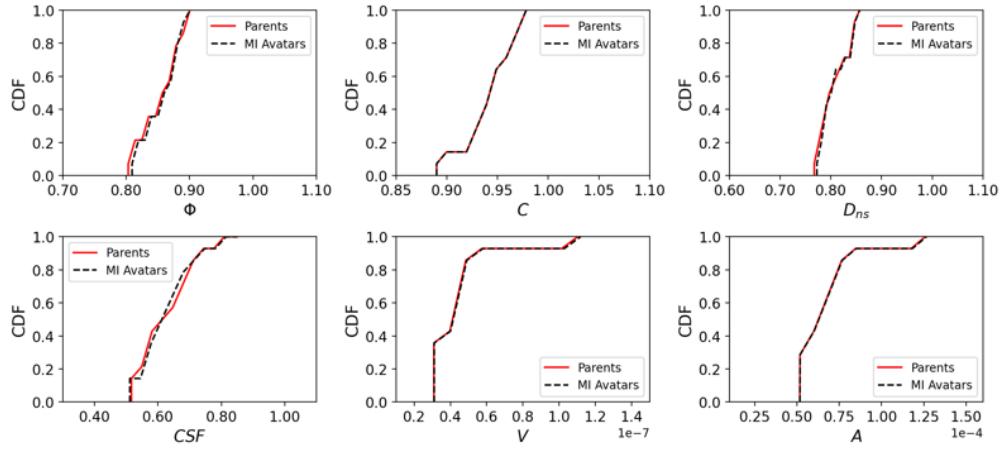


(a) Cobblestone

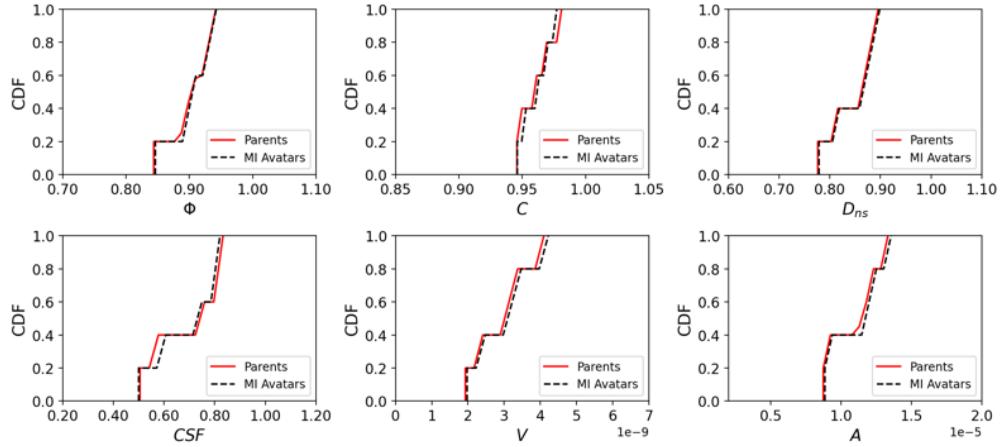


(b) Ottawa sand

Figure 8 I: Reconstruction quality of morphological metrics on particle groups



(c) Angular sand



(d) Concave sand

Figure 8 II: Reconstruction quality of morphological metrics on particle groups

431 4. Generating 3D Style-similar Granular Particles with Metaball
432 Descriptor

433 4.1. MetaballVAE

434 The variational autoencoder(VAE)(Kingma and Welling, 2013) is a neural-
435 network based generative model. The framework of it is to first learn the
436 distribution P of the targeted data \mathbf{x} and then generate through sampling
437 with some unobserved variable \mathbf{z} , which is called the latent variable. And the
438 collection of them is named the latent space. In implementation, the learn-
439 ing of $P(\mathbf{x})$ is carried out with an assumed distribution $\int Q(\mathbf{z} \mid \mathbf{x})Q(\mathbf{x})d\mathbf{z}$
440 (Q is the assumed distribution of two parts. Since these two parts are im-
441 plemented in one neural network system, they share the same notation) ,
442 which is in the form of neural network. This distribution corresponds to two
443 important components of VAE: the encoder $\mathbf{z} = E(\mathbf{x})$ (For $Q(\mathbf{z} \mid \mathbf{x})$) and
444 decoder $\bar{\mathbf{x}} = D(\mathbf{z})$ (For $Q(\mathbf{x})$), where \mathbf{x} represents the input, $\bar{\mathbf{x}}$ for the gen-
445 erated(reconstructed). For particle generation, \mathbf{x} and $\bar{\mathbf{x}}$ refer to the shape
446 representation, for example, the Metaball descriptor \mathbf{M} or XRCT images.
447 The encoder and decoder consist of the major steps in VAE: encoding and
448 decoding. In encoding, the input shape representation \mathbf{x} is compressed and
449 mapped into the latent variable \mathbf{z} , a multidimensional shape-representation
450 tensor. Then \mathbf{z} is decoded to reconstruct the input particle $\bar{\mathbf{x}}$. Through
451 minimizing the difference between \mathbf{x} and $\bar{\mathbf{x}}$, morphologies and shape-feature
452 distribution of input particles can be learned effectively (The theory behind
453 this is briefly stated in Appendix A). Then, the trained decoder $\bar{\mathbf{x}} = D(\mathbf{z})$
454 can be applied to generate particles by inputting random \mathbf{z} . Assisted by the
455 powerful learning ability of neural network, VAE can inference new particles,

456 which are not included in the training set but maintain coessential morpho-
457 logical features and distributions with the parental particles. This make it
458 have the potential to provide a more practical solution to the particle gener-
459 ation task than previous studies.

460 It is worth noting that latent variables z in VAE are regularized (chosen
461 to be multivariable normal distribution) to encourage similar input samples
462 compress at closer positions in the latent space. This property allows the
463 model to learn a more flexible and general distribution, rather than sim-
464 ply adapting to the specific patterns present in the training data. As a
465 result, generation by sampling from sperate, regularized z can help to avoid
466 overfitting and underfitting problems on the shape-feature distributions of
467 generated particles(Jerves et al., 2017; Shi et al., 2021). More importantly,
468 it enables high-level controls on the generated particle morphologies(Macedo
469 et al., 2023).

470 Based on VAE, we propose a Metaball-based particle generation frame-
471 work, called Metaball Variational Autoencoder(MetaballVAE). It can resolve
472 the correlation between XRCT images and morphologies of input particles
473 with a regularized latent space, where complex calculations are converted
474 into one-step solutions, reducing variance in generated models and improv-
475 ing controls on generation process. This model requires no prior knowledge
476 (e.g. particle shape-feature distributions) but only XRCT images of the tar-
477 get particles to generate non-existent style-similar avatars, particles in the
478 form of Metaball, which can be put into simulations directly. Note that the
479 style-learning is not limited on the morphology but also the shape-feature
480 distribution, including surface area, volume, sphericity etc.

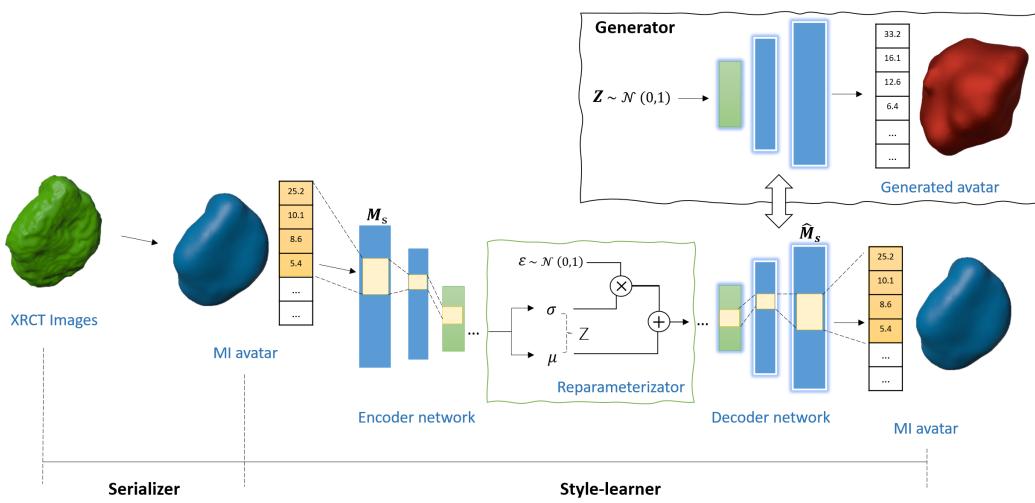


Figure 9: MetaballVAE for generation of complex-shaped particles. Serializer: interpreting and transforming XRCT images of granular matters. Style-learner: analysing the serialized Metaball descriptor and learning the morphological characteristics and distribution. Generator: generating style-similar granular media in Metaball form

481 The MetaballVAE consists of three major parts as illustrated in Fig-
482 ure 9: serializer, style-learner and generator. The serializer interprets and
483 transforms XRCT images of target granular-particles into Metaball descrip-
484 tors. Then, the style-leaner analyses those distilled descriptors, capture ma-
485 jor shape characteristics, conducts inference on feature distribution and de-
486 vise style-similiar avatars. In the end, the generator outputs designed style-
487 similar, Metaball-based avatars.

