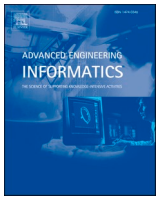
Advanced Engineering Informatics 69 (2026) 103994



Contents lists available at ScienceDirect

Advanced Engineering Informatics

journal homepage: www.elsevier.com/locate/aei



Augmentation of 3D virtual aggregate database using deep convolutional Wasserstein generative adversarial networks

Dong Feng [a], Jinchao Guan [b], Chaoliang Fu [a,*], Yuanyuan Hu [a], Frederic Otto [c], Alvaro García Hernandez [a], Pengfei Liu [a]

[a] Institute of Highway Engineering, RWTH Aachen University, Mies-van-der-Rohe-Straße 1, 52074 Aachen, Germany
[b] Zhejiang Expressway Co., Ltd., 310016 Hangzhou, China
[c] Federal Highway and Transport Research Institute (BASt), Brüderstr. 53, 51427 Bergisch Gladbach, Germany

ARTICLE INFO

Keywords:
3D Virtual aggregate
Database augmentation
Deep learning
Stereo vision
DC-WGAN

ABSTRACT

The generalization capability of asphalt mixture numerical models incorporating virtual aggregates is often constrained by the limited morphological diversity of the virtual aggregate database, which reduces the ability of the model to reliably predict the mechanical response of asphalt pavements. To address this, a novel framework based on deep convolutional Wasserstein generative adversarial networks (DC-WGANs) is proposed in this study to augment the 3D virtual aggregate database. The original 3D virtual aggregate database was created using stereo vision and multi-view matching techniques. Six DC-WGAN variants were subsequently employed for data augmentation, and a multi-criteria evaluation approach was introduced to comprehensively assess both model performance and the augmented databases. In addition, aggregate packing and Superpave Gyratory Compaction (SGC) simulations were performed to investigate the effect of data augmentation on the prediction of mechanical response. Finally, the feasibility of applying lightweight processing to the database augmented by the best-performing DC-WGAN variant was investigated. Results show that the DC-WGAN-div model achieved the best overall performance, as evidenced by its lowest Wasserstein Distance and Fréchet Inception Distance values and morphological fidelity in terms of true sphericity, angularity index, and fractal dimension. Additionally, the augmented aggregate database can reproduce the mechanical response of asphalt mixtures with high fidelity, as evidenced by the highly consistent aggregate packing density and gyration-height curve. Moreover, the lightweight database, although incurring minor quality loss, substantially improved training efficiency. These findings demonstrate the feasibility of incorporating artificial intelligence-driven data augmentation techniques into virtual aggregate modeling.

1. Introduction

Finite element modeling and discrete element modeling, which are based on the geometric characteristics of virtual aggregates, have been increasingly adopted in pavement engineering to conduct mesoscale simulations of asphalt mixtures [1–3]. These methods enable a more realistic analysis of the mechanical behavior of asphalt materials by accounting for their inherent heterogeneity and complex internal interactions. However, the generalizability and accuracy of such simulations are highly dependent on the quality and representativeness of the aggregate input data [4,5]. Therefore, establishing a comprehensive virtual aggregate database that accurately reflects the geometric and morphological characteristics of real aggregates is essential for supporting reliable and robust numerical modeling.

In response to this need, various techniques have been developed to create virtual aggregate databases. These methods can be broadly categorized into traditional image-based reconstruction approaches and emerging artificial intelligence (AI)-driven data augmentation techniques. On the one hand, image-based methods have long been used to capture aggregate morphology. Among them, X-ray computed tomography (CT) has been widely used due to its high resolution and accuracy. For example, Wang et al. [6] and Jin et al. [7] established virtual aggregate databases by combining CT scanning with computer vision algorithms, enabling precise morphological characterization. However, the high equipment cost and labor intensity of CT remain major barriers to widespread application. Alternatively, photogrammetry has been explored as a more economical solution. Ozturk and Rashidzade [8] as

* Corresponding author.
E-mail address: fu@isac.rwth-aachen.de (C. Fu).

https://doi.org/10.1016/j.aei.2025.103994
Received 12 June 2025; Received in revised form 26 August 2025; Accepted 15 October 2025
Available online 24 October 2025
1474-0346/© 2025 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Nomenclature

Abbreviation

AIMS	Aggregate Imaging System
AI	Artificial intelligence
BC	Bhattacharyya Coefficient
CT	Computed tomography
CT-WGAN	Wasserstein generative adversarial network with consistency regularization
D	Discriminator
DC	Deep convolution
DC-WGAN	Deep convolutional Wasserstein generative adversarial network
DEM	Discrete element method
FID	Fréchet Inception Distance
G	Generator
GAN	Generative adversarial network
Lp-WGAN	l_p-norm Wasserstein generative adversarial network
PD	Packing density
RWGAN	Relaxed Wasserstein generative adversarial network
SGC	Superpave Gyratory Compaction
WD	Wasserstein Distance
WGAN	Wasserstein generative adversarial network
WGAN-div	Wasserstein generative adversarial network based on Wasserstein divergence
WGAN-GP	Wasserstein generative adversarial network with gradient penalty

Symbols

α	Scaling level for Lp-WGAN
α_1, β_1	Rotation angle
β	Critical damping ratio
β_n	Normal critical damping ratio
β_t	Shear critical damping ratio
ε	Random number for WGAN-GP
δ	Side length used to cover the fractal object
$\overline{\delta}_n$	Normal overlap at the contact
$\overline{\delta}_t$	Tangential overlap at the contact
$\overline{\theta}^{rel}$	Relative rotation angle
λ	Hyperparameter for WGAN-GP
λ_1, λ_2	Hyperparameters for CT-WGAN
μ	Mean for random noise
μ_r	Rolling friction coefficient
μ_s	Static friction coefficient
μ_G	Mean of augmented data for feature vector/ Mean of the indicator values in augmented data
μ_R	Mean of real data for feature vector/ Mean of indicator values in real data
σ	Standard deviation for random noise
σ_G	Standard deviation of augmented data for feature vector/ Standard deviation of indicator values in augmented data
σ_R	Standard deviation of real data for feature vector/ Standard deviation of indicator values in real data
Σ_G	Covariance of augmented data for the feature vector
Σ_R	Covariance of real data for the feature vector
Φ	Strictly convex and continuously differentiable function
c	Clipping parameter for WGAN
e	Discrete value of angle
k	Hyperparameter for WGAN-div
k_t	Tangential stiffness coefficient
m	Total number of samples in a batch
m_{ij}	Equivalent mass of the particle
n	Hyperparameter for WGAN-div
$\vec{n}$	Outward unit normal vector
p	p-norm for Lp-WGAN
$p_G(x)$	Generated data distribution
$p_R(x)$	Real data distribution
r_{ij}	Contact effective radius
s	Scaling level for RWGAN
$\vec{u}_n^{rel}$	Relative normal translational velocity
$\vec{u}_t^{rel}$	Relative tangential translational velocity
x', x''	Perturbed data
z	Random noise
A	Fitting parameter
A_i	Area of the projected surface
AI	3D angularity index value
AI_i	2D angularity index value
B	Fitting parameter
BC_Φ	BC value
COR	Coefficient of restitution
$D(\cdot)$	Discriminator function
E_{ij}	Equivalent Young's modulus
$\vec{F}$	Vector field
$\vec{F}^c_{n,Hertz}$	Normal contact force for Hertz-Mindlin contact model
$\vec{F}^d_{n,Hertz}$	Normal damping force for Hertz-Mindlin contact model
$\vec{F}^c_{t,Hertz}$	Tangential contact force for Hertz-Mindlin contact model
$\vec{F}^d_{t,Hertz}$	Tangential damping force for Hertz-Mindlin contact model
FD	Fractal dimension value
FID	Multivariate Fréchet Inception Distance value
FID_Φ	Univariate Fréchet Inception Distance value
G	Gyration number of specimen
$G(\cdot)$	Generator function
$G(x)$	Discrete probability value at point x in the augmented data distribution
G_{ij}	Equivalent shear modulus
H	Height of specimen
H_0	Null hypothesis
L_D	Discriminator loss function value
L_G	Generator loss function value
$\vec{M}^r_{Hertz}$	Rolling friction moment for Hertz-Mindlin contact model
M'	Constant for CT-WGAN
$N(\delta)$	Number of counting boxes
PD	Packing density
$P(e)$	Angle change frequency of each vertex relative to the previous vertex
$R(x)$	Discrete probability value at point x in the real data distribution
R^2	Determination coefficient
S	Total surface area
Tr	Trace of the matrix
TS	True sphericity value
V	Enclosed volume
$V_{Aggregate}$	Total volume of aggregates
$V_{Container}$	Volume of filling container
W_D	Wasserstein Distance value
W_{div}	Wasserstein Distance value for WGAN-div
X_i, Y_i, Z_i	Point cloud coordinates
$\mathscr{D}$	Set of 1-Lipschitz functions
∇	Gradient
$\nabla\cdot$	Divergence

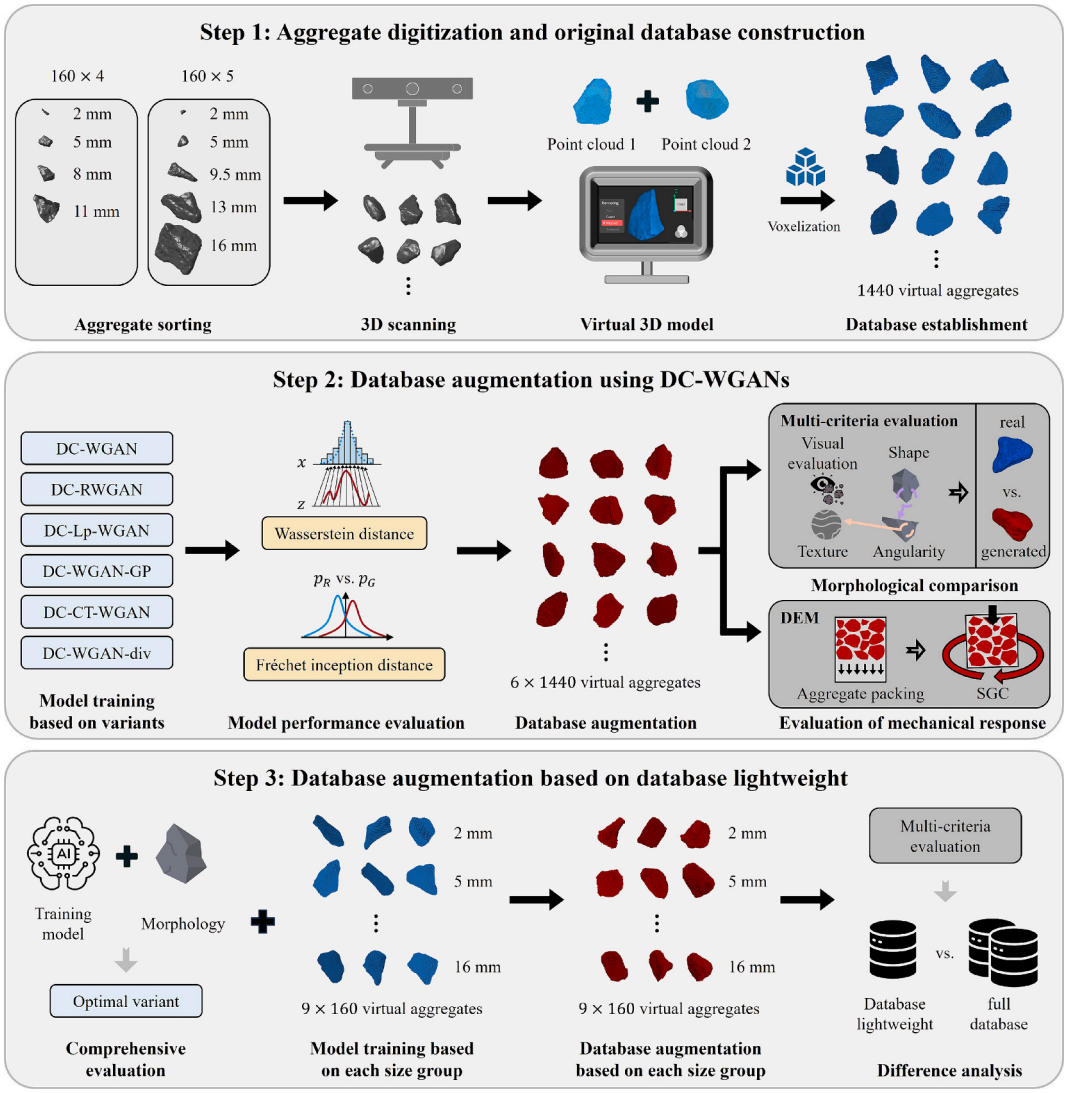


Fig. 1. Schematic for technical framework.

well as Zhao et al. [9] used multi-view image reconstruction to generate 3D virtual aggregates for random stacking simulations. While photogrammetry is more cost-effective, its data acquisition efficiency is relatively low. The Aggregate Imaging System (AIMS) is another commonly used tool for analyzing aggregate shape features. Wang et al. [10] and Rajan and Singh [11] applied AIMS to evaluate angularity and study the influence of crushing processes. Although AIMS is efficient and cost-effective, it is limited to 2D representations. This method inherently fails to capture the true volumetric morphology of aggregates, neglecting surface texture variations and spatial interlocking patterns that govern key mechanical responses in asphalt mixtures. Consequently, critical factors such as load transfer efficiency, contact force distribution, and particle rotation behavior cannot be fully represented in subsequent numerical simulations [12,13]. In comparison, 3D laser scanning methods, as utilized by Suhr et al. [14] and Liu et al. [15], can partially overcome the limitations associated with traditional 2D approaches by enabling the reconstruction and morphological characterization of 3D aggregate models. Nevertheless, despite the advances achieved through conventional image-based techniques, their inherent limitations in efficiently augmenting database size and morphological diversity persist, thereby constraining their ability to support comprehensive and robust numerical simulations.