488 *4.2. The serializer*

489 The serializer is designed to abstract the particle morphology, extract
490 shape-feature distributions and code them structurally. It can significantly
491 reduce the dimension of XRCT images while keeping all the vital morpholog-
492 ical information for generation. The reasons for implementing it are two-fold.
493 On the one hand, structured data are more suitable for generation tasks(Para-
494 et al., 2021), which can improve the generation quality. On the other hand,
495 this enables a direct generation of particles in Metaball function form, which
496 can be put into simulations without bridging or transformation, avoiding
497 unnecessary information loss and computational cost.

498 In this paper, the serialization is accomplished with the Metaball-Imaging
499 technique as introduced in Section 2.2. The serialized particle is in MI avatar
500 form, which is noted as M_s .

501 *4.3. The style-learner*

502 The style-learner is a modification of the aforementioned VAE. It can
503 digest the structured data S and learn how to generate particles. Main

504 components involved are: encoder, decoder, reparameterizer, loss function
505 and distribution annealer.

506 **Encoder** and **Decoder** are multi-layer perceptions, which are connected
507 in a bottle-neck form as shown in Figure 9. They are implemented to ap-
508 proximate the real distribution $P(\mathbf{x})$ as a learnable, assumed distribution
509 $\int Q(\mathbf{z} | \mathbf{x})Q(\mathbf{x})dz$. On the topic of particle generation, \mathbf{x} refers to the \mathbf{M}_s .
510 The encoder takes serialized particles \mathbf{M}_s as input and outputs parameters (μ
511 - the mean, σ - the standard deviation) of the corresponding latent variable
512 \mathbf{z} , mapping morphologies and shape-feature distributions of particles into
513 a regularized latent space. On the contrary, the decoder interprets \mathbf{z} to
514 restore $\bar{\mathbf{M}}_s$, reconstructing particle morphologies and shape-feature distribu-
515 tions from that regularized space.

516 **Reparameterizer** locates halfway between the encoder and decoder.
517 It is designed to regularize the latent space by creating a map between the
518 encoded information and a normal distribution. Instead of direct sampling
519 (Fig, 10, a), a more deterministic pattern is utilized:

$$z = \mu + \sigma \odot \epsilon \quad (8)$$

520 where the ϵ is the assumed normal distribution. This enables continuous
521 gradient calculation on the mapping relationship, making MetaballVAE a
522 learnable system.

523 **The loss function** is the global optimization objective for MetaballVAE:

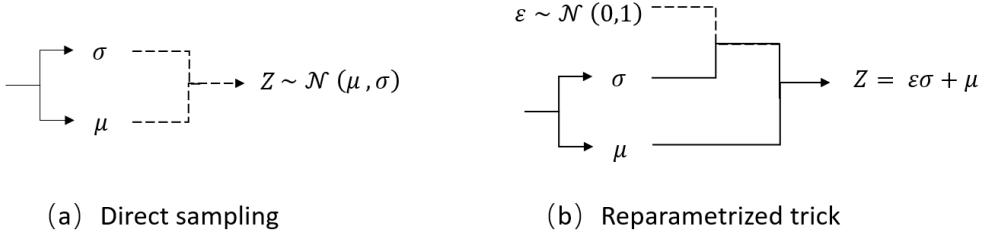


Figure 10: Visualizations of the direct sampling and reparameterization trick. The solid line stands for relationship capable of back propagation. The dash line represent relationship where backpropagation can not be carried out.

$$L(\mathbf{M}_s) = \underbrace{\frac{1}{d} \sum_{k=1}^d \|\mathbf{M}_s - \hat{\mathbf{M}}_s\|^2}_{\text{Reconstruction Item}} + \underbrace{\frac{1}{2} \sum_{k=1}^d \left(\mu_{(k)}^2(\mathbf{M}_s) + \sigma_{(k)}^2(\mathbf{M}_s) - \ln \sigma_{(k)}^2(\mathbf{M}_s) - 1 \right)}_{\text{Distribution Item}} \quad (9)$$

524 where d is the dimension of \mathbf{M}_s . This function is modified from the original
 525 VAE theory based on the particle generation problem. The deduction of it
 526 is stated in Appendix B. It consists of two items: the distribution item and
 527 reconstruction item. The distribution Item measures the difference between
 528 real and learned distributions of particle morphologies. The Reconstruction
 529 Item evaluates the quality of learned morphological characteristics. The com-
 530 bination of them forces MetaballVAE to learn not only morphologies but also
 531 shape-feature distributions of the input particles.

532 **The distribution annealer** is proposed to tackle the training challenge
 533 of VAE. A well-trained model possesses a relatively small reconstruction item
 534 and a non-zero distribution item. However, most direct training will yield

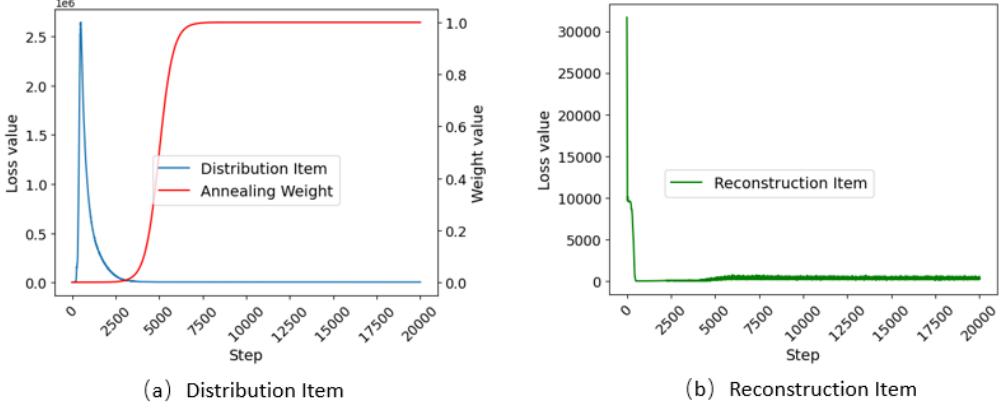


Figure 11: The impact of the distribution annealer on the loss value of different items in the training of cobblestone dataset

535 a model with a zero distribution item. Such tendency in learning is caused
 536 by the sensitivity of decoder to variation introduced by the mapping process
 537 of reparameterizer. This makes the decoder ignore the latent variable pro-
 538 vided by the encoder and output the average optimal with distribution item
 539 equal to zero. For this reason, the distribution annealer is implemented by
 540 adding a weight to the distribution item of the loss function(Eq. 9). This
 541 weight starts from zero, where the weighted loss function equals the recon-
 542 struction loss. Then, the weight is increased gradually to one, where this
 543 weighted function satisfies the true loss function definition. With such a pro-
 544 cess, the model will be forced to use the learned latent space to achieve good
 545 likelihood in prediction.

546 Figure 11 is an example of the distribution annealer in the training of the
 547 cobblestone dataset during the first 20k steps. It can be observed that the
 548 distribution item first spikes as the reconstruction item drops significantly,

549 where the model is encoding shape features into the latent space cheaply.
550 Then, the distribution item starts to decrease rapidly as more attention is
551 paid to the divergence penalty. Correspondingly, the decrease of reconstruc-
552 tion item slows down. Finally, the distribution item gradually converges
553 and the reconstruction item enters fluctuation, where more morphological
554 information is compressed into the model.