On the other hand, with recent advances in AI, researchers have begun to explore deep learning as a tool to augment virtual aggregate databases. Wang et al. [16] proposed a deep learning-based framework for generating 2D aggregate images that maintain morphological features statistically similar to those of real aggregates. Their results demonstrated that the angularity distribution of generated aggregates closely matched that of real ones, confirming the feasibility of AI-driven augmentation. However, 2D representations are insufficient to capture the complex 3D geometry of real aggregates, as the mechanical behavior of asphalt mixtures is influenced not only by aggregate outlines but also by 3D surface features and spatial interlocking effects [17,18]. Recent studies have shown that deep learning techniques can be applied to generate voxel-based 3D models [19,20] or perform 3D reconstruction [21,22], largely employing variants of generative adversarial networks (GANs) [23]. These studies highlight a promising direction for leveraging deep generative models to construct 3D virtual aggregates for numerical simulations.

In summary, although traditional techniques have been widely used to construct virtual aggregate databases, they are inherently limited by the morphological characteristics of the measured samples. Consequently, it is difficult to capture the full range of aggregate shapes and sizes encountered in practice, which restricts the morphological diversity of the database and, in turn, constrains the generalization capability of asphalt mixture numerical models based on such data. Furthermore, although recent studies have introduced AI as a promising tool to augment aggregate morphological diversity in virtual databases,

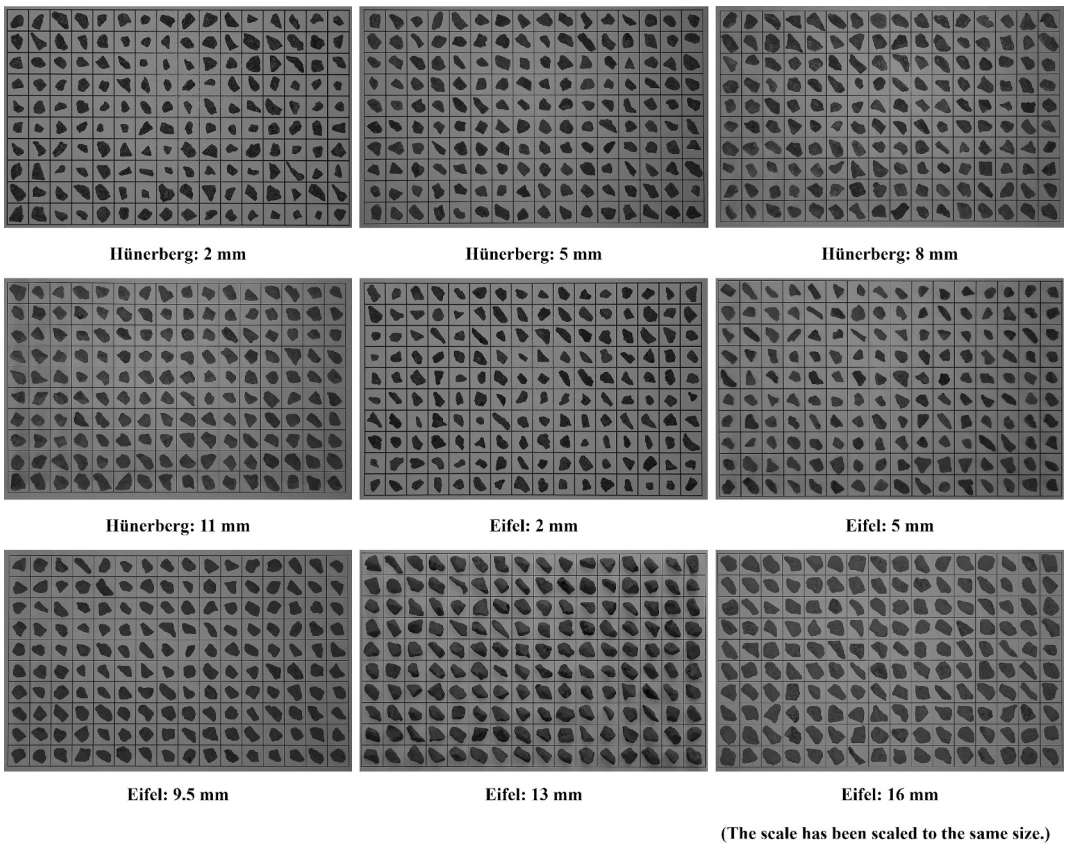


Fig. 2. Collection of original aggregates.

the research has been limited to 2D aggregate data, which fails to capture the complex 3D characteristics required for accurate numerical modeling.

To overcome these limitations, this study proposes a machine learning-based framework for augmenting 3D virtual aggregate databases. The original database is established using a 3D scanning system based on stereo vision and multi-view matching techniques. Building upon this, six deep convolutional Wasserstein generative adversarial networks (DC-WGANs) are employed to augment the original database. To comprehensively evaluate the effectiveness of the proposed augmentation framework, the generative performance of the six DC-WGAN variants is compared first using Wasserstein Distance (WD) and Fréchet Inception Distance (FID) to identify the most effective model. Subsequently, the augmented databases produced by each variant are evaluated using morphological indices such as true sphericity, angularity index, and fractal dimension. Based on these evaluations, the optimal DC-WGAN training model is identified. Additionally, aggregate packing and Superpave Gyratory Compaction (SGC) simulations are conducted to investigate the impact of augmented data on mechanical response prediction. Finally, the feasibility of applying lightweight processing to the database augmented by the best-performing DC-WGAN is investigated to assess the trade-off between efficiency and morphological fidelity.

2. Methodology

The technical framework of this study comprised three sequential steps: (1) aggregate digitization and original database construction, (2) database augmentation using DC-WGANs, and (3) database augmentation based on database lightweight. In this framework, the third step serves as an extension of the second, aiming to evaluate whether high-quality augmentation can be achieved using significantly fewer training samples, thereby improving computational efficiency without significantly compromising data fidelity. In addition, the sample data, codes, and trained models were provided in an open-source manner to ensure reproducibility, with specific information given in the Appendix A.

As illustrated in Fig. 1, the first step involved the construction of the original virtual aggregate database. Aggregates used in the asphalt mixture were sieved into predefined size ranges, and their surface geometries were captured using a stereo vision-based 3D scanner. The acquired data were processed through multi-view matching techniques to reconstruct point clouds, which were then converted into virtual 3D models and voxelized. This process yielded an original database containing 1440 aggregate voxel models.

In the second step, the original database was used to train six DC-WGAN variants, which were then employed to generate the augmented aggregate database. The generative performance of each model was first evaluated using indicators such as WD and FID. Subsequently, a morphological comparison between the augmented and real aggregates was conducted to assess geometric fidelity. This two-stage evaluation framework ensured both the statistical robustness and structural accuracy of the augmented aggregates. Finally, based on the discrete element method (DEM), aggregate packing and SGC simulations were conducted to investigate the impact of the augmented data on the prediction of mechanical responses.

In the third step, the optimal DC-WGAN generative model was selected based on the evaluation results from Step 2. To assess the feasibility of reducing the training data volume while maintaining acceptable quality of augmentation, the optimal DC-WGAN model was retrained using only 160 voxelized samples for each particle size group. A comparative analysis was then conducted between the optimal DC-WGAN models trained on the reduced and full databases, focusing on generative performance and morphological characteristics of the resulting aggregates. The detailed methodological procedure is presented below.

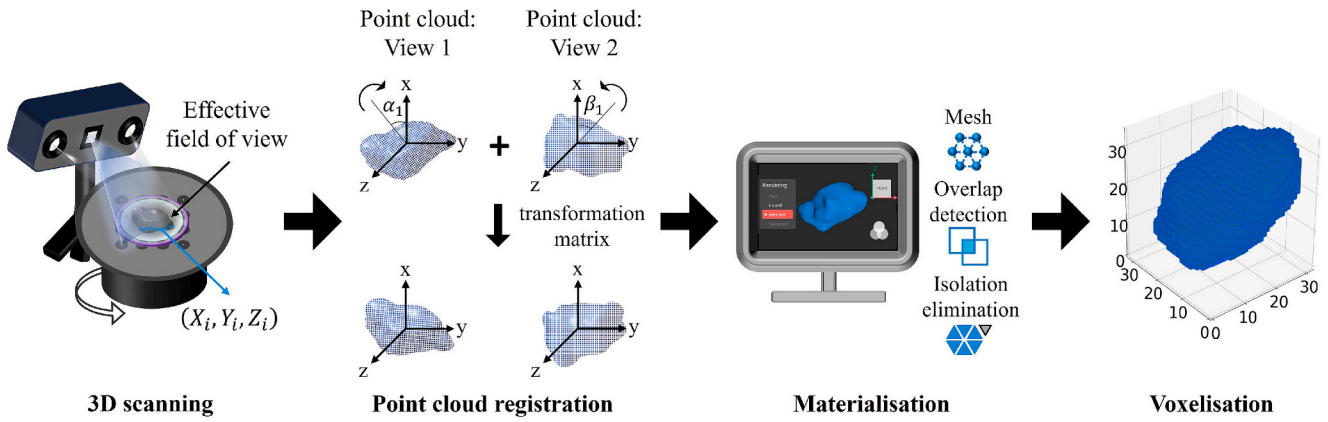


Fig. 3. Principle of 3D virtual aggregate reconstruction.

2.1. Establishment of original aggregate database

The original aggregates used in this study were basalt sourced from Hünerberg and Eifel in Germany, as shown in Fig. 2. Aggregates from Hünerberg were screened into size fractions of 2 mm, 5 mm, 8 mm, and 11 mm, corresponding to the coarse aggregate gradation of an AC-11 asphalt mixture. Meanwhile, aggregates from Eifel were screened and sorted into 2 mm, 5 mm, 9.5 mm, 13 mm, and 16 mm, making them appropriate for AC-13 or AC-16 coarse aggregate gradations. Each grain size contained 160 aggregates, and a total of 1440 aggregates were scanned to construct the original aggregate database. The 3D scanner used in this study was a high-precision binocular stereo scanner (Revopoint MINI 3D Scanner), allowing a minimum point spacing of 0.05 mm and a single-frame accuracy of 0.02 mm. The former indicates the scanner's spatial resolution, while the latter reflects its measurement accuracy. The scanner operated with an acquisition frequency of 10 frames per second, and the effective working distance was 120–130 mm, providing sufficient coverage for aggregates of all sizes.

To construct a high-fidelity 3D original aggregate database, a standardized workflow was developed, as shown in Fig. 3. First, individual aggregates were placed on a motorized turntable rotating at 2 revolutions per minute (rpm). A binocular stereo vision system equipped with a blue-light projector was used to capture the surface geometry of each aggregate. Spatial localization markers were arranged around the scanning area to facilitate precise calibration. As the turntable rotated, stereo cameras continuously captured point cloud data (X_i, Y_i, Z_i) from multiple viewpoints. The effective scanning region was defined by the overlapping field of view between the two cameras, ensuring sufficient coverage and high geometric accuracy for 3D reconstruction. Since the bottom surface could not be fully recorded in a single view, the aggregate was flipped and rescanned to ensure complete surface acquisition. Ideally, a full 360° rotation of the turntable would yield a complete point cloud, but due to visual shadows caused by shape variations, the turntable usually required a slight additional rotation. Each aggregate was scanned at least twice from multiple viewpoints to ensure a complete and detailed point cloud.

Next, the raw point clouds captured from multiple views were processed to reconstruct a continuous and complete 3D model of each aggregate. Preprocessing included removing outliers, applying a statistical noise filter, and downsampling with a voxel grid filter (voxel size: 0.05 mm) to balance resolution and computational efficiency. Point cloud registration was then performed to align multiple views, with the core step being the determination of a transformation matrix to align one point cloud (rotation angle α_1) with another (rotation angle β_1), thereby maximizing surface feature overlap [24,25]. This step was implemented using MATLAB's built-in function pcmatchfeatures(), which enabled efficient batch processing of point cloud registration tasks. After point cloud registration, data from all viewing angles were merged and converted into triangular mesh models to accurately represent aggregate geometry. The resulting mesh models were saved in. stl format to ensure compatibility with mainstream simulation platforms. Overlap detection was applied to remove duplicate points, which occur when points appear more than once due to multi-view matching. Duplicate coordinates were identified and removed from the point cloud matrix to prevent errors in mesh construction. In addition, the mesh was further refined by eliminating isolated points and filtering out surface noise introduced by dust or background interference. This was achieved by identifying and removing discrete noise points that deviated from the local spatial distribution of their neighboring points.

Finally, the cleaned mesh models were voxelized into a standardized 3D binary tensor with dimensions of $32 \times 32 \times 32$ voxels, where voxels occupied by the aggregate were assigned a value of 1, and empty voxels were assigned a value of 0. This voxel-based format facilitates efficient storage, consistent data structure, and broad applicability in numerical simulation, machine learning, and morphological analysis. Through this integrated process, a comprehensive and high-resolution digital original aggregate database was constructed, providing essential data for generating augmented aggregate databases.

2.2. Augmentation of original aggregate database

2.2.1. Construction of DC-WGAN model

Since the original GAN was proposed [23], GAN-based models have demonstrated strong capabilities in generating data. The classical GAN framework consists of a generator (G) and a discriminator (D), where G aims to produce data that resembles the real data distribution, while D is trained to distinguish between real and augmented data. However, two major limitations frequently hinder the training of standard GANs: (1) D often outperforms G, leading to vanishing gradients for G, which results in unstable training; and (2) mode collapse, where G learns only a limited subset of the data distribution, reducing the diversity of the augmented samples [26].