555 *4.4. The generator*

556 Before formal generation, the decoder of style-learner should be well-
557 trained. The generation task requires the trained decoder and a normal
558 distribution $N(0, 1)$, which represents the regularized latent space. A typical
559 generation process is illustrated in Figure 9. The latent variable z is sampled
560 from a normal distribution, serving as the input matrix to the decoder. Then,
561 the decoder can devise style-similar particles unseen in the training dataset.
562 The distribution of shape features can be well reconstructed when the number
563 of generated particles is large enough. It is worth noting that the generated
564 particle is in the form of Metaball descriptor, which can be applied directly
565 into simulations.

566 **5. Evaluation of MetaballVAE**

567 *5.1. Dataset and Metrics*

568 Previous studies on particle generation are often carried out on hundreds
569 of thousands of samples(Shi et al., 2021; Medina and Jerves, 2019; Macedo
570 et al., 2023). However, particle reconstruction with XRCT requires consider-
571 able time and computational resources. In actual engineering, it is very often

572 to have only a dozen scanned particles. Therefore, it's crucial to evaluate the
573 performance of algorithms on smaller datasets. To this end, we tested the
574 effectiveness of the MetaballVAE model on four distinct sets of XRCT data,
575 which included particles of different types and sizes: 290, 100, and 10 sam-
576 ples of Ottawa sands, as well as 20 cobblestones. These testing datasets were
577 selected randomly from the Ottawa sand and cobblestone datasets listed in
578 Section 3.1. For better learning performance, data augmentation are im-
579 plemented on training datasets, where slightly modified synthetic data is
580 introduced based on the real one. Here, particle rotating and parameter
581 shuffling are implemented. Particle rotating is a popular strategy based on
582 rotational-invariant property. For example, Shi et al. (Shi et al., 2021) ap-
583 plied nine rotations to each particle and enlarged his dataset by ten times in a
584 particle generation task, which effectively enhanced the model performance.
585 Parameter shuffling means random recombination of $\{k_i, \mathbf{x}_i\}$ in the serialized
586 particle \mathbf{M}_s . This is because the sequence change of control spheres will not
587 modify the corresponding Metaball model. Such processing can effectively
588 avoid the overfitting problem and enhance convergence performance. During
589 augmentation, each particle is rotated 5 times and the corresponding Meta-
590 ball parameter shuffled 50 times. The augmented datasets then each contain
591 145,000 Ottawa sand and 6,000 cobblestone samples.

592 Accurate evaluation is a challenge for particle generation tasks. Apart
593 from the rationality of particle shape, another important content of eval-
594 uation is the quantitative difference between parents and clones. Thus, the
595 same metrics used in reconstruction are also adopted here for further verifi-
596 cation on the shape feature distributions.

597 *5.2. The setting of serializer and style-learner*

598 In this evaluation, we apply the following hyper-parameter setting. In se-
599 rializer, the control point number, n, is set to be 40 for both cobblestone and
600 Ottawa sand samples. The learning rate for the gradient update is set to be
601 0.001. In style-learner, the encoder is a 4 layer full-connected network with
602 leaky ReLU activation function. The size of it is: $160 \times 1024 \times 512 \times 256 \times 128$.
603 The decoder is also a 4 layer full-connected network with leaky ReLU ac-
604 tivation function yet in reverse form. The reparameterizator is set to be
605 one fully connected layer with size 128. The above networks are trained by
606 Adam(Zhang, 2018; Wang et al., 2018) with learning rate $\eta = 0.0001$.

607 Since the setting of hyper-parameters is not a focus of this paper, how to
608 obtain them is not included here for the sake of brevity. A detailed procedure
609 can be referenced in (Shi et al., 2021).

610 *5.3. Particle Generation*

611 In validation, the training datasets are denoted as "Parents". And 1000
612 particle avatar are devised by the generator, denoted as "Clones". The num-
613 ber of generated particles is set to larger for better evaluation on the mor-
614 phological distribution.

615 Figure 12 and Figure 13 displays several cloned examples of each type. It
616 can be concluded that these clones exhibit reasonable shape features of both
617 Ottawa sands (angled features) and cobblestones (round features). Note
618 that these particles are in Metaball function form and the meshes are only
619 for visualization.

620 To further exam the regeneration ability of the proposed method on
621 shape-feature distributions, probability density functions (PDF) of selected

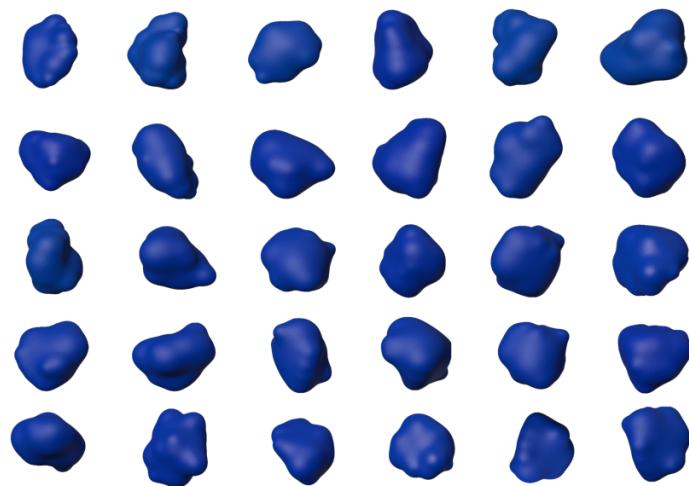


Figure 12: Examples of cloned Ottawa sands

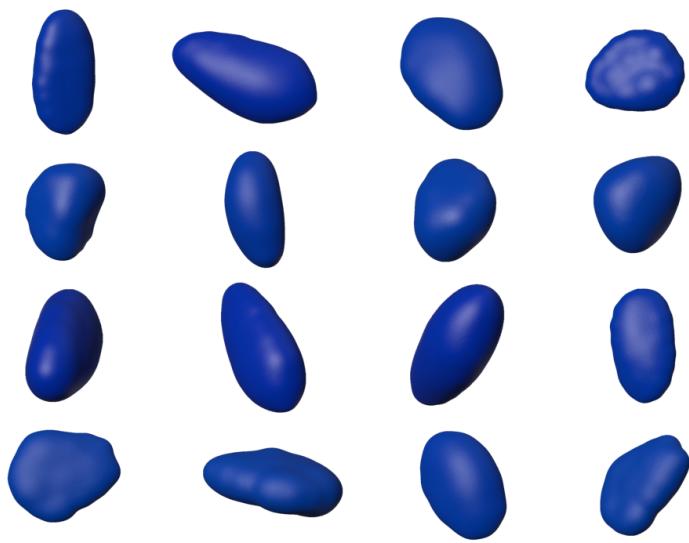


Figure 13: Examples of cloned cobblestone

622 metrics are calculated on both parents and their clones (See Figure 14 to 17).

623 All shape features of parents and clones in four datasets of Ottawa-sand
624 and cobblestone datasets share similar distributions. In term of ϕ , we observe
625 deviation errors of 2.54%, 1.33%, 1.09% and 0.92% on means between parents
626 and clones, for the four datasets separately. In the case of standard deviation,
627 errors of 3.12%, 5.90%, 3.77% and 4.45% are obtained. For C of the Ottawa
628 sand and cobblestone, the distribution means are off with errors of 0.03%,
629 0.01%, 0.10% and 0.25%. While errors in standard deviation are 2.56%,
630 8.32%, 4.41% and 3.46% separately. In the case of D_{ns} , errors of 2.57%,
631 0.67%, 1.29% and 0.98% are observed on the mean misalignments for the
632 Ottawa sand and cobblestone. And the errors coming from the standard
633 deviation are 2.69%, 1.53%, 9.35% and 5.87%. As for CSF , the distributions
634 have errors of 3.87%, 5.31%, 2.80% and 3.25% on the mean, as well as errors
635 of 11.58%, 10.70%, 1.94% and 4.88% on the standard deviations. Finally, we
636 obtain errors of 2.77%, 6.49%, 3.72% and 0.66% from V , as well as errors
637 of 4.22%, 7.45%, 3.50% and 1.41% from A on means of distributions of the
638 Ottawa sand and cobblestone. In the case of standard deviation, the errors
639 are 20.61%, 9.87%, 1.88% and 5.64% given by V , as well as 24.21%, 14.94%,
640 6.5% and 10.65% given by A .