To address these issues, the WGAN [27] was introduced by replacing the Jensen-Shannon divergence with the WD as its objective function. The objective function of WGAN is defined as follows [28]:

$$\begin{aligned}\min_{G}\max_{D\in\mathscr{D}} V(D,G) &= E_{x\ p_R(x)}[D(x)] - E_{z\ p_Z(z)}[D(G(z))] \\ &= E_{x\ p_R(x)}[D(x)] - E_{x\ p_G(x)}[D(x)] \end{aligned} \quad (1)$$

where $\mathscr{D}$ is the set of 1-Lipschitz functions, $p_R(x)$ is the real data distribution, $p_G(x)$ is the augmented data distribution, and z is the random noise. The Lipschitz constraint restricts the gradient of the discriminator function $D(\cdot)$, ensuring training stability and smoothness. Specifically, the "1-Lipschitz" condition requires that the gradient magnitude of $D(\cdot)$ does not exceed 1, thus promoting continuity of transformation for D. In this study, to maintain consistency of the value function, the

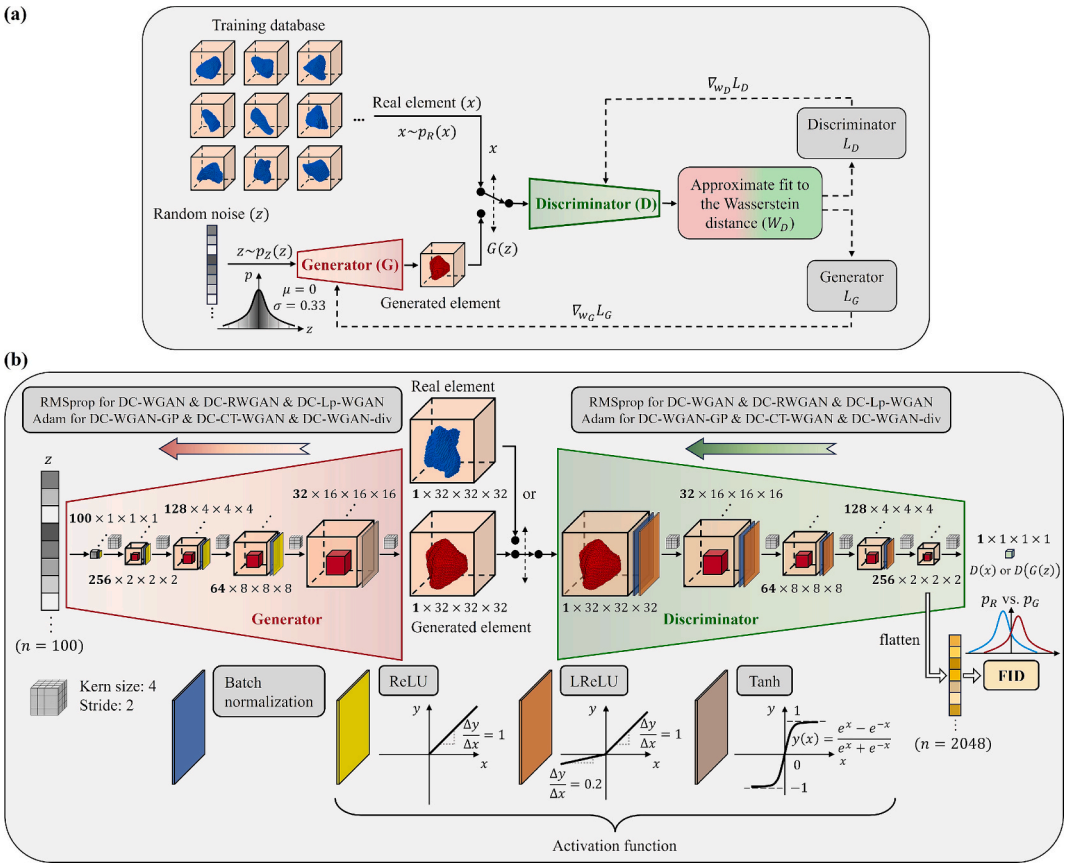


Fig. 4. Schematic for model of DC-WGANs: (a) training process, (b) composition of structure.

maximization problem was converted to a minimization problem. Accordingly, for the value function of D, $\max_D V(D,G)$ is equivalent to $-\min_D V(D,G)$.

Building on this theoretical foundation, six WGAN variants were considered to investigate how generator performance affects the morphology of augmented aggregates. For these WGANs, the loss function of G is identical across variants and is uniformly computed as follows:

$$L_{G_WGANs} = -\frac{1}{m}\sum_{i=1}^{m} D\left(G\left(z^{(i)}\right)\right) \tag{2}$$

where m is the total number of samples in a batch.

The key to using WGANs is to ensure that $D(\cdot)$ satisfies the Lipschitz continuity condition. The six WGAN variants differ primarily in how they enforce the Lipschitz continuity in $D(\cdot)$, and each modification is designed to improve training stability, convergence speed, or the quality of augmented data. The original WGAN [27] used a simple weight clipping method (see Eq. (3)). This method is straightforward but may limit the capacity of D and lead to slow convergence. Relaxed WGAN (RWGAN) [29] introduced a gradient-scaled clipping strategy based on a strictly convex function (see Eq. (4)). Compared to WGAN, RWGAN permits more flexible weight updates, thereby improving training stability. The l_p-norm WGAN (Lp-WGAN) [30] enforced Lipschitz continuity by normalizing the weight vector using the p-norm (see Eq. (5)). This method standardizes weight scales, reduces sensitivity to parameter initialization, and offers better convergence behavior than simple clipping.

$$L_{D_WGAN} = \frac{1}{m}\sum_{i=1}^{m}\left[-D\left(x^{(i)}\right) + D\left(G\left(z^{(i)}\right)\right)\right] \text{with} w \leftarrow clip(w, -c, c) \tag{3}$$

where $clip(w,-c,c)$ is a clipping operation applied to all parameter weights to ensure they remain within a bounded range, and c is the clipping parameter, which is a constant.

$$L_{D_RWGAN} = \frac{1}{m}\sum_{i=1}^{m}\left[-D\left(x^{(i)}\right) + D\left(G\left(z^{(i)}\right)\right)\right] \text{with} w \leftarrow clip\left(w, -s(\nabla\Phi)^{-1}c, s(\nabla\Phi)^{-1}c\right) \tag{4}$$

where s denotes the scaling level, and Φ denotes a strictly convex and continuously differentiable function. In this study, the Euclidean form of Φ was used (see [29] for details).

$$L_{D_Lp-WGAN} = \frac{1}{m}\sum_{i=1}^{m}\left[-D\left(x^{(i)}\right) + D\left(G\left(z^{(i)}\right)\right)\right] \text{with} w \leftarrow \alpha\frac{w}{\|w\|_p} \tag{5}$$

where α denotes the scaling level, and p denotes the p-norm. For example, $p = 2$ corresponds to the Euclidean norm.

The above three variants ensure Lipschitz continuity through weight clipping; however, this approach may restrict the smoothness of D. To address this, Lipschitz continuity can also be enforced using a gradient penalty [31]. This approach is called WGAN with gradient penalty (WGAN-GP). This is achieved by adding a gradient penalty term to the loss function of D:

$$L_{D_WGAN-GP} = \frac{1}{m}\sum_{i=1}^{m}\left[-D\left(x^{(i)}\right) + D\left(G\left(z^{(i)}\right)\right) + \lambda\left(\|\nabla_{\hat{x}} D\left(\hat{x}^{(i)}\right)\|_2 - 1\right)^2\right] \text{with} \hat{x}^{(i)} = \varepsilon x^{(i)} + (1-\varepsilon)G\left(z^{(i)}\right) \tag{6}$$

where ε is a random number from the uniform distribution $U[0,1]$. It ensures that the gradient penalty is applied across both the real data distribution $p_R(x)$ and the augmented data distribution $p_G(x)$. λ is a hyperparameter, which is a constant. Compared to weight clipping methods, WGAN-GP provides smoother gradients and better augmen-

tation quality, particularly in regions between real and augmented samples.

However, if the real data distribution differs significantly from the augmented data distribution, the gradient penalty term in WGAN-GP cannot ensure the continuity of the sampled points between the two distributions. To resolve this, a consistency regularization term was proposed in CT-WGAN [32], which added local perturbation constraints to ensure smooth transitions near real samples:

$$L_{D_CT-WGAN} = \frac{1}{m}\sum_{i=1}^{m}\left[-D\left(x^{(i)}\right)+D\left(G\left(z^{(i)}\right)\right)+\lambda_1 GP|_{\hat{x}}+\lambda_2 CT|_{x',x''}\right]$$

$$\text{with} GP|_{\hat{x}} = \left(\|\nabla_{\hat{x}} D\left(\hat{x}^{(i)}\right)\|_2 - 1\right)^2 \&$$

$$CT|_{x',x''} = max\left(0, \frac{\|D\left(x'^{(i)}\right) - D\left(x''^{(i)}\right)\|_2}{\|x'^{(i)} - x''^{(i)}\|_2} - M'\right) \quad (7)$$

where $\hat{x}^{(i)}$ is seen in Eq. (6), x' and x'' denote perturbed data in the vicinity of the real sample, M' is a constant used to limit the value of the consistency term, λ_1 and λ_2 are hyperparameters. Compared to WGAN-GP, CT-WGAN explicitly enforces local smoothness near real samples, thereby further stabilizing training in sparse data regions.

Furthermore, Wu et al. [33] proposed a new approach to optimize WGANs based on the Wasserstein divergence (WGAN-div), which can completely eliminate the Lipschitz continuity limitation. This is achieved by modifying the original WGAN objective function, as shown in Eq. (1). The loss function of D is expressed as follows:

$$L_{D_WGAN-div} = \frac{1}{m}\sum_{i=1}^{m}\left[-D\left(x^{(i)}\right)+D\left(G\left(z^{(i)}\right)\right)+k\|\nabla_{\hat{x}} D\left(\hat{x}^{(i)}\right)\|_2^n\right] \quad (8)$$

where $\hat{x}^{(i)}$ is also defined in Eq. (6), and k and n are hyperparameters. Compared to the previous variants, WGAN-div relaxes the strict Lipschitz constraint, thereby improving training efficiency and enhancing the diversity of the augmented samples.

While WGAN and its variants offer significant advantages in stabilizing adversarial training and improving the quality of augmented samples, they do not prescribe a specific network architecture for G and D. In practice, the architecture design plays a crucial role, especially when dealing with high-dimensional data such as 3D voxel representations. To address this, the deep convolutional structure proposed in DCGAN [34] was integrated with WGAN to form the DC-WGAN model. The convolutional layers in DCGAN are particularly effective in learning spatial hierarchies and local features, which are essential for capturing the geometric and morphological characteristics of 3D aggregates. Compared to alternative GAN architectures such as original GAN [23], conditional GAN (cGAN) [35], or StyleGAN [36], DC-WGANs provide a balance of training stability, gradient smoothness, and efficient feature extraction in three-dimensional space. Original GANs often suffer from mode collapse; cGANs require explicit conditional information that may not be readily available for 3D aggregates; and StyleGAN may produce pseudo-structures when applied directly to 3D volumetric data. This makes DC-WGAN particularly well-suited for generating realistic and diverse 3D aggregate structures. Based on this insight, six different variants of WGAN were constructed using a shared deep convolutional architecture, and their generator performance was systematically compared in the context of 3D aggregate morphology augmentation.

2.2.2. Training of DC-WGAN model

As illustrated in Fig. 4 (a), the training process of the DC-WGAN followed an iterative adversarial learning procedure. To generate 3D voxelized representations of aggregate particles, random noise sampled from a normal distribution ($\mu = 0, \sigma = 0.33$) with an initial size of $1 \times 1 \times 1$ was first fed into G. The mean $\mu = 0$ centers the latent space around zero, which helps prevent bias in initial outputs of G. The standard deviation $\sigma = 0.33$ was chosen to provide sufficient variation for G to explore diverse aggregate shapes without causing instability in training. Then, G upsampled this input through five layers of 3D transposed convolution, outputting a voxel grid of size $32 \times 32 \times 32$. Subsequently, D received $32 \times 32 \times 32$ voxel grids as input and progressively downsampled them to output $D(x)$ or $D(G(z))$ for loss. These losses were subsequently backpropagated to update the parameters of G and D alternately, thereby completing one training epoch.

Table 1
Hyperparameters for DC-WGANs.

DC-WGAN variants	Hyperparameters
DC-WGAN	$c = 0.04$
DC-RWGAN	$s = 8, c = 0.04$
DC-Lp-WGAN	$\alpha = 10, p = 2$
DC-WGAN-GP	$\lambda = 5$
DC-CT-WGAN	$\lambda_1 = 5, \lambda_2 = 5, M' = 0$
DC-WGAN-div	$k = 2, n = 6$

The training epochs were set to 200, while the batch sizes were set to 20. For both D and G, the optimizer used a uniform learning rate of 1e-3. After every five training cycles of D, G underwent one training cycle. As all training was performed within the framework of WGANs, there was no risk of vanishing gradients for G caused by D achieving too high accuracy. Since only a well-performing D can improve the performance of G, D is trained more frequently to prioritize its performance. Based on the references [27,29–33] and numerical tests, different hyperparameters were selected to ensure the convergence of the results. Table 1 shows the different hyperparameters used in this study for numerical experiments.

In addition, the structure of G and D is shown in Fig. 4 (b). The number of channels in each layer of G was set to 256, 128, 64, 32, and 1, respectively, with ReLU used in hidden layers and Tanh in the output layer. D had a symmetric structure consisting of five layers of 3D convolution. The number of channels in the discriminator layers was set in reverse order (1, 32, 64, 128, 256), and LeakyReLU was used as the activation function. Both networks used $4 \times 4 \times 4$ convolution kernels with a stride of 2, and batch normalization [37] was applied in the first four layers of both G and D to stabilize training and accelerate convergence. Moreover, optimization strategies were selected based on the characteristics of each model's loss function. Specifically, RMSprop [38] was used for DC-WGAN, DC-RWGAN, and DC-Lp-WGAN due to their simpler gradient constraints, while Adam [39] was employed for DC-WGAN-GP, DC-CT-WGAN, and DC-WGAN-div to better accommodate more complex regularization terms. At the final stage of the network, a flatten operation converted the output features into a 2048-dimensional vector, which was used to calculate the FID. This metric quantitatively evaluates the similarity between the distributions of augmented samples and real samples, providing a reliable basis for assessing the model's generative performance and diversity.

2.2.3. Augmentation and post-processing of original database

To enrich the original 3D aggregate database with high-quality and morphologically diverse augmented samples, this study established a comprehensive augmentation-to-application pipeline based on trained DC-WGAN models. This process not only extended the database volume but also introduced greater geometric variability, thereby improving the robustness and generalization of downstream numerical simulations.

Once the DC-WGAN model was fully trained, numerous augmented aggregate samples were generated by feeding randomly sampled latent vectors into G. These outputs were voxel-based representations with a resolution of $32 \times 32 \times 32$, exported in.mat format for further processing, as shown in Fig. 5 (a). However, the raw voxel data may contain structural artifacts, such as isolated noise voxels and disconnected clusters, which can compromise the geometric integrity of the aggregates. To mitigate these issues, a 3D median filtering process was applied in MATLAB to enhance structural continuity and eliminate non-

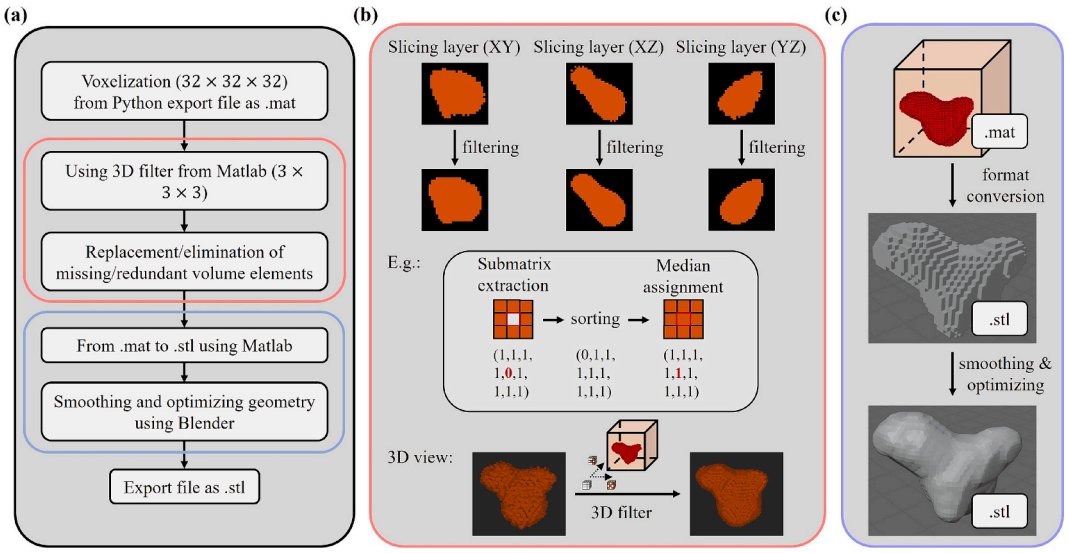


Fig. 5. Post-processing of the augmented data: (a) overview of the process, (b) processing of raw data using 3D filter, (c) processing and format transposition for voxel file.

connected components. As illustrated in Fig. 5 (b), filtering was performed across three orthogonal planes (XY, XZ, YZ), using a 3 × 3 × 3 sliding window. At each step, a local submatrix was extracted, its voxel values were sorted, and the median was assigned to the central voxel. This process effectively smoothed the voxel structure and improved local coherence.

Following voxel refinement, the cleaned.mat files were converted into.stl format to ensure compatibility with engineering simulation and visualization platforms, as shown in Fig. 5 (c). However, due to the discrete nature of voxel representations, the initial.stl models exhibit blocky and unrealistic surfaces. To address this, further mesh optimization was carried out in Blender, including surface smoothing, triangular mesh subdivision, and curvature-preserving reconstruction. These operations were automated through scripting to support efficient batch

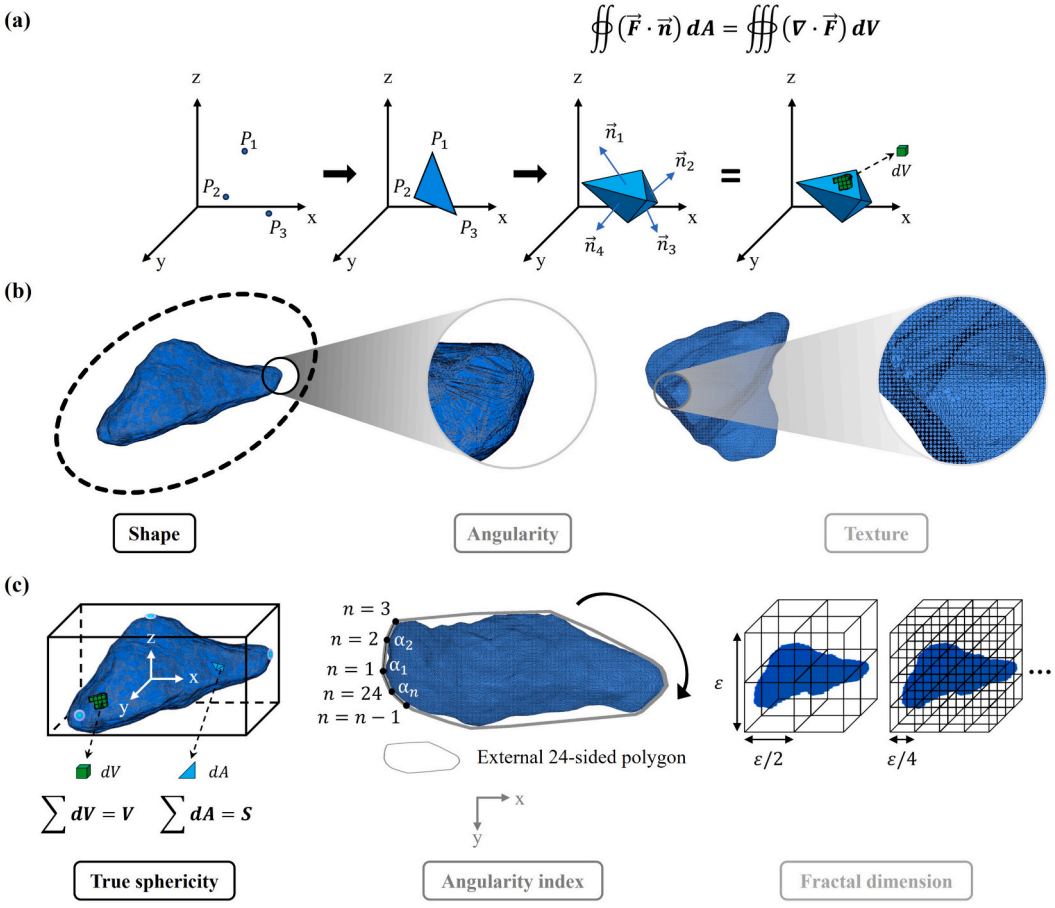


Fig. 6. Morphological investigation of aggregates: (a) 3D mesh reconstruction, (b) multi-scale morphology of aggregate, (c) morphological evaluation indicators.

processing and ensure consistency across all samples.

This comprehensive augmentation and post-processing pipeline ensured the geometric realism and structural fidelity of generated aggregates, thereby laying a solid foundation for their integration into numerical modeling.

2.2.4. A multi-criteria framework for evaluating augmented database

To assess the quality and reliability of the augmented aggregate database, a multi-criteria evaluation framework was proposed, incorporating two categories of indicators. The first category included training model indicators such as the WD and the FID, which evaluated the performance of the generative model. The second category was based on morphological indicators of aggregates, including true sphericity, angularity index, and fractal dimension, which were derived from 3D mesh models and described the shape features of individual aggregates at different scales. In this study, no weighting scheme was applied; all indicators were treated equally. Each indicator quantifies a specific aspect of database quality, without implying a hierarchy of importance. The use of weights could introduce subjective bias and compromise the objectivity of the overall evaluation. This comprehensive framework enabled a robust and systematic evaluation of the augmented database.

2.2.4.1. Evaluation indicators for DC-WGANs model performance. The WD measures the distance between two probability distributions and can be interpreted as the minimal cost required to transform one distribution into the other [26]. A lower WD indicates a closer match between the real and augmented data distributions. When the discriminator $D(\cdot)$ satisfies the 1-Lipschitz continuity condition (excluding DC-WGAN-div), the discrete form of WD is:

$$W_D(p_R, p_G) = \frac{1}{m}\sum_{i=1}^{m}\left[D\left(x^{(i)}\right) - D\left(G\left(z^{(i)}\right)\right)\right] \tag{9}$$

For DC-WGAN-div, the WD is calculated as follows:

$$W_{div}(p_R, p_G) = \frac{1}{m}\sum_{i=1}^{m}\left[D\left(x^{(i)}\right) - D\left(G\left(z^{(i)}\right)\right) + k\|\nabla_{\hat{x}} D\left(\hat{x}^{(i)}\right)\|_2^n\right] \tag{10}$$

The FID is widely used to evaluate the quality and diversity of data augmented by GANs [40–42]. It compares the means and covariances of features extracted from real and augmented data distributions. Lower FID values indicate higher similarity between the two distributions. In this study, FID was calculated using a 2048-dimensional feature vector extracted by flattening the final layer of D. For the multivariate normal distribution, the FID can be computed as follows:

$$FID = \|\mu_G - \mu_R\|_2^2 + Tr\left(\Sigma_G + \Sigma_R - 2(\Sigma_G\Sigma_R)^{1/2}\right) \tag{11}$$

where μ_G, Σ_G and μ_R, Σ_R are the means and covariances of the augmented and real data for the feature vectors, respectively; Tr is the trace of the matrix.

2.2.4.2. Evaluation indicators for augmented aggregate database. To evaluate the geometric fidelity of the generated aggregates in the augmented aggregate database, this study employed three scale-sensitive morphological indicators, including shape on the coarse scale, angularity on the medium scale, and texture on the fine scale. These morphology-related indicators rely on a valid 3D surface representation of each aggregate, which is obtained via mesh reconstruction from post-processed data. The mesh reconstruction process is essential, as it enables the transition from discrete data to a geometrically valid and analyzable 3D model.

As illustrated in Fig. 6 (a), the mesh reconstruction process involves connecting adjacent points to form triangular facets, thus producing a closed 3D surface model. The total surface area S is obtained by summing all triangular elements, and the enclosed volume V is calculated using the divergence theorem:

$$\oiint\left(\vec{F}\cdot\vec{n}\right)dA = \iiint\left(\nabla\cdot\vec{F}\right)dV \tag{12}$$

where $\vec{F}$, $\vec{n}$, $\nabla\cdot$ denote the vector field, the outward unit normal vector, and the divergence, respectively.

Based on the reconstructed 3D models, the morphological properties of the augmented aggregates can be quantitatively characterized at multiple scales. As shown in Fig. 6 (b), shape at the coarse scale, angularity at the medium scale, and texture at the fine scale were used to characterize aggregate morphology. The three scales, from largest to smallest, represent the overall contour, the sharpness of edges, and the microscopic surface roughness of the aggregates. As shown in Fig. 6 (c), specific morphological evaluation indicators for three scales were demonstrated, including true sphericity for shape, angularity index for angularity, and fractal dimension for texture. True sphericity reflects how closely the shape of an aggregate resembles a perfect sphere. It is defined as [43]:

$$TS = \frac{\pi}{S}\left(\frac{6V}{\pi}\right)^{\frac{2}{3}} \tag{13}$$

A higher true sphericity value indicates a smoother and more regular shape, closer to an ideal sphere. Moreover, the angularity index [44] was used to quantify the angularity. A higher angularity index value indicates a greater angularity of the aggregate. For each aggregate, the angularity index was first calculated in 2D based on the outline of a 24-sided polygon from each of the three orthogonal views (front, side, top), and then averaged using area weighting:

$$AI_i = \sum_{e=0}^{170} eP(e)$$

$$AI = \frac{\sum_{i=1}^{3} AI_i A_i}{\sum_{i=1}^{3} A_i} \tag{14}$$

where $e = 0^\circ, 10^\circ, 20^\circ, \cdots, 170^\circ$, $P(e)$ denotes the frequency of angular changes at each vertex relative to the previous vertex within the range from e to $(e + 10^\circ)$, and A_i denotes the area of the projected surface of any view.

Meanwhile, the fractal dimension characterizes the surface roughness and complexity at a fine scale. A larger fractal dimension indicates a more irregular and rough surface. The box-counting method was used in this study [45]:

$$FD = \lim_{\delta\to 0}[logN(\delta)/log(1/\delta)] \tag{15}$$

where $N(\delta)$ denotes the number of counting boxes with side length δ required to cover the fractal object. In this study, fractal dimension was computed based on $32 \times 32 \times 32$ voxel models and averaged across four scales (lengths $\delta/2$, $\delta/4$, $\delta/8$, and $\delta/16$) to reduce variability.

Finally, to assess the similarity between augmented and real aggregates, the above indicators served as the basis for statistical comparison. This study employed two feature-level similarity indicators: the univariate Fréchet Inception Distance (FID_Φ) and the Bhattacharyya Coefficient (BC_Φ). For each indicator $\Phi \in \{TS, AI, FD\}$, FID_Φ is calculated as follows:

$$FID_\Phi(G, R) = (\mu_G - \mu_R)^2 + (\sigma_G - \sigma_R)^2 \tag{16}$$

where μ_G, σ_G and μ_R, σ_R are the means and standard deviations of the indicator values in the augmented and real data.

Two different FIDs were used in this study. Therefore, it is necessary to emphasize their difference. The FID in Eq. (11) compared the similarity between the augmented data and the real data in a high dimension, while the FID_Φ directly compared the similarity between the morphological features of the augmented aggregates and the real aggregates based on the one-dimensional data distribution.

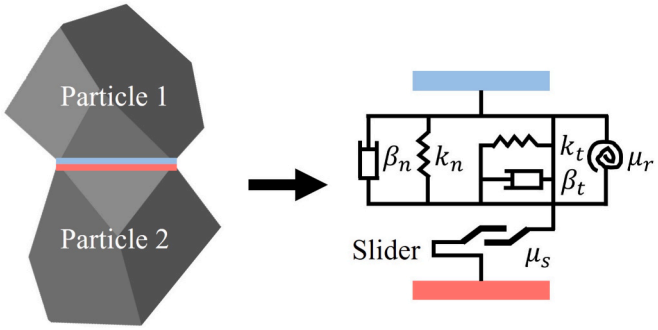


Fig. 7. Contact model for Hertz-Mindlin (with slip) model.