641 Note that the MetaballVAE is capable of learning and representing non-
642 Gaussian shape-feature distributions. Examples of this can be seen in the
643 histogram of the feature V and A of the 10 Ottawa sands and 20 cobble-
644 stone(Figure 17 and 16), where multiple peaks are observed in the distri-
645 bution (although they may not be clear in the PDF curve due to limited
646 samples). The MetaballVAE capture this feature effectively and generate

647 particles with similar distributions, demonstrates the effectiveness of the
648 MetaballVAE in cloning grains of complex shape-feature distributions. While
649 an interesting phenomenon arises in the smalleset dataset of 10 Ottawa sand
650 (Figure 16). It can be observed that the distributions of shape features of
651 parental particles are discrete and discontinuous. Although the shape-feature
652 distributions of the clones match well with the parents, theirs distribution
653 are continuous with peaks around the discrete values of the parents. This
654 is because MetaballVAE is designed to interpret and map the discrete par-
655 ticle morpholgoies into continuous latent space. Such architecture can not
656 only facilitates the generation of a large variety of particles with similar
657 shape-feature distributions and allows for more efficient exploration and ma-
658 nipulation of the generated particles, which will be discussed in Section 5.4.

659 Compared with previous studies (Zhou and Wang, 2017; Jerves et al.,
660 2017; Medina and Jerves, 2019; Macedo et al., 2023), MetaballVAE does not
661 require prior knowledge of particle morphologies or shape-feature distribu-
662 tions as input, but only XRCT images of different sizes. Despite this, the
663 satisfactory comparison results show that MetaballVAE is capable of accu-
664 rately capturing overall shape features from parental XRCT particles and
665 cloning new, morphologically similar avatars with similar feature distribu-
666 tions. Furthermore, this algorithm is effective even with small datasets of
667 only a few dozen samples, which is often the case in engineering problems.

668 Besides, the above system is able to learn continuously. XRCT images
669 can be kept added to the training dataset to improve the model perfor-
670 mance. After short training (5 - 10mins), new morphological characteristics
671 and shape-feature distributions can be learned by the original models.

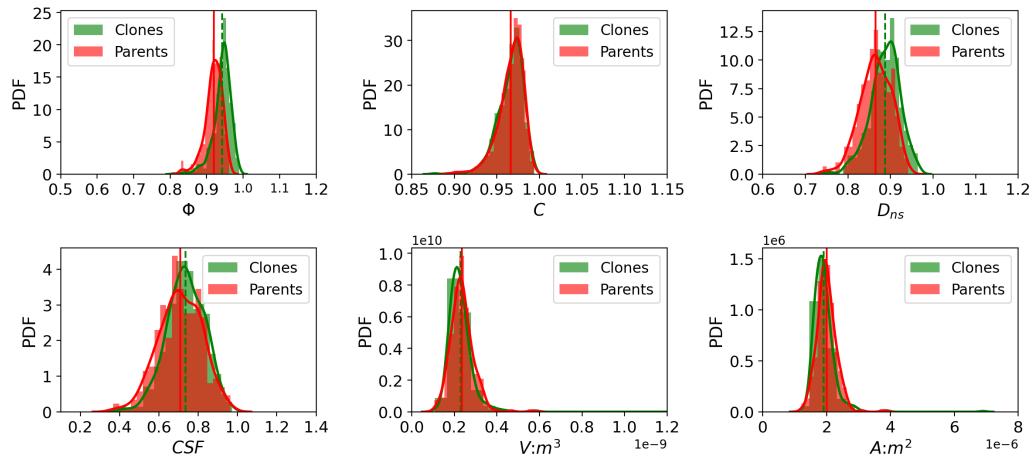


Figure 14: Comparison of feature distributions between parental and cloned particles in the Ottawa sand dataset of size 290

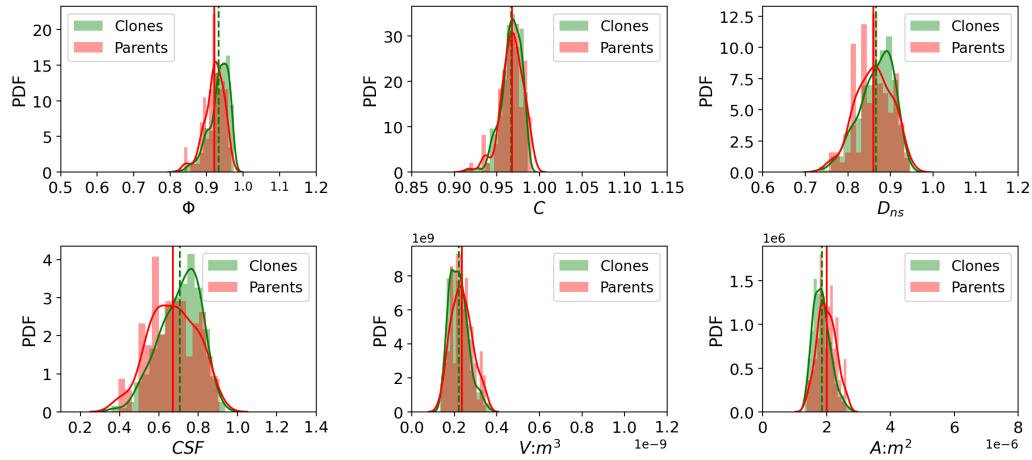


Figure 15: Comparison of feature distributions between parental and cloned particles in the Ottawa sand dataset of size 100

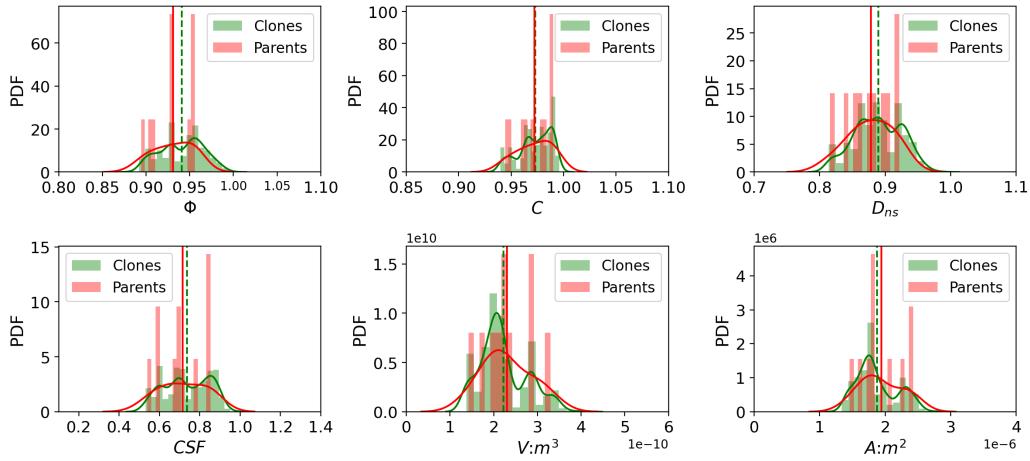


Figure 16: Comparison of feature distributions between parental and cloned particles in the Ottawa sand dataset of size 10

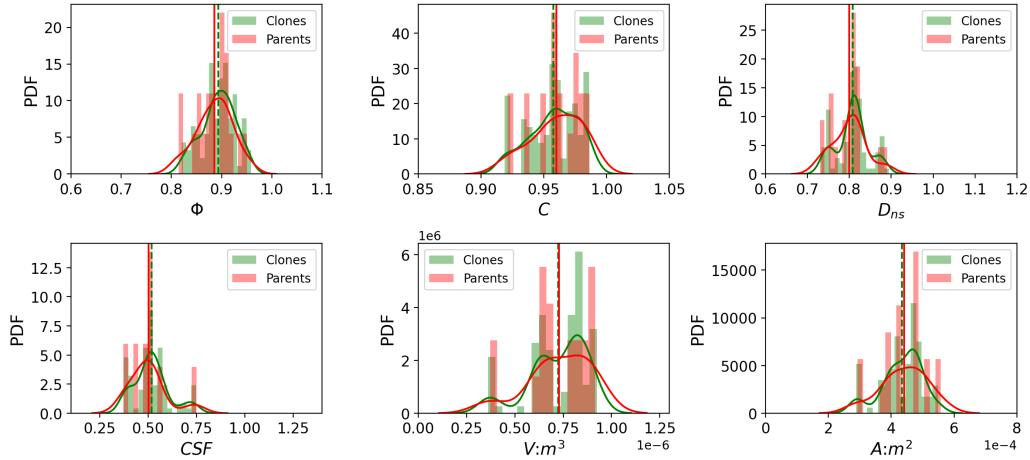


Figure 17: Comparison of feature distributions between parental and cloned particles in the cobblestone dataset of size 20

672 5.4. Latent space arithmetic

673 A well-trained MetaballVAE can achieve integration of discrete training
674 samples into continuous, regularized latent space, where the number of gener-
675 ated particles can be infinite. Such regularized latent space avoids overfitting
676 and underfitting problems in generation, ensuring the rationality of generated
677 morphologies(Shi et al., 2021). More importantly, it also realizes a certain
678 control on the generated geometric feature(Zamorski et al., 2020).