In addition, the Bhattacharyya Coefficient BC_Φ [46] was used to quantify the overlap between the distributions of real and augmented features:

$$BC_\Phi(G,R) = \sum_x \sqrt{G(x)R(x)} \tag{17}$$

where $G(x)$ and $R(x)$ are the discrete probability values at point x for the augmented and real data distributions. The value of BC_Φ ranges from 0 to 1, where higher values indicate greater similarity between the distributions.

2.2.5. Evaluation of mechanical response based on DEM

To further verify the practical significance of the augmented aggregate database, a comparative discrete element method (DEM) simulation was conducted. While the previous sections primarily focused on evaluating morphological fidelity, it is also essential to examine whether the augmented data can enhance the accuracy and robustness of mechanical response predictions in asphalt mixtures. By incorporating both the real and augmented databases into DEM simulations, the influence of data augmentation on mixture-level behavior can be systematically assessed. The DEM, originally developed by Cundall and Strack [47,48], simulates granular media by representing them as collections of discrete particles. Particle motions are determined by Newtonian dynamics, while particle–particle and particle–boundary interactions are computed through contact models, enabling accurate representation of packing structure and load transfer mechanisms. Considering the incompressibility of aggregates, this study adopted the Hertz-Mindlin contact model to simulate the contact behavior between aggregates (see Fig. 7). The Hertz-Mindlin contact model consists of a nonlinear formulation based on an approximation of the theory of Mindlin and Deresiewicz [49]. However, the original Hertz-Mindlin contact model does not include a slip component, making it incapable of realistically simulating the rolling characteristics of particles. Therefore, the Hertz-Mindlin contact model used in this study additionally contained a slip component to consider particle rolling. Table 2 summarizes the force and moment formulas for the Hertz-Mindlin contact model. Table 3 describes the meaning of each parameter.

In this study, aggregate packing and Superpave Gyratory Compaction (SGC) simulations were performed to create standard virtual Marshall specimens with a diameter of 101.6 mm and a height of 63.5 mm. All simulations used the AC-13 gradation, with the selected gradation relationships shown in Fig. 8. Considering computational efficiency, only aggregates with particle sizes larger than 2.36 mm were included in the DEM simulations. The parameters used for the

Table 3
Description of parameters for Hertz-Mindlin (with slip) contact model.

Parameter	COR	E_{ij}	G_{ij}	r_{ij}
Description	Coefficient of restitution	Equivalent Young's modulus	Equivalent shear modulus	Contact effective radius
Parameter	β_n	β_t	$\overline{u}_n^{rel}$	$\overline{u}_t^{rel}$
Description	Normal critical damping ratio	Shear critical damping ratio	Relative normal translational velocity	Relative tangential translational velocity
Parameter	$\overline{\theta}^{rel}$	m_{ij}	$\overline{\delta}_n$	$\overline{\delta}_t$
Description	Relative rotation angle	Equivalent mass of the particle	Normal overlap at the contact	Tangential overlap at the contact
Parameter	μ_s	μ_r	k_t	
Description	Static friction coefficient	Rolling friction coefficient	Tangential stiffness coefficient	

Table 2
Force and moment formulas for Hertz-Mindlin (with slip) contact model.

Normal force	Contact	$\overline{F}^c_{n,Hertz} = \frac{4}{3}E_{ij}\sqrt{r_{ij}}\overline{\delta}_n^{\frac{3}{2}}$ (18)
	Damping	$\overline{F}^d_{n,Hertz} = -2\beta_n\overline{u}_n^{rel}\sqrt{\overline{S}_n m_{ij}}$ (19)
Tangential force	Contact	$\overline{F}^c_{t,Hertz} = \begin{cases} \overline{S}_t\overline{\delta}_t, \lVert\overline{S}_t\overline{\delta}_t\rVert \le \lvert\mu_s\overline{F}^c_{n,Hertz}\rvert \\ \mu_s\overline{F}^c_{n,Hertz}, otherwise \end{cases}$ (20)
	Damping	$\overline{F}^d_{t,Hertz} = -2\beta_t\overline{u}_t^{rel}\sqrt{\overline{S}_t m_{ij}}$ (21)
Moment	Rolling friction	$\overline{M}^r_{Hertz} = \begin{cases} -\overline{S}_t r_{ij}^2\overline{\theta}^{rel}, \lVert\overline{M}^r_{Hertz}\rVert \le \lvert\mu_r r_{ij}\overline{F}^c_{n,Hertz}\rvert \\ \mu_r r_{ij}\overline{F}^c_{n,Hertz}, otherwise \end{cases}$ (22)
with		$\beta = \frac{-ln(COR)}{\sqrt{\pi^2 + (ln(COR))^2}}$ (23)
		$E_{ij} = \left(\frac{(1-\nu_i^2)}{E_i} + \frac{(1-\nu_j^2)}{E_j}\right)^{-1}$ (24)
		$r_{ij} = \left(\frac{1}{r_i} + \frac{1}{r_j}\right)^{-1}$ (25)
		$m_{ij} = \left(\frac{1}{m_i} + \frac{1}{m_j}\right)^{-1}$ (26)
		$\overline{S}_n = 2E_{ij}\sqrt{r_{ij}\overline{\delta}_n}$ (27)
		$\overline{S}_t = 8G_{ij}\sqrt{r_{ij}\overline{\delta}_t}$ (28)
		$G_{ij} = \left(\frac{2-\nu_i}{G_i} + \frac{2-\nu_j}{G_j}\right)^{-1}$ (29)

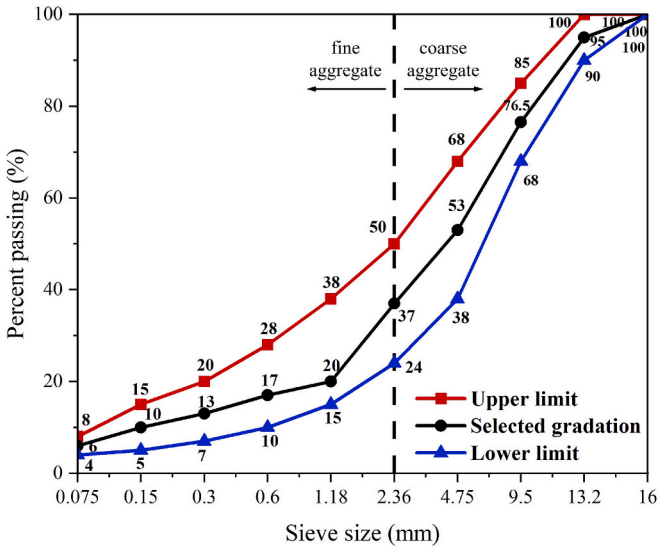


Fig. 8. AC-13 gradation for aggregates.

Table 4
Simulation parameters for discrete element models.

Young's modulus [Pa]	Density [kg/m^3]	Stiffness ratio	Poisson's ratio	Coefficient of restitution
5.5×10^{10}	2650	2.6	0.3	0.2
Local damping	Static friction coefficient	Rolling friction coefficient	Time step [s]	
0	0.5	0.35	8.0×10^{-8}	

simulations were determined by referring to previous studies [50–52], as shown in Table 4. Aggregate packing simulation was used to investigate the influence of the augmented database on aggregate packing structure, as illustrated in Fig. 9 (a). 9000 aggregates were randomly generated within a predefined container and allowed to settle under gravity until mechanical equilibrium was reached. Then, the aggregate pile was pre-compacted to a height of 70 mm using a disk applying a pressure of 600 kPa. This process captured *meso*-scale packing characteristics, including aggregate orientation, void distribution, and inter-particle contact network, which are critical for representing realistic mixture skeletons. After pre-compaction, packing density was used to characterize the proportion of space occupied by aggregates in the pile system:

$$PD = \frac{V_{Aggregate}}{V_{Container}} \tag{30}$$

where $V_{Aggregate}$ and $V_{Container}$ are the total volume of the aggregate and the volume of the filling container, respectively.

Furthermore, a Superpave Gyratory Compaction (SGC) simulation was performed to reproduce the laboratory compaction process of asphalt mixtures. The procedure of the SGC simulation is illustrated in Fig. 9 (b). A cylindrical mold was created to confine the loose asphalt mixture model, with two disks placed at the top and bottom ends. The cylinder was subjected to a controlled gyratory motion at a fixed angle (0.82°) and rotational speed (30 rpm). The bottom disk was fixed, while a constant vertical pressure (600 kPa) was applied through the top disk until the asphalt mixture specimen was compacted to a target height of 63.5 mm. This process enabled direct observation of aggregate rearrangement, load transfer, and the evolution of air voids during compaction, thereby providing insights into the mechanical response of mixtures prepared from both the real and augmented aggregate databases. During this process, both the packing density of the compacted asphalt mixture model (see Eq. (30)) and the gyration-height curve of

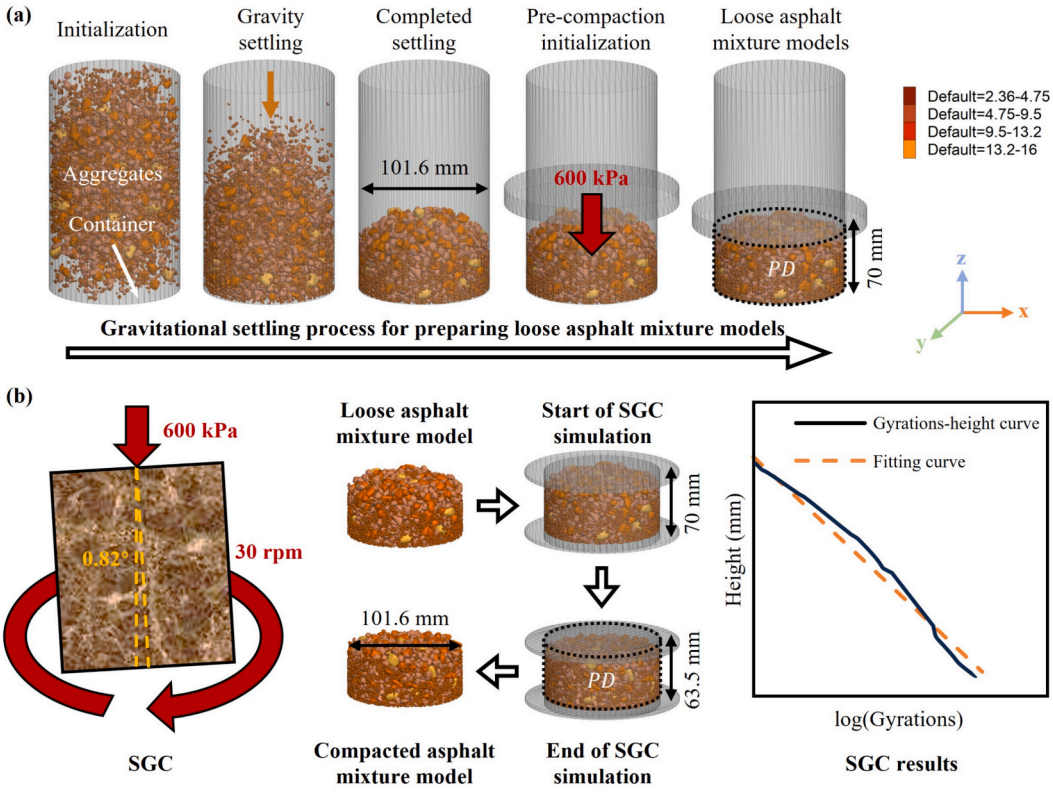


Fig. 9. Preparation of virtual Marshall specimens based on the DEM: (a) preparation of loose asphalt mixture model, (b) SGC simulation.

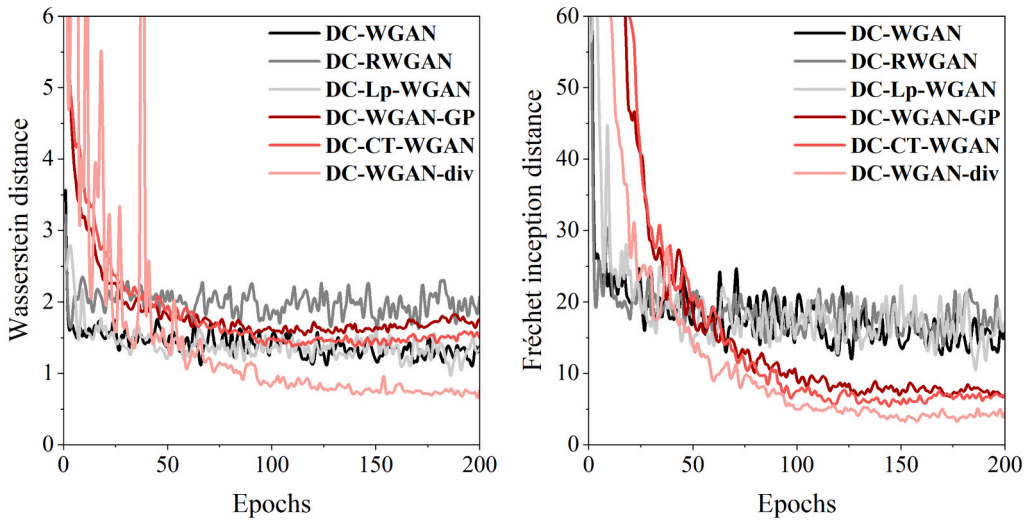


Fig. 10. Wasserstein distance and Fréchet inception distance for the training process of DC-WGANs.

the sample were recorded. Then, the gyration-height curve was fitted using the following equation:

$$H = AG^B \tag{31}$$

where H and G denote the height and gyration number of the specimen, respectively; A and B are the parameters to be fitted.

2.3. Database augmentation based on database lightweight

Following the multi-criteria evaluation presented above, an optimal variant of the DC-WGAN model can be identified based on augmentation quality and morphological fidelity. On this basis, the feasibility of a lightweight database augmentation strategy was further investigated, in which the generative model was trained and applied separately for each aggregate size category using a reduced subset of the original data. While previous studies have demonstrated that reducing training data may degrade model performance [53,54], adopting a lightweight configuration can significantly improve training efficiency and reduce computational burden. Therefore, it is necessary to evaluate whether this trade-off impacts the quality and diversity of the augmented aggregates. In this study, the same augmentation pipeline was applied as in the full-database scenario, with the only difference being the reduction in data volume per size group. A systematic comparison was conducted to evaluate whether the lightweight approach can maintain satisfactory augmentation performance while enhancing computational efficiency, offering a practical solution for database design under resource-constrained conditions.

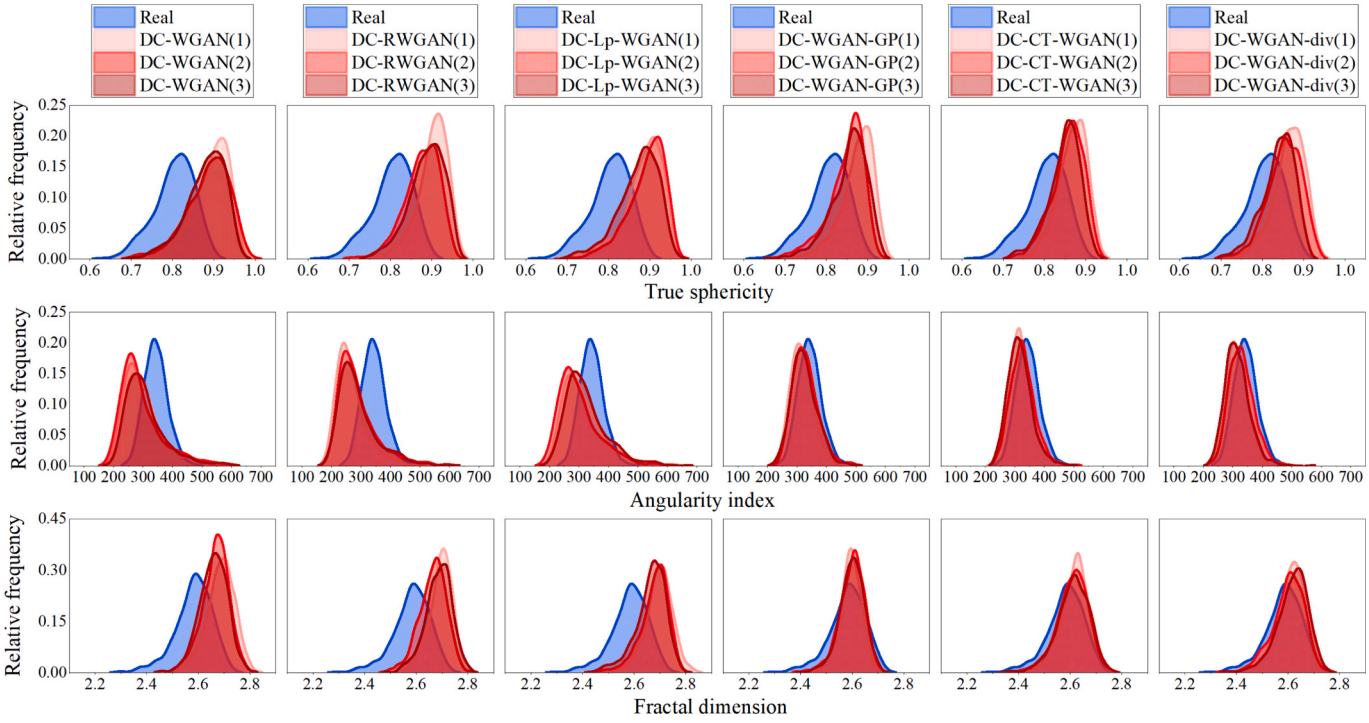


Fig. 11. Diversity validation of generator for DC-WGANs.

Table 5
Analysis of variance for true sphericity.

	Source	Degree of freedom	Sum of square	Mean square	F-value	P-value	Significant
DC-WGAN	Groups	3	7.238	2.413	982.588	2.187e-16	**Yes**
	Error	5756	14.134	0.003			
	Total	5759	21.372	0.004			
DC-RWGAN							
	Groups	3	7.923	2.641	1345.938	4.441e-16	**Yes**
	Error	5756	11.294	0.002			
	Total	5759	19.217	0.003			
DC-Lp-WGAN							
	Groups	3	6.924	2.308	961.512	2.221e-16	**Yes**
	Error	5756	13.817	0.002			
	Total	5759	20.741	0.004			
DC-WGAN-GP							
	Groups	3	3.796	1.265	627.585	1.443e-15	**Yes**
	Error	5756	11.605	0.002			
	Total	5759	15.402	0.003			
DC-CT-WGAN							
	Groups	3	3.329	1.110	657.116	2.219e-16	**Yes**
	Error	5756	9.719	0.002			
	Total	5759	13.048	0.002			
DC-WGAN-div							
	Groups	3	2.986	0.995	543.580	1.110e-16	**Yes**
	Error	5756	10.540	0.002			
	Total	5759	13.526	0.002			

Table 6
Analysis of variance for angularity index.

	Source	Degree of freedom	Sum of square	Mean square	F-value	P-value	Significant
DC-WGAN	Groups	3	2.144e6	7.147e5	172.136	2.442e-15	**Yes**
	Error	5756	2.390e7	4.152e3			
	Total	5759	2.604e7	4.522e3			
DC-RWGAN							
	Groups	3	4.480e6	1.493e6	421.217	4.335e-16	**Yes**
	Error	5756	2.041e7	3.545e3			
	Total	5759	2.489e7	4.321e3			
DC-Lp-WGAN							
	Groups	3	1.382e6	4.606e5	104.716	1.655e-15	**Yes**
	Error	5756	2.532e7	4.398e3			
	Total	5759	2.670e7	4.636e3			
DC-WGAN-GP							
	Groups	3	5.187e5	1.729e5	94.152	4.422e-16	**Yes**
	Error	5756	1.057e7	1.836e3			
	Total	5759	1.109e7	1.925e3			
DC-CT-WGAN							
	Groups	3	5.876e5	1.959e5	116.472	1.665e-15	**Yes**
	Error	5756	9.680e6	1.681e3			
	Total	5759	1.027e7	1.782e3			
DC-WGAN-div							
	Groups	3	6.273e5	2.091e5	116.306	4.441e-16	**Yes**
	Error	5756	1.034e7	1.797e3			
	Total	5759	1.097e7	1.905e3			

Table 7
Analysis of variance for fractal dimension.

	Source	Degree of freedom	Sum of square	Mean square	F-value	P-value	Significant
DC-WGAN	Groups	3	9.120	3.040	838.094	4.444e-16	**Yes**
	Error	5756	20.879	0.003			
	Total	5759	29.999	0.005			
DC-RWGAN							
	Groups	3	11.045	3.682	1076.768	2.220e-16	**Yes**
	Error	5756	19.681	0.003			
	Total	5759	30.726	0.005			
DC-Lp-WGAN							
	Groups	3	10.861	3.620	927.988	1.221e-15	**Yes**
	Error	5756	22.456	0.004			
	Total	5759	33.317	0.006			
DC-WGAN-GP							
	Groups	3	0.258	0.086	27.813	7.772e-15	**Yes**
	Error	5756	17.797	0.003			
	Total	5759	18.055	0.003			
DC-CT-WGAN							
	Groups	3	0.959	0.320	77.054	1.554e-15	**Yes**
	Error	5756	23.871	0.004			
	Total	5759	24.830	0.004			
DC-WGAN-div							
	Groups	3	1.031	0.344	79.199	1.112e-16	**Yes**
	Error	5756	24.986	0.004			
	Total	5759	26.017	0.005			

3. Results and discussion

3.1. Evaluation of augmented database

3.1.1. Training model evaluation of DC-WGAN variants

After setting all the hyperparameters according to Table 1, each DC-WGAN variant was trained separately. Fig. 10 shows the trends of WD and FID for DC-WGANs. The WD values for all DC-WGANs except DC-WGAN-div dropped to near-stable values before 50 epochs. Then, the trends flattened out. The WD values of DC-WGAN-div declined until 150 epochs, after which they flattened out. The first five variants exhibited the same pattern due to the loss function of D always satisfying the "1-Lipschitz" continuity constraint. This strong constraint caused the WD value to stabilize earlier. DC-WGAN-div exhibited a different pattern because it did not need to satisfy the Lipschitz continuity constraint. This relaxed condition caused the WD value to stabilize later (see Section 2.2.1). At epoch 200, the WD values for the six variants were 1.3776, 2.0765, 1.3480, 1.7579, 1.4916, and 0.7665. The FID values of DC-WGAN, DC-RWGAN, and DC-Lp-WGAN flattened out before 50 epochs. In comparison, the FID values of DC-WGAN-GP, DC-CT-WGAN, and DC-WGAN-div started to flatten out only after 100 epochs. The values of the first three variants were very close, while the values of the last three variants were relatively close. This can be attributed to the fact that the first three variants use the weight penalty method, which directly restricts all parameters to a compact space, resulting in curves that stabilize more quickly but with less smooth transitions and greater oscillations. In contrast, the latter three variants add normalization terms to the loss function of D, which essentially constrains the loss function, resulting in smoother curves with smaller oscillations. At epoch 200, the FID values for the six variants were 16.0753, 17.1154, 17.2799, 6.7502, 6.5897, and 4.5973. Overall, based on the results alone, the augmented data from DC-WGAN-div showed the highest similarity to the real data in terms of mathematical abstraction, i.e., the model training performance was the best. At epoch 200, the WD and FID values of DC-WGAN-div corresponded to the smallest values among all variants, which were 0.7665 and 4.5973, respectively.

After the alternate training of the model is completed, it is only necessary to generate new data using G. However, to ensure the generalization ability of G, the diversity of the augmented data should first be verified, i.e., the distributions corresponding to each augmented database should be different. For each DC-WGAN variant, each G was run 3 times to generate new data. Each time, 1440 voxel models were generated to match the original data volume, ensuring comparability. Then, after a series of post-processing steps (cf. Section 2.2.3), the aggregates augmented from the full database were obtained. To compare the differences between the different databases, the true sphericity, angularity index, and fractal dimension of each aggregate in the databases were calculated. Fig. 11 shows the distribution of each indicator in each database under the different variants, where the blue curves represent the data distributions for the real aggregates, and the different red curves represent the data distributions for the augmented aggregates. Although some overlap existed among the augmented data distributions, variability in their shapes remained. Specifically, while the distributions from G in each plot exhibited similar peaks, their ranges differed, demonstrating that the well-trained G generalized well and produced diverse outputs. Furthermore, the distribution of the augmented data resembled that of the real data, indicating that the model effectively learned the distribution patterns in the database.

After validating the generalization ability of G through distributional comparisons, a statistical analysis was conducted to rigorously assess whether the differences observed among the databases were statistically significant. To this end, an analysis of variance (ANOVA) was conducted on the morphological indicators of aggregates, including true sphericity, angularity index, and fractal dimension, across the different augmented and real databases. The significance level was set at 0.05. Table 5, Table 6, and Table 7 present the ANOVA results for the three morphological indicators, respectively. In all cases, the computed p-values were well below the 0.05 threshold, while the corresponding F-values were consistently large and exceeded the critical values associated with the 95 % confidence interval. These results indicate that the null hypothesis

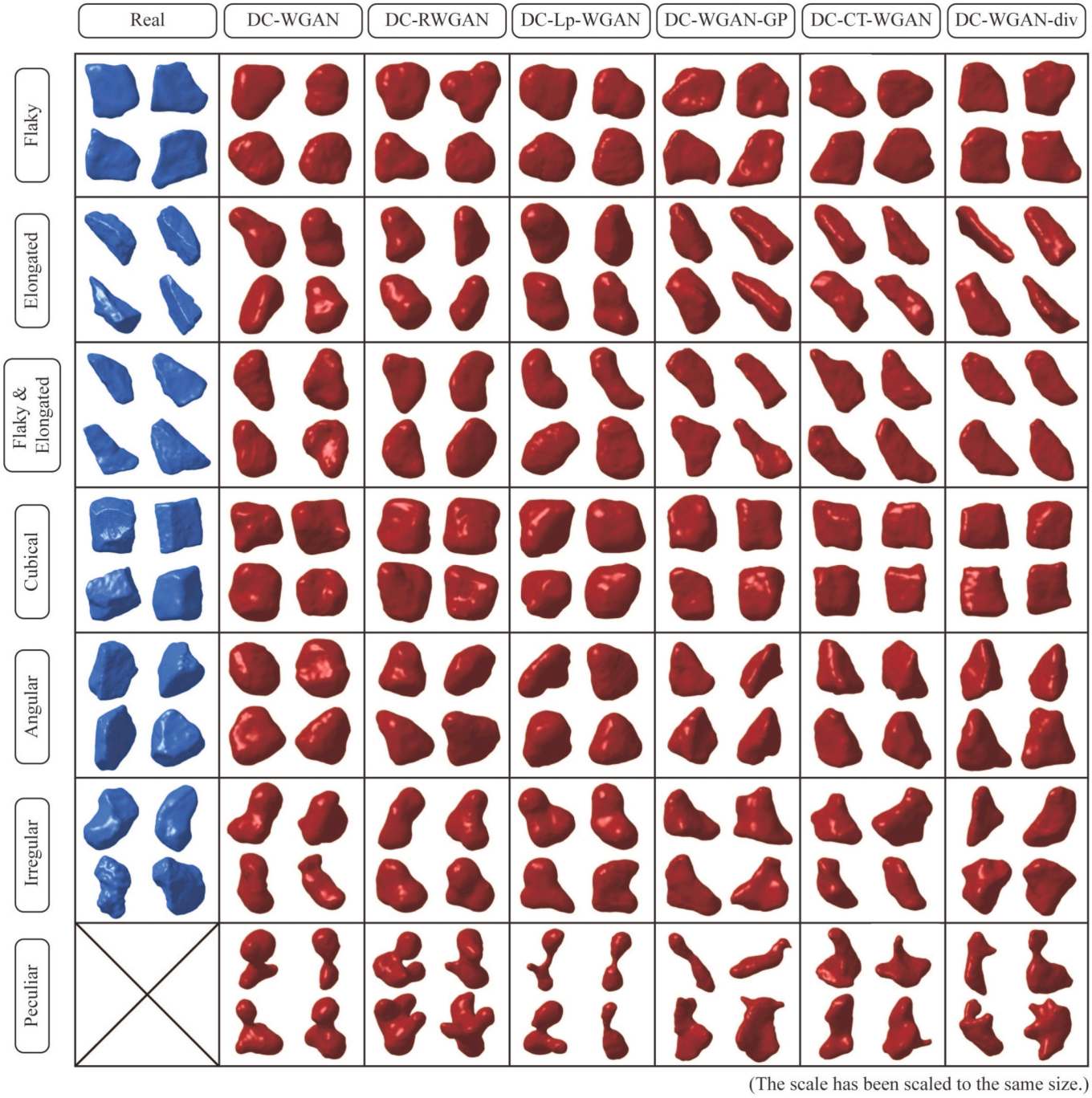


Fig. 12. Visualization for real and augmented aggregates.