679 Figure 18 illustrates that the addition of Gaussian random noises δ into
680 the latent variable \mathbf{z} can introduce morphological changes of different degrees
681 into generated avatars. Here, δ is set to have the same dimension of \mathbf{z} , with
682 zero mean value and different variance σ . Then, δ s are added to a randomly
683 selected latent variable \mathbf{z} to produce modified ones $\mathbf{z} + \delta$. Finally, these
684 latent variables are fed into the generator as inputs. From the corresponding
685 generated results, it can be observed that the addition of δ s with small σ can
686 slightly adjust the particle morphologies. As the increase in σ , the degree of
687 modification becomes larger, resulting in less similar avatars to the original
688 one. This is from the property of regularized latent space. The addition of
689 δ s can create new latent variable adjacent to the original one, while the σ
690 of δ s decides distances between them in the latent space. Since the latent
691 space is regularized, such adjacent relationships control the shape similarity
692 in generated avatars. This phenomenon is also observed in generating digital
693 sand particles with VAE(Shi et al., 2021). It can be very useful when particles
694 in certain morphologies are needed in simulations. We can first select the
695 template avatars and then add δ s of small magnitude into its latent variables.
696 In this way, slightly modified avatars can be generated, avoiding repetitive

697 particle morphologies in simulation.

698 Figure 19 indicates that interpolation between latent variables can pro-
699 duce smooth shape transitions in corresponding generated avatars. In these
700 examples, two latent variable z_1 and z_2 are randomly selected to create in-
701 terpolations $z_1 + \alpha(z_2 - z_1)$. Then, these latent variables are fed as input
702 to the generator. It is clear that as the increase of α , those interpolated
703 avatars gradually transform from z_1 avatar to z_2 avatar. Note that such
704 a change occurs simultaneously in multiple characteristics including shape,
705 volume and surface area. This is also result of the regularized latent space.
706 Those interpolated variables possess adjacent locations in the latent space
707 with z_1 and z_2 , resulting in avatars of similar shape. With change in the
708 location of latent variables, the generated avatars show smooth modification
709 in shape from z_1 to z_2 avatars. This phenomenon can be applied to obtain
710 avatars of combined features. We can first select two template avatars of
711 specific morphologies then do interpolation between them.

712 Figure 20 shows that the avatar shape can be modified by applying ad-
713 dition or subtraction in the latent space. Here, arithmetic operations are
714 implemented on latent variables, z_1 and z_2 , corresponding to avatars of dis-
715 tinct shapes. Under such operation, specific shape features can be added or
716 removed from the generated avatar of z_3 . This also results from regularized
717 latent space. The proper mapping between latent variables and morpho-
718 logical features provides a powerful method to modify the shape. Through
719 proper arithmetic operation, avatars of specific features can be generated
720 according to need.

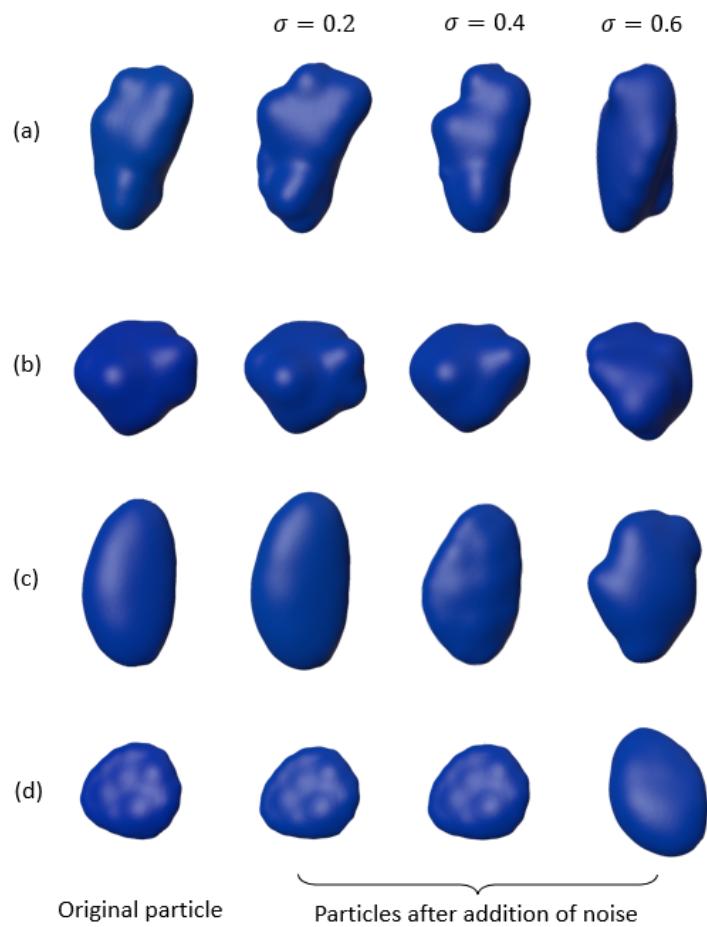


Figure 18: Comparison of generated particles before and after addition of Gaussian noise in the latent variables. (a) and (b) are Ottawa sand examples. (c) and (d) are cobblestone examples. The variance of Gaussian noise is noted as σ

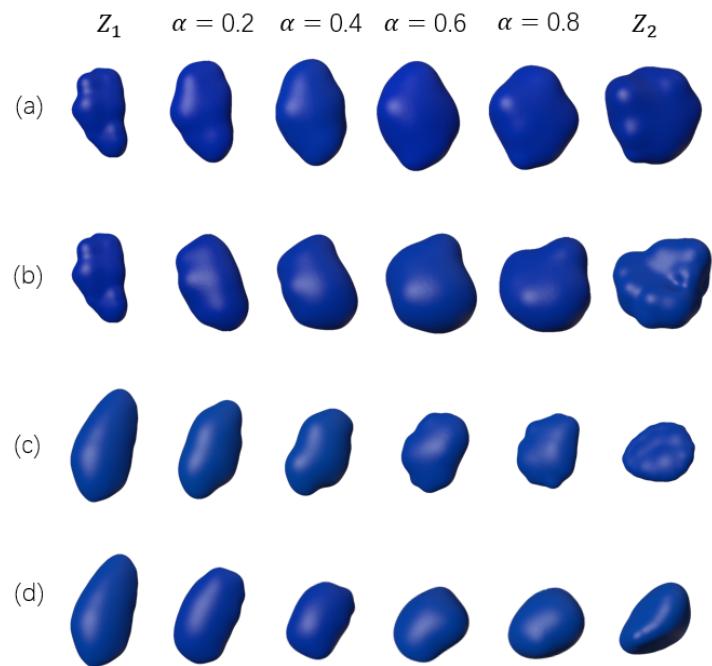


Figure 19: Comparison of generated particles of interpolated latent variables. (a) and (b) are cobblestone examples of the same z_1 . (c) and (d) are Ottawa sand examples of the same z_1 . The interpolated coefficient is noted as α .

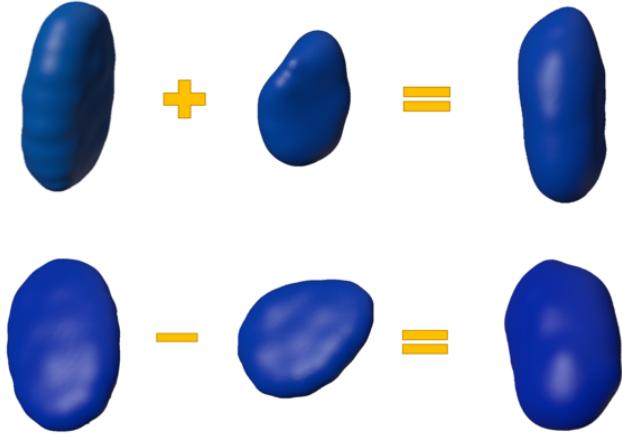


Figure 20: Manipulation of particle shape through addition and subtraction in latent variables. Top: adding angled features into a smooth cobblestone. Bottom: removing flattening features from a thin cobblestone.

721 6. Conclusions

- 722 (1) We propose a geometric-based Metaball-Imaging (MI) algorithm, which
- 723 can transform X-Ray Computed Tomography (XRCT) images of real parti-
- 724 cles into Metaball-based avatars.
- 725 (2) The impact of the core hyperparameter, the number of control points,
- 726 on particle characterization was carefully investigated. It was discovered that
- 727 the characterization fidelity enhance with the increase of control-point num-
- 728 ber until reaching equilibrium. For major morphological features of common
- 729 particles, 40 control points are sufficient to do characterization. As for par-
- 730 ticles with more complex shapes, such as angular or concave features, only
- 731 120 control points are required to accurately represent these features.
- 732 (3) By comparing four types of commonly-encountered particles with dis-
- 733 tinct characteristics, the efficacy and precision of Metaball imaging (MI) in

734 particle reconstruction are verified. The Metaball descriptor has been found
735 to excel at depicting granular materials with round, smooth characteristics,
736 such as cobblestones. The use of a lightweight Metaball avatar with only a
737 handful of control points has been shown to be sufficient for creating high-
738 fidelity representations of such particles. Overall, the findings of this evalua-
739 tion suggest that MI is a valuable tool for accurately characterizing particles
740 with distinct features.

741 (4) We also present a variational-autoencoder (VAE) based particle gen-
742 eration algorithm called MetaballVAE. It can generate style-similar particles
743 in Metaball function form with XRCT images of parental particles.

744 (5) The MetaballVAE was evaluated through a comparison of two groups
745 of particles with different sizes. It was found that the parental and cloned
746 particles exhibited good agreement in terms of their morphologies and shape-
747 feature distributions. These results provide evidence that MetaballVAE is a
748 reliable and practical tool for characterizing particles with varying sizes and
749 morphologies.

750 (6) The regularized latent space of MetaballVAE allows for control over
751 the generation process. Particles with specific morphologies can be gener-
752 ated through arithmetic operations on the latent space. This feature makes
753 MetaballVAE a versatile and useful tool for generating particles with desired
754 characteristics.

755 (7) The authors have previously developed metaball-based simulation
756 frameworks that have proven to be a powerful tool in comprehending the in-
757 tricacies of fluid-particle systems involving realistic soil particles(Zhao et al.,
758 2023; Zhang et al., 2021). With the addition of reconstruction and genera-

759 tion methods proposed in this paper, the integration of these tools has the
760 potential to unveil new insights into soil mechanics and provide valuable in-
761 formation for a wide range of applications, including soil erosion modeling,
762 soil contamination analysis, and soil moisture modeling.

763 **Acknowledgement**

764 We gratefully acknowledge the funds from National Natural Science Foun-
765 dation of China (Project No.12172305), Natural Science Foundation of Zhe-
766 jiang Province, China (LHZ21E090002), Key Research and Development Pro-
767 gram of Zhejiang Province (Grant No.2021C02048) and Westlake University.
768 We also thank Westlake High-performance Computing Center for computa-
769 tional sources and corresponding assistance.

770 **Appendix A. The theory of VAE**

771 VAE is designed to learn the distribution of the input data x , which is
772 controlled by some unobserved latent variable z :

$$P(x) = \int P(x, z) dz \quad (.1)$$

773 Since the above distribution is hard to learn directly for high dimensional
774 problems, VAE approximates $P(x)$ with an assumed distribution $Q_\phi(x)$. The
775 parameter ϕ is obtained through minimzing the KL divergence() between the
776 real joint distribution $P(x, z)$ and the assumed one $Q_\phi(x, z)$:

$$\arg \min_{\phi} KL(Q_\phi(x, z) || P(x, z)) \quad (.2)$$

777 The above divergence can be rewritten as:

$$KL(P(x, z) \| Q(x, z)) = \mathbb{E}_{x \sim P(x)} [\ln P(x)] + \mathbb{E}_{x \sim P(x)} \left[\int P(z | x) \ln \frac{P(z | x)}{Q(z, x)} dz \right] \quad (.3)$$

778 where the first item is a constant. Then, a lower bound \mathcal{L} of this divergence,
 779 which is called the evidence lower bound in variational inference, can be
 780 obtained:

$$\begin{aligned} \mathcal{L} &= KL(Q_\phi(x, z) || P(x, z)) - \text{constant} \\ &= \mathbb{E}_{x \sim P(x)} \left[\int P(z | x) \ln \frac{P(z | x)}{Q(z, x)} dz \right] \end{aligned} \quad (.4)$$

781 This transforms the optimization objective(Eq. .2) into minimizing the
 782 above \mathcal{L} .

783 To simplify the calculation of \mathcal{L} , we further apply the variational infer-
 784 ence:

$$\mathcal{L} = \underbrace{\mathbb{E}_{x \sim P(x)} [\mathbb{E}_{z \sim P(z|x)} [-\ln Q(z | x)]]}_{\text{Reconstruction Item}} + \underbrace{\mathbb{E}_{x \sim P(x)} [KL(P(z | x) || Q(z))] }_{\text{Distribution Item}} \quad (.5)$$

785 This joint expression consists of tow items: the reconstruction item and
 786 the distribution item. The first item evaluates the reconstructive performance
 787 of the input and the second item measures the distribution similarity between
 788 the generation and the input.

789 Appendix B. Derivation of the loss function

790 The form of loss function for MetaballVAE is determined from the evi-
 791 dence lower bound of VAE(Eq. .5). In this equation, there are three unde-
 792 cided items: $P(z)$, $Q(x | z)$, and $P(z | x)$. We assume z to be multivariable
 793 normal distribution for regularization, which provides controls on the gener-
 794 ated shape.

795 In present study, x refers to the shape representation, Metaball descriptor.
 796 Thus, Gaussian distribution is taken as the distribution of $Q(x | z)$. A neural
 797 network is adopted to represent its average $\mu(z)$ and variance $\sigma(z)$ with the
 798 reparameterizer, which gives:

$$Q(x | z) = \frac{1}{\prod_{k=1}^D \sqrt{2\pi\tilde{\sigma}_{(k)}^2(z)}} \exp\left(-\frac{1}{2} \left\| \frac{x - \tilde{\mu}(z)}{\tilde{\sigma}(z)} \right\|^2\right) \quad (.6)$$

799 where D is the dimension of latent variable z .

800 This allows $-\ln Q(x | z)$ to be transformed into:

$$-\ln Q(x | z) = \frac{1}{2} \left\| \frac{x - \tilde{\mu}(z)}{\tilde{\sigma}(z)} \right\|^2 + \frac{D}{2} \ln 2\pi + \frac{1}{2} \sum_{k=1}^D \ln \tilde{\sigma}_{(k)}^2(z) \quad (.7)$$

801 In actual implementations, the variance $\sigma(z)$ is taken as a constant. This
 802 simplifies Eq. .7 into:

$$-\ln Q(x | z) \sim \frac{1}{2\sigma^2} \|x - \tilde{\mu}(z)\|^2 \quad (.8)$$

803 The above deductions enables approximation of the Reconstruction item
 804 with Mean Square Error form:

$$\mathbb{E}_{x \sim P(x)} [\mathbb{E}_{z \sim P(z|x)} [-\ln Q(x | z)]] \sim \frac{1}{d} \sum_{k=1}^d \|x - \hat{x}\|^2 \quad (.9)$$

805 where the d stands for the dimension of input x .