H_0, which states that no significant differences exist among the groups, can be rejected at the 5 % significance level. Hence, the morphological characteristics derived from different databases follow distinct distributions, and the variations are statistically significant. Such findings provide robust evidence that the proposed models achieve diversity in the augmented aggregates while maintaining consistency with real data distributions.

3.1.2. *Morphological comparison of augmented virtual aggregates*

To evaluate the quality of augmented aggregates from different DC-WGANs, it is necessary to select representative real and augmented aggregates for visualization. As shown in Fig. 12, six common shape types of aggregates (flaky, elongated, flaky & elongated, cubical, angular, irregular) were selected for visualization. Then, four representative instances of each type were selected. To ensure the comparability of the visualization results, all aggregates were visualized from the same viewpoint. Overall, aggregates augmented by DC-WGAN, DC-RWGAN, and DC-Lp-WGAN were of lower quality compared with those produced by DC-WGAN-GP, DC-CT-WGAN, and DC-WGAN-div. The aggregates augmented based on the latter three variants exhibited finer details and were more similar to real aggregates in appearance. Specifically, the aggregates augmented by the latter three variants exhibited more pronounced angularity and unevenness, indicating better generalization of the augmented data. In contrast, the aggregates augmented by the former three variants were smoother and closer to spherical, indicating that the augmented data had slightly weaker generalizability. However, regardless of which DC-WGAN was used, some aggregates with atypical shapes were always generated. Although

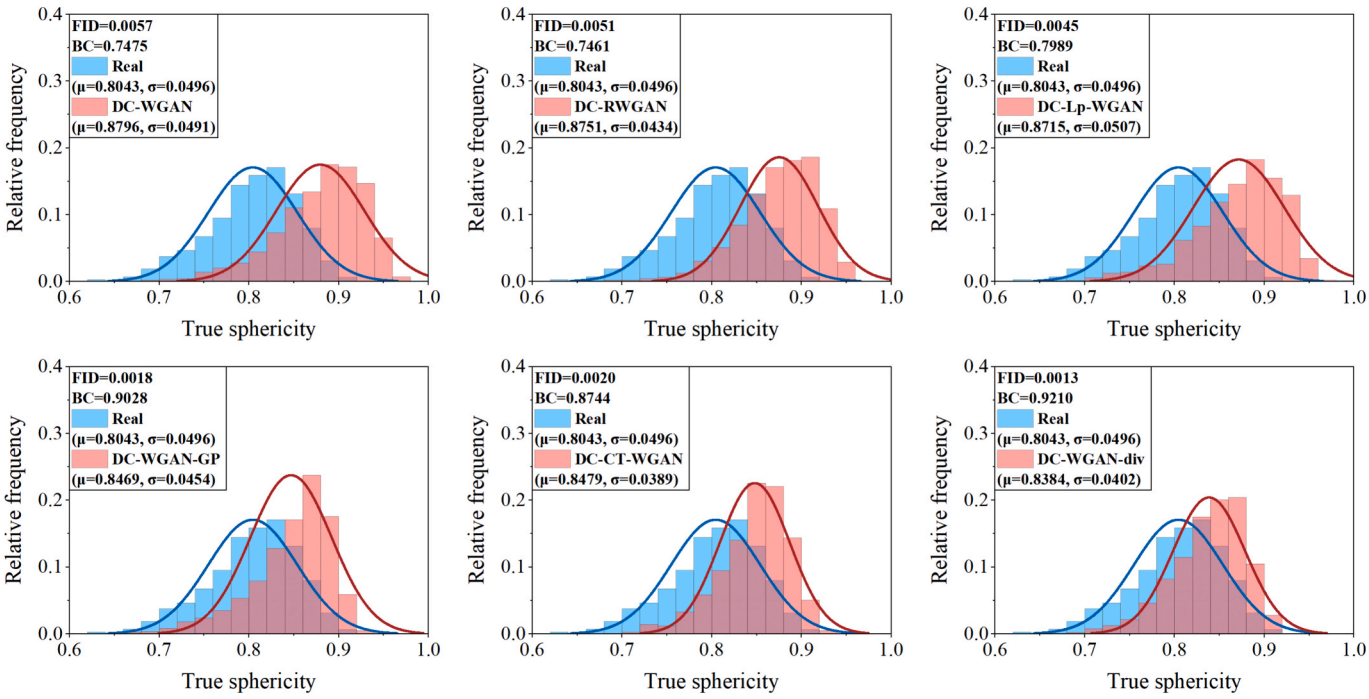


Fig. 13. Comparison of true sphericity for virtual aggregates from different DC-WGANs.

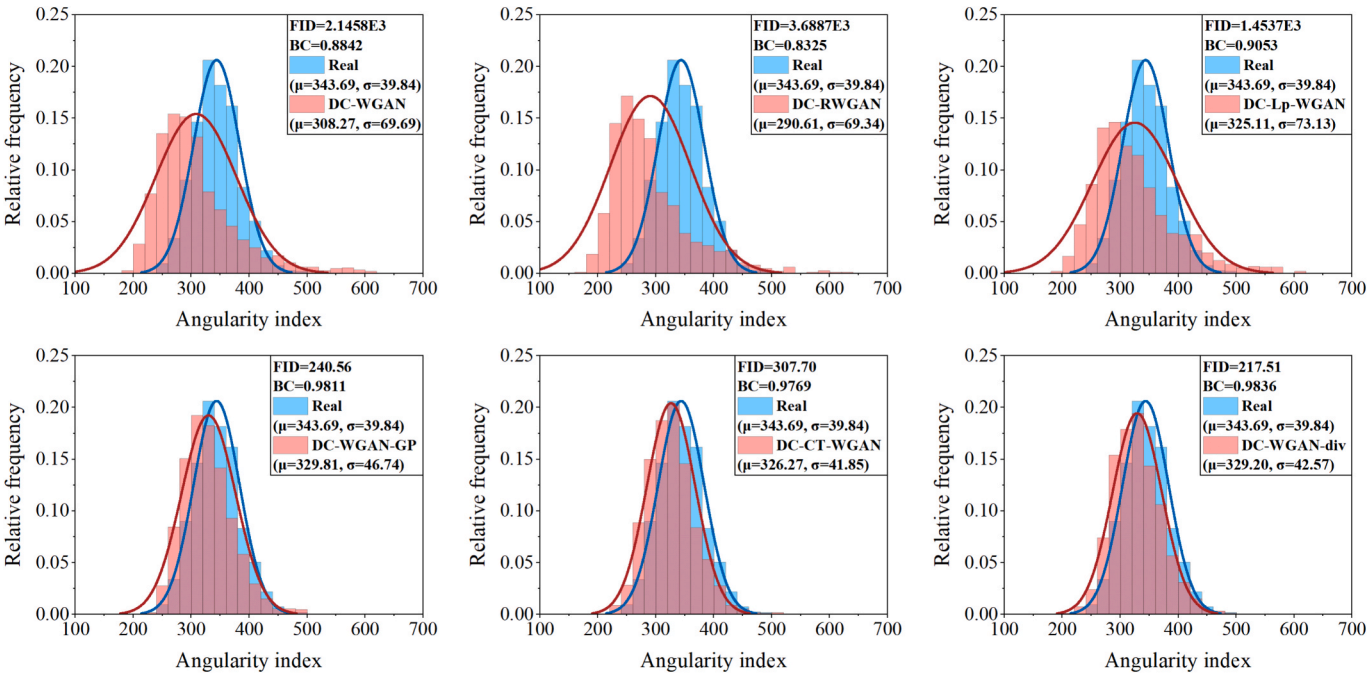


Fig. 14. Comparison of angularity index for virtual aggregates from different DC-WGANs.

these shapes are mathematically reasonable because they are generated based on the data distribution, they may lack practical relevance. Therefore, such aggregates should be screened out in actual engineering applications.

Based on Fig. 12, the augmented aggregates generated by different DC-WGANs were qualitatively evaluated via visual inspection. Following the visual evaluation, a subsequent quantitative analysis compared aggregates at three scales. For each augmented database, the instance most similar to the real database was selected. Then, FID_{Φ} (see Eq. (16)) and BC_{Φ} (see Eq. (17)) were used to assess differences between the two databases across three scales. To facilitate the quantitative analysis, all data were assumed to follow a normal distribution.

Fig. 13, Fig. 14, and Fig. 15 show the distribution of virtual aggregates augmented using different DC-WGAN variants in terms of true sphericity, angularity index, and fractal dimension, as well as comparisons with real aggregate data. Overall, the distributions of true sphericity for aggregates generated by all variants were consistent with the real data in terms of overall contour, exhibiting a unimodal distribution characteristic. However, there were still obvious differences between the variants in terms of mean, standard deviation, and distribution shift

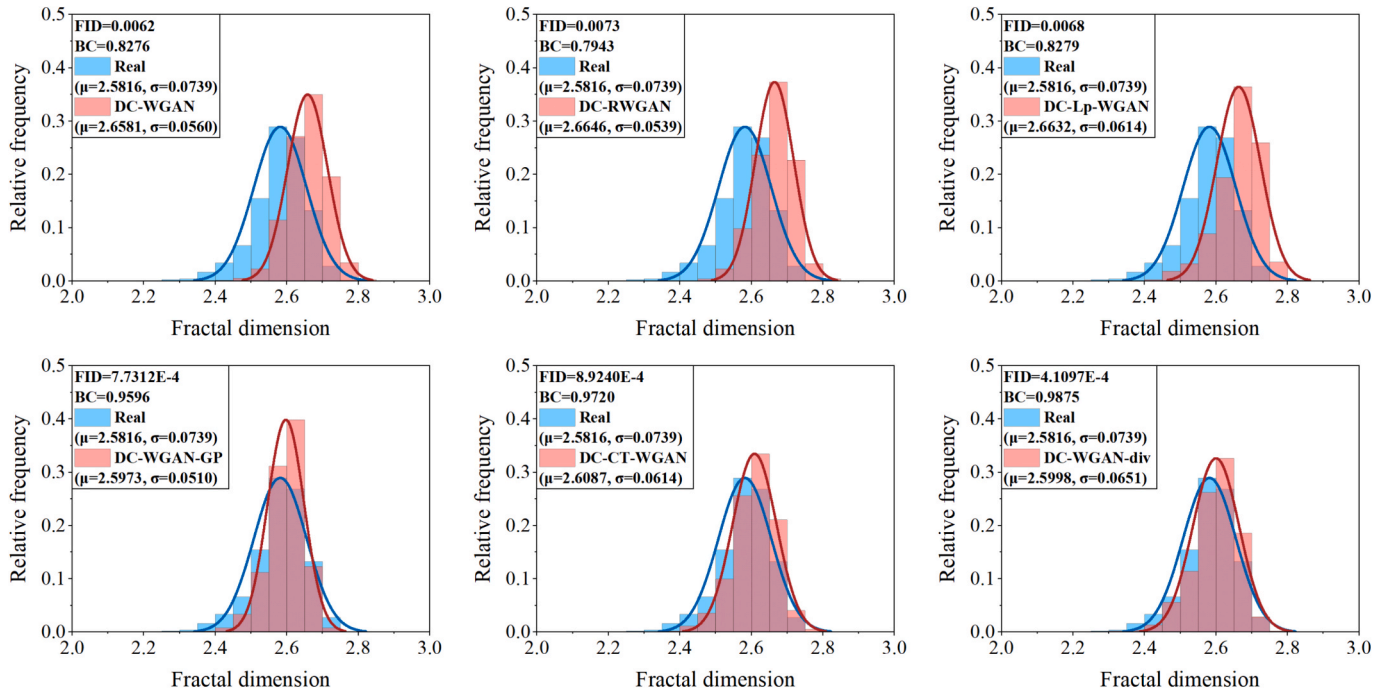


Fig. 15. Comparison of fractal dimension for virtual aggregates from different DC-WGANs.

direction.

Specifically, for the true sphericity, compared with the real data ($\mu_R = 0.8043$, $\sigma_R = 0.0496$), the mean values of the data augmented by DC-WGAN, DC-RWGAN, and DC-Lp-WGAN were 0.8796, 0.8751, and 0.8715, respectively, indicating a significant rightward shift, while the standard deviations were similar (0.0491, 0.0434, and 0.0507, respectively), indicating that the augmented data generally tended toward higher sphericity with no significant changes in variability. In contrast, the mean values of the data augmented by DC-WGAN-GP, DC-CT-WGAN, and DC-WGAN-div were 0.8469, 0.8479, and 0.8384, respectively, which were close to the real data, and the standard deviations were smaller (0.0454, 0.0389, and 0.0402, respectively), indicating that the augmented samples exhibited better concentration and consistency in sphericity features. From the quantitative indicators, DC-WGAN-div corresponded to the lowest FID value (0.0013) and the highest BC value (0.9210), indicating that the augmented data were most similar to the real data in terms of distribution shape and statistical characteristics. In contrast, DC-RWGAN exhibited a higher FID value (0.0051) and a lower BC value (0.7461), indicating the greatest distribution divergence between the augmented and real data, which suggests that its augmentation quality was the poorest.