806 We also assume $P(z | x)$ to be Gaussian. Similarly, a neural network is
 807 adopted to decide it:

$$P(z | x) = \frac{1}{\prod_{k=1}^d \sqrt{2\pi\sigma_{(k)}^2(x)}} \exp\left(-\frac{1}{2} \left\| \frac{z - \mu(x)}{\sigma(x)} \right\|^2\right) \quad (.10)$$

808 where $\mu(x)$ and $\sigma(x)$ are the corresponding average and variance. This allows
809 the Distribution item to be expressed as:

$$KL(p(z | x) \| q(z)) = \frac{1}{2} \sum_{k=1}^d \left(\mu_{(k)}^2(x) + \sigma_{(k)}^2(x) - \ln \sigma_{(k)}^2(x) - 1 \right) \quad (11)$$

810 **References**

- 811 Bailey, D.G., 2004. An efficient euclidean distance transform, in: Interna-
812 tional workshop on combinatorial image analysis, Springer. pp. 394–408.
- 813 Blinn, J.F., 1982. A generalization of algebraic surface drawing. *j-tog*, 1 (3):
814 235–256.
- 815 Bourilkov, D., 2019. Machine and deep learning applications in particle
816 physics. *International Journal of Modern Physics A* 34, 1930019.
- 817 Bouwman, A.M., Bosma, J.C., Vonk, P., Wesselingh, J.H.A., Frijlink, H.W.,
818 2004. Which shape factor (s) best describe granules? *Powder Technology*
819 146, 66–72.
- 820 Chen, J., Li, R., Mo, P.Q., Zhou, G., Cai, S., Chen, D., 2022. A modi-
821 fied method for morphology quantification and generation of 2d granular
822 particles. *Granular Matter* 24, 1–18.
- 823 Chen, R.P., Liu, Q.W., Wu, H.N., Wang, H.L., Meng, F.Y., 2020. Effect
824 of particle shape on the development of 2d soil arching. *Computers and*
825 *Geotechnics* 125, 103662.
- 826 Coetzee, C., 2016. Calibration of the discrete element method and the effect
827 of particle shape. *Powder Technology* 297, 50–70.

- 828 Cundall, P.A., Strack, O.D., 1979. A discrete numerical model for granular
829 assemblies. *geotechnique* 29, 47–65.
- 830 Dietrich, W.E., 1982. Settling velocity of natural particles. *Water resources*
831 research 18, 1615–1626.
- 832 Galindo-Torres, S., 2013. A coupled discrete element lattice boltzmann
833 method for the simulation of fluid–solid interaction with particles of gen-
834 eral shapes. *Computer Methods in Applied Mechanics and Engineering*
835 265, 107–119.
- 836 Garcia, X., Latham, J.P., XIANG, J.s., Harrison, J., 2009. A clustered
837 overlapping sphere algorithm to represent real particles in discrete element
838 modelling. *Geotechnique* 59, 779–784.
- 839 Golombek, M., Warner, N.H., Grant, J.A., Hauber, E., Ansan, V., Weitz,
840 C.M., Williams, N., Charalambous, C., Wilson, S.A., DeMott, A., et al.,
841 2020. Geology of the insight landing site on mars. *Nature communications*
842 11, 1014.
- 843 Gong, J., Nie, Z., Zhu, Y., Liang, Z., Wang, X., 2019. Exploring the effects
844 of particle shape and content of fines on the shear behavior of sand-fines
845 mixtures via the dem. *Computers and Geotechnics* 106, 161–176.
- 846 Grabowski, A., Nitka, M., Tejchman, J., 2021. Comparative 3d dem sim-
847 ulations of sand–structure interfaces with similarly shaped clumps versus
848 spheres with contact moments. *Acta Geotechnica* 16, 3533–3554.

- 849 Grigoriu, M., Garboczi, E., Kafali, C., 2006. Spherical harmonic-based ran-
850 dom fields for aggregates used in concrete. Powder Technology 166, 123–
851 138.
- 852 Higgins, I., Matthey, L., Pal, A., Burgess, C., Glorot, X., Botvinick, M.,
853 Mohamed, S., Lerchner, A., 2016. beta-vae: Learning basic visual concepts
854 with a constrained variational framework .
- 855 Höhner, D., Wirtz, S., Scherer, V., 2012. A numerical study on the influence
856 of particle shape on hopper discharge within the polyhedral and multi-
857 sphere discrete element method. Powder technology 226, 16–28.
- 858 Iwashita, K., Oda, M., 1998. Rolling resistance at contacts in simulation of
859 shear band development by dem. Journal of engineering mechanics 124,
860 285–292.
- 861 Jerves, A.X., Kawamoto, R.Y., Andrade, J.E., 2017. A geometry-based al-
862 gorithm for cloning real grains. Granular Matter 19, 1–10.
- 863 Jiang, M., Yu, H.S., Harris, D., 2005. A novel discrete model for granular
864 material incorporating rolling resistance. Computers and Geotechnics 32,
865 340–357.
- 866 Jin, X., Li, Y., Peng, Q., 2000. General constrained deformations based on
867 generalized metaballs. Computers & Graphics 24, 219–231.
- 868 Kawamoto, R., Andò, E., Viggiani, G., Andrade, J.E., 2016. Level set dis-
869 crete element method for three-dimensional computations with triaxial
870 case study. Journal of the Mechanics and Physics of Solids 91, 1–13.

- 871 Kingma, D.P., Welling, M., 2013. Auto-encoding variational bayes. arXiv
872 preprint arXiv:1312.6114 .
- 873 Lai, Z., Zhao, S., Zhao, J., Huang, L., 2022. Signed distance field framework
874 for unified dem modeling of granular media with arbitrary particle shapes.
875 Computational Mechanics , 1–21.
- 876 Liu, X., Garboczi, E., Grigoriu, M., Lu, Y., Erdogan, S.T., 2011. Spherical
877 harmonic-based random fields based on real particle 3d data: improved
878 numerical algorithm and quantitative comparison to real particles. Powder
879 Technology 207, 78–86.
- 880 Lu, Z., Yao, A., Su, A., Ren, X., Liu, Q., Dong, S., 2019. Re-recognizing the
881 impact of particle shape on physical and mechanical properties of sandy
882 soils: a numerical study. Engineering Geology 253, 36–46.
- 883 Buarque de Macedo, R., Marshall, J.P., Andrade, J.E., 2018. Granular ob-
884 ject morphological generation with genetic algorithms for discrete element
885 simulations. Granular Matter 20, 1–12.
- 886 Macedo, R., Monfared, S., Karapiperis, K., Andrade, J., 2023. What is
887 shape? characterizing particle morphology with genetic algorithms and
888 deep generative models. Granular Matter 25, 1–12.
- 889 Medina, D.A., Jerves, A.X., 2019. A geometry-based algorithm for cloning
890 real grains 2.0. Granular Matter 21, 1–15.
- 891 Mollon, G., Zhao, J., 2012. Fourier–voronoi-based generation of realistic
892 samples for discrete modelling of granular materials. Granular matter 14,
893 621–638.

- 894 Mollon, G., Zhao, J., 2013. Generating realistic 3d sand particles using fourier
895 descriptors. *Granular Matter* 15, 95–108.
- 896 Mollon, G., Zhao, J., 2014. 3d generation of realistic granular samples based
897 on random fields theory and fourier shape descriptors. *Computer Methods*
898 in Applied Mechanics and Engineering
- 279, 46–65.
- 899 Mora, C., Kwan, A., 2000. Sphericity, shape factor, and convexity mea-
900 surement of coarse aggregate for concrete using digital image processing.
901 *Cement and concrete research* 30, 351–358.
- 902 Nie, J.Y., Li, D.Q., Cao, Z.J., Zhou, B., Zhang, A.J., 2020. Probabilistic
903 characterization and simulation of realistic particle shape based on sphere
904 harmonic representation and nataf transformation. *Powder Technology*
905 360, 209–220.
- 906 Para, W., Bhat, S., Guerrero, P., Kelly, T., Mitra, N., Guibas, L.J., Wonka,
907 P., 2021. Sketchgen: Generating constrained cad sketches. *Advances in*
908 *Neural Information Processing Systems* 34, 5077–5088.
- 909 Plankers, R., Fua, P., 2001. Articulated soft objects for video-based body
910 modeling, in: *Proceedings Eighth IEEE International Conference on Com-*
911 *puter Vision. ICCV 2001, IEEE.* pp. 394–401.
- 912 Regueiro, R., Zhang, B., Shahabi, F., Soga, K., 2014. Micromorphic contin-
913 uum stress measures calculated from three-dimensional ellipsoidal discrete
914 element simulations on granular media. *IS-Cambridge* 2014, 1–6.
- 915 Santamarina, J., Cho, G.C., 2004. Soil behaviour: The role of particle shape,

- 916 in: Advances in geotechnical engineering: The Skempton conference: Pro-
917 ceedings of a three day conference on advances in geotechnical engineering,
918 organised by the Institution of Civil Engineers and held at the Royal Ge-
919 ographical Society, London, UK, on 29–31 March 2004, Thomas Telford
920 Publishing. pp. 604–617.
- 921 Shi, J.j., Zhang, W., Wang, W., Sun, Y.h., Xu, C.y., Zhu, H.h., Sun, Z.x.,
922 2021. Randomly generating three-dimensional realistic schistous sand par-
923 ticles using deep learning: Variational autoencoder implementation. Engi-
924 neering Geology 291, 106235.
- 925 Shinohara, K., Oida, M., Golman, B., 2000. Effect of particle shape on angle
926 of internal friction by triaxial compression test. Powder technology 107,
927 131–136.
- 928 Sommer, C., Straehle, C., Koethe, U., Hamprecht, F.A., 2011. Ilastik: In-
929 teractive learning and segmentation toolkit, in: 2011 IEEE international
930 symposium on biomedical imaging: From nano to macro, IEEE. pp. 230–
931 233.
- 932 Su, D., Yan, W., 2018. 3d characterization of general-shape sand particles
933 using microfocus x-ray computed tomography and spherical harmonic func-
934 tions, and particle regeneration using multivariate random vector. Powder
935 Technology 323, 8–23.
- 936 Sun, Q., Zheng, J., 2021. Realistic soil particle generation based on lim-
937 ited morphological information by probability-based spherical harmonics.
938 Computational Particle Mechanics 8, 215–235. Publisher: Springer.

- 939 Tan, Q., Gao, L., Lai, Y.K., Xia, S., 2018. Variational autoencoders for
940 deforming 3d mesh models, in: Proceedings of the IEEE conference on
941 computer vision and pattern recognition, pp. 5841–5850.
- 942 Thomas, M., Wiltshire, R., Williams, A., 1995. The use of fourier descriptors
943 in the classification of particle shape. *Sedimentology* 42, 635–645.
- 944 Tolomeo, M., McDowell, G.R., 2022. Modelling real particle shape in dem:
945 a comparison of two methods with application to railway ballast. *International
946 Journal of Rock Mechanics and Mining Sciences* 159, 105221.
- 947 Vlahinić, I., Kawamoto, R., Andò, E., Viggiani, G., Andrade, J.E., 2017.
948 From computed tomography to mechanics of granular materials via level
949 set bridge. *Acta Geotechnica* 12, 85–95.
- 950 Wang, X., Yin, Z.Y., Su, D., Wu, X., Zhao, J., 2022. A novel approach of ran-
951 dom packing generation of complex-shaped 3d particles with controllable
952 sizes and shapes. *Acta Geotechnica* 17, 355–376.
- 953 Wang, Y., Zhou, P., Zhong, W., 2018. An optimization strategy based on
954 hybrid algorithm of adam and sgd, in: MATEC Web of Conferences, EDP
955 Sciences. p. 03007.
- 956 Wei, D., Wang, J., Nie, J., Zhou, B., 2018. Generation of realistic sand par-
957 ticles with fractal nature using an improved spherical harmonic analysis.
958 *Computers and Geotechnics* 104, 1–12.
- 959 Wensrich, C., Katterfeld, A., 2012. Rolling friction as a technique for mod-
960 elling particle shape in dem. *Powder Technology* 217, 409–417.

- 961 Xiao, Y., Yuan, Z., Lin, J., Ran, J., Dai, B., Chu, J., Liu, H., 2019. Effect of
962 particle shape of glass beads on the strength and deformation of cemented
963 sands. *Acta Geotechnica* 14, 2123–2131.
- 964 Xiong, W., Wang, J., 2021. Gene mutation of particle morphology through
965 spherical harmonic-based principal component analysis. *Powder Technol-*
966 *ogy* 386, 176–192.
- 967 Yin, Z.Y., Wang, P., Zhang, F., 2020. Effect of particle shape on the pro-
968 gressive failure of shield tunnel face in granular soils by coupled fdm-dem
969 method. *Tunnelling and Underground Space Technology* 100, 103394.
- 970 Zamorski, M., Zięba, M., Klukowski, P., Nowak, R., Kurach, K., Stokowiec,
971 W., Trzciński, T., 2020. Adversarial autoencoders for compact representa-
972 tions of 3d point clouds. *Computer Vision and Image Understanding* 193,
973 102921.
- 974 Zhang, H., Zhang, X., Zhang, G., Dong, K., Deng, X., Gao, X., Yang, Y.,
975 Xiao, Y., Bai, X., Liang, K., et al., 2022a. Size, morphology, and com-
976 position of lunar samples returned by chang'e-5 mission. *Science China*
977 *Physics, Mechanics & Astronomy* 65, 1–8.
- 978 Zhang, P., Dong, Y., Galindo-Torres, S., Scheuermann, A., Li, L., 2021.
979 Metaball based discrete element method for general shaped particles with
980 round features. *Computational Mechanics* 67, 1243–1254.
- 981 Zhang, P., Galindo-Torres, S., Tang, H., Jin, G., Scheuermann, A., Li, L.,
982 2016. Lattice boltzmann simulations of settling behaviors of irregularly
983 shaped particles. *Physical Review E* 93, 062612.

- 984 Zhang, P., Qiu, L., Galindo-Torres, S., Chen, Y., Scheuermann, A., Li, L.,
985 2022b. Coupled metaball discrete element lattice boltzmann method for
986 fluid-particle systems with non-spherical particle shapes: A sharp interface
987 coupling scheme. arXiv preprint arXiv:2206.11634 .
- 988 Zhang, T., Zhang, C., Zou, J., Wang, B., Song, F., Yang, W., 2020. Dem
989 exploration of the effect of particle shape on particle breakage in granular
990 assemblies. Computers and Geotechnics 122, 103542.
- 991 Zhang, Z., 2018. Improved adam optimizer for deep neural networks, in: 2018
992 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS),
993 Ieee. pp. 1–2.
- 994 Zhao, S., Zhao, J., 2019. A poly-superellipsoid-based approach on particle
995 morphology for dem modeling of granular media. International Journal for
996 Numerical and Analytical Methods in Geomechanics 43, 2147–2169.
- 997 Zhao, S., Zhao, J., Guo, N., 2020. Universality of internal structure charac-
998 teristics in granular media under shear. Physical Review E 101, 012906.
- 999 Zhao, X., Elsworth, D., He, Y., Hu, W., Wang, T., 2021. A grain tex-
1000 ture model to investigate effects of grain shape and orientation on macro-
1001 mechanical behavior of crystalline rock. International Journal of Rock
1002 Mechanics and Mining Sciences 148, 104971.
- 1003 Zhao, Y., Zhang, P., Lei, L., Kong, L., Galindo-Torres, S., Li, S.Z., 2023.
1004 Metaball-imaging discrete element lattice boltzmann method for fluid-
1005 particle system of complex morphologies with case studies. Physics of
1006 Fluids 35, 023308.

- 1007 Zheng, J., Hryciw, R.D., 2015. Traditional soil particle sphericity, roundness
1008 and surface roughness by computational geometry. *Géotechnique* 65, 494–
1009 506. Publisher: Thomas Telford Ltd.
- 1010 Zheng, J., Hryciw, R.D., 2016. A corner preserving algorithm for realistic
1011 DEM soil particle generation. *Granular Matter* 18, 84. Publisher: Springer.
- 1012 Zhou, B., Wang, J., 2017. Generation of a realistic 3d sand assembly using
1013 x-ray micro-computed tomography and spherical harmonic-based principal
1014 component analysis. *International Journal for Numerical and Analytical
1015 Methods in Geomechanics* 41, 93–109.
- 1016 Zhou, B., Wang, J., Wang, H., 2018. Three-dimensional sphericity, roundness
1017 and fractal dimension of sand particles. *Géotechnique* 68, 18–30.
- 1018 Zhou, B., Wang, J., Zhao, B., 2015. Micromorphology characterization and
1019 reconstruction of sand particles using micro x-ray tomography and spher-
1020 ical harmonics. *Engineering geology* 184, 126–137.
- 1021 Zhou, B., Wei, D., Ku, Q., Wang, J., Zhang, A., 2020. Study on the effect
1022 of particle morphology on single particle breakage using a combined finite-
1023 discrete element method. *Computers and Geotechnics* 122, 103532.
- 1024 Zuo, L., Lourenco, S.D., Baudet, B.A., 2019. Experimental insight into the
1025 particle morphology changes associated with landslide movement. *Land-
1026 slides* 16, 787–798.