For the angularity index, compared with real data ($\mu_R = 343.69$, $\sigma_R = 39.84$), the mean values of the data augmented by DC-WGAN, DC-RWGAN, and DC-Lp-WGAN were 308.27, 290.61, and 325.11, respectively, indicating an overall leftward shift in the distribution. This reflects a reduction in the angularity of the augmented aggregates. Moreover, the standard deviations increased substantially (69.69, 69.34, and 73.13, respectively), suggesting greater variability in the angularity of the augmented aggregates. In contrast, the mean values of the data augmented by DC-WGAN-GP, DC-CT-WGAN, and DC-WGAN-div were 329.81, 326.27, and 329.20, respectively, which were close to the real data, and the corresponding standard deviations were smaller (46.74, 41.85, and 42.57, respectively), reflecting greater central tendency and consistency in angularity features. DC-WGAN-div exhibited the lowest FID value (217.51) and the highest BC value (0.9836), demonstrating superior augmentation performance. In contrast, the DC-RWGAN showed the highest FID value (3.6887E3) and the lowest BC value (0.8325), indicating the poorest augmentation quality.

For the fractal dimension, compared to the real data ($\mu_R = 2.5816$, $\sigma_R = 0.0739$), the mean values of the augmented data from DC-WGAN, DC-RWGAN, and DC-Lp-WGAN were 2.6581, 2.6646, and 2.6632, respectively, indicating a rightward shift. This suggested that the augmented data was highly likely to exhibit distortion at the fine scale. However, the differences in standard deviation were not significant (0.0560, 0.0539, and 0.0614). In contrast, the mean values of the data augmented by DC-WGAN-GP, DC-CT-WGAN, and DC-WGAN-div were 2.5973, 2.6087, and 2.5998, respectively, which were close to those of the real data, and the corresponding standard deviations were also similar (0.0510, 0.0614, and 0.0651, respectively), indicating that the augmented samples closely resembled the real samples in terms of fine scale representation. DC-WGAN-div showed the lowest FID value (4.1097e-4) and highest BC value (0.9875), demonstrating superior augmentation quality. In contrast, DC-RWGAN exhibited the highest FID value (0.0073) and lowest BC value (0.7943), indicating the poorest augmentation performance.

In summary, the distribution of aggregate morphology indicators augmented by different variants exhibited significant differences. These differences can be attributed to the different constraint strategies applied to the loss function of D. DC-WGAN-div achieved higher fitting flexibility for D and G during adversarial training by eliminating the Lipschitz continuity constraint, thereby better capturing the complex distribution of morphological features in the real database. In contrast, traditional DC-WGAN and its slightly improved variants (e.g., DC-RWGAN, DC-Lp-WGAN) adopted overly simplistic weight penalty strategies, which separately limited the ability of D and G to recognize and express complex data distributions, resulting in overall distribution shifts and increased distribution dispersion in the augmented data. Overall, based on these comparisons, DC-WGAN-div demonstrated the best performance in terms of data quality and distributional consistency, providing a more reliable data foundation for subsequent studies on the mechanical behavior of virtual aggregates.

3.1.3. Mechanical response evaluation

To evaluate the mechanical response of asphalt mixtures constructed from both the real and augmented aggregate databases, two types of discrete element simulations were performed: aggregate packing and

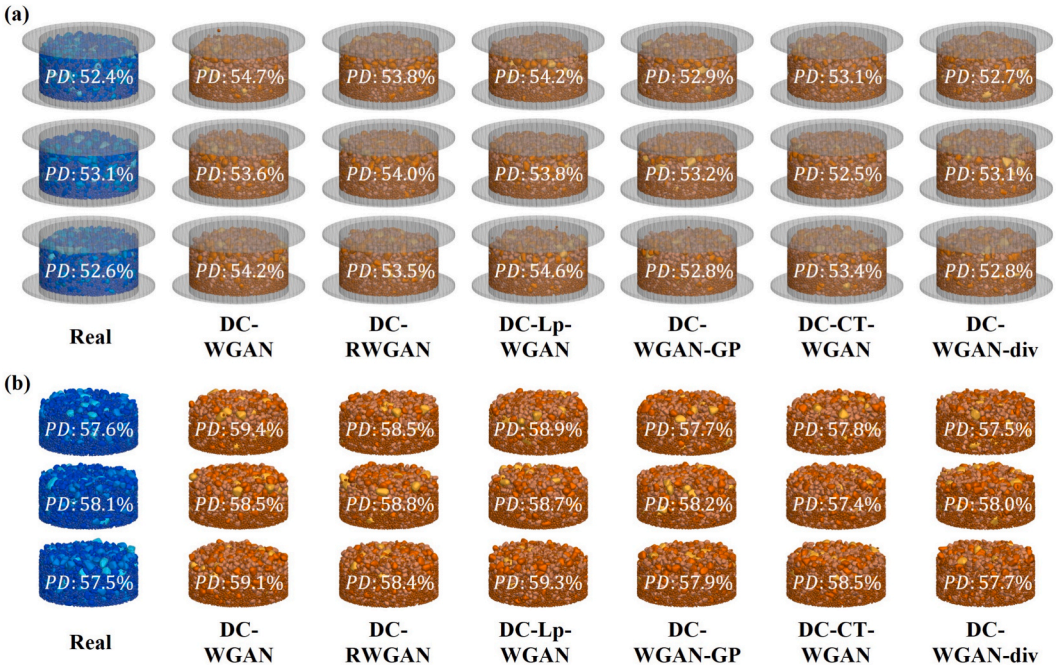


Fig. 16. Aggregate packing density results for (a) loose asphalt mixture model, (b) compacted asphalt mixture model.

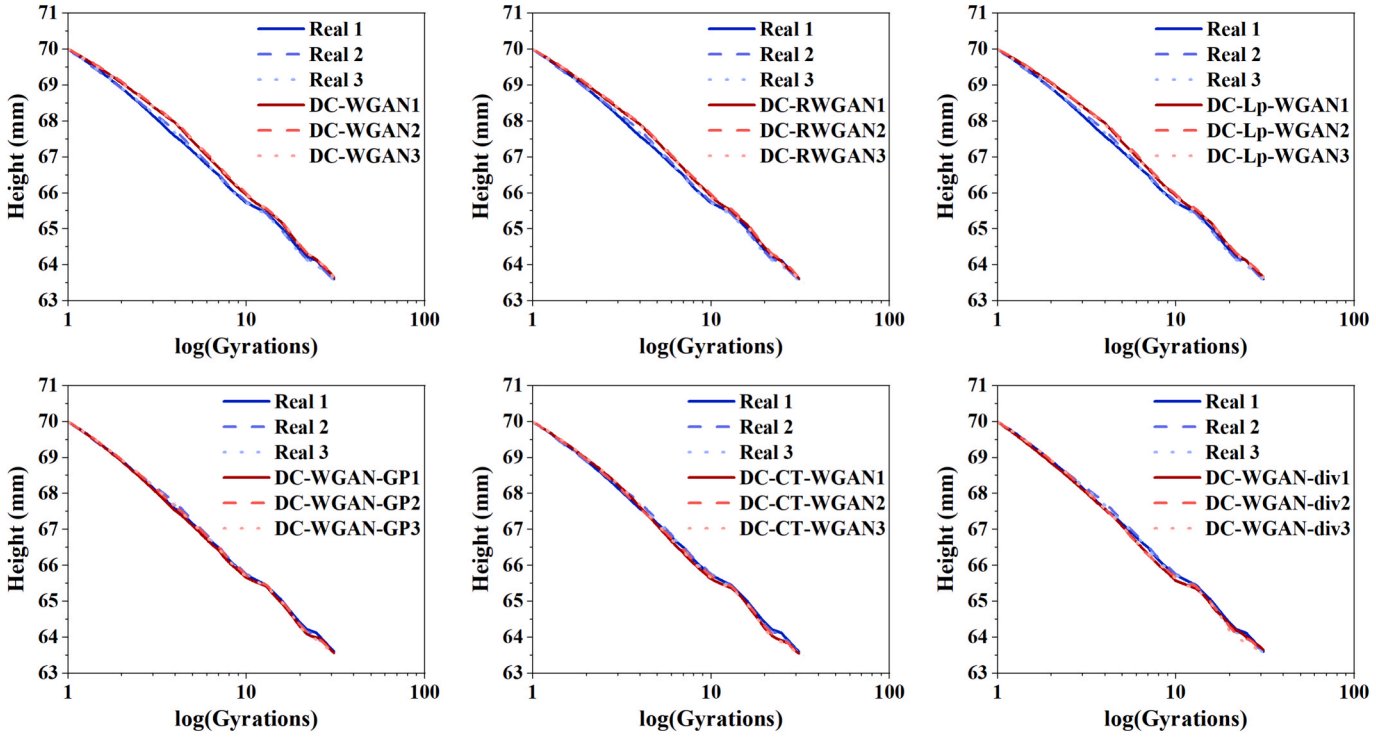


Fig. 17. Gyration-height curves for SGC simulations.

Superpave Gyratory Compaction (SGC). To reduce the variability of simulation results, each case was simulated three times. The aggregate packing simulation focused on *meso*-scale packing characteristics, while the SGC simulation reproduced the laboratory compaction process, enabling analysis of aggregate reorientation, densification, and air void evolution.

Fig. 16 (a) and Fig. 16 (b) show the packing density (PD) results for loose asphalt mixture and compacted asphalt mixture, respectively. As shown in Fig. 16 (a), the aggregate packing simulations of loose asphalt mixtures yielded PD in the range of approximately 52.4 %–54.7 %. The values obtained from mixtures generated using different augmentation methods were generally consistent with the real aggregate packing structure, with variations within about 2 %. Among the models, DC-WGAN and DC-Lp-WGAN produced slightly higher PD values, suggesting a tendency to form denser packing structures, whereas DC-WGAN-div and DC-CT-WGAN were closer to the reference real mixtures. This indicates that while augmented databases can reproduce the overall *meso*-scale packing characteristics, subtle differences in particle

Table 8
Fitting of gyration-height curves.

	Sample 1	Sample 2	Sample 3
Real	$H = 70.2844G^{-0.0288}R^2 = 0.9969$	$H = 70.3530G^{-0.0292}R^2 = 0.9962$	$H = 70.3407G^{-0.0294}R^2 = 0.9965$
DC-WGAN	$H = 70.5396G^{-0.0294}R^2 = 0.9936$	$H = 70.5946G^{-0.0295}R^2 = 0.9925$	$H = 70.5563G^{-0.0295}R^2 = 0.9926$
DC-RWGAN	$H = 70.5123G^{-0.0295}R^2 = 0.9941$	$H = 70.5535G^{-0.0295}R^2 = 0.9932$	$H = 70.5383G^{-0.0298}R^2 = 0.9940$
DC-Lp-WGAN	$H = 70.5485G^{-0.0296}R^2 = 0.9939$	$H = 70.5932G^{-0.0297}R^2 = 0.9929$	$H = 70.5284G^{-0.0296}R^2 = 0.9939$
DC-WGAN-GP	$H = 70.2630G^{-0.0290}R^2 = 0.9964$	$H = 70.3704G^{-0.0294}R^2 = 0.9959$	$H = 70.3659G^{-0.0296}R^2 = 0.9954$
DC-CT-WGAN	$H = 70.3264G^{-0.0295}R^2 = 0.9949$	$H = 70.3421G^{-0.0294}R^2 = 0.9958$	$H = 70.3805G^{-0.0298}R^2 = 0.9945$
DC-WGAN-div	$H = 70.1700G^{-0.0285}R^2 = 0.9972$	$H = 70.2606G^{-0.0291}R^2 = 0.9958$	$H = 70.3105G^{-0.0296}R^2 = 0.9951$

arrangement exist depending on the augmentation strategy.

In the case of compacted asphalt mixtures, illustrated in Fig. 16 (b), the PD values increased markedly to a range of 57.4 %–59.4 % after SGC. The results showed good consistency across both real and augmented databases, with only minor deviations (<2%). Similar to the loose mixtures, DC-WGAN and DC-Lp-WGAN tended to yield slightly higher packing densities compared to the real mixtures, reflecting their stronger ability to simulate aggregate densification during compaction. Meanwhile, DC-WGAN-div and DC-CT-WGAN again exhibited the closest match with the reference, suggesting these models better preserved the realistic compaction behavior.

Overall, the comparative analysis demonstrates that the augmented aggregate databases can not only replicate the *meso*-scale characteristics of loose mixtures but also reliably capture the densification trends observed in laboratory compaction. The small deviations across different generative models highlight that while augmentation introduces minor variability, the mechanical responses inferred from the simulations remain consistent with those of real aggregates.

Fig. 17 illustrates the gyration–height curves obtained from the SGC simulations, while Table 8 presents the corresponding fitting equations and determination coefficients. Overall, the simulated mixtures constructed using the augmented aggregate databases reproduced the compaction behavior of real mixtures with high fidelity. The height reduction trends with increasing gyration number showed strong overlap between the real and augmented databases, indicating that the augmented aggregate databases effectively captured the densification process of asphalt mixtures during SGC.

The fitting results in Table 8 further confirmed this consistency. All databases achieved coefficients of determination (R^2) above 0.992, demonstrating excellent agreement between the simulated gyration–height curves and the fitted models. The exponents of the fitting equations, which characterized the compaction rate, remained within a narrow range of –0.0285 to –0.0298 across both real and augmented mixtures, suggesting only minor differences introduced by database augmentation. Importantly, among the generated models, DC-WGAN-GP, DC-CT-WGAN, and DC-WGAN-div exhibited the closest alignment with the real mixtures. Their fitted exponents and initial heights were highly consistent with those of the real databases, while their R^2 values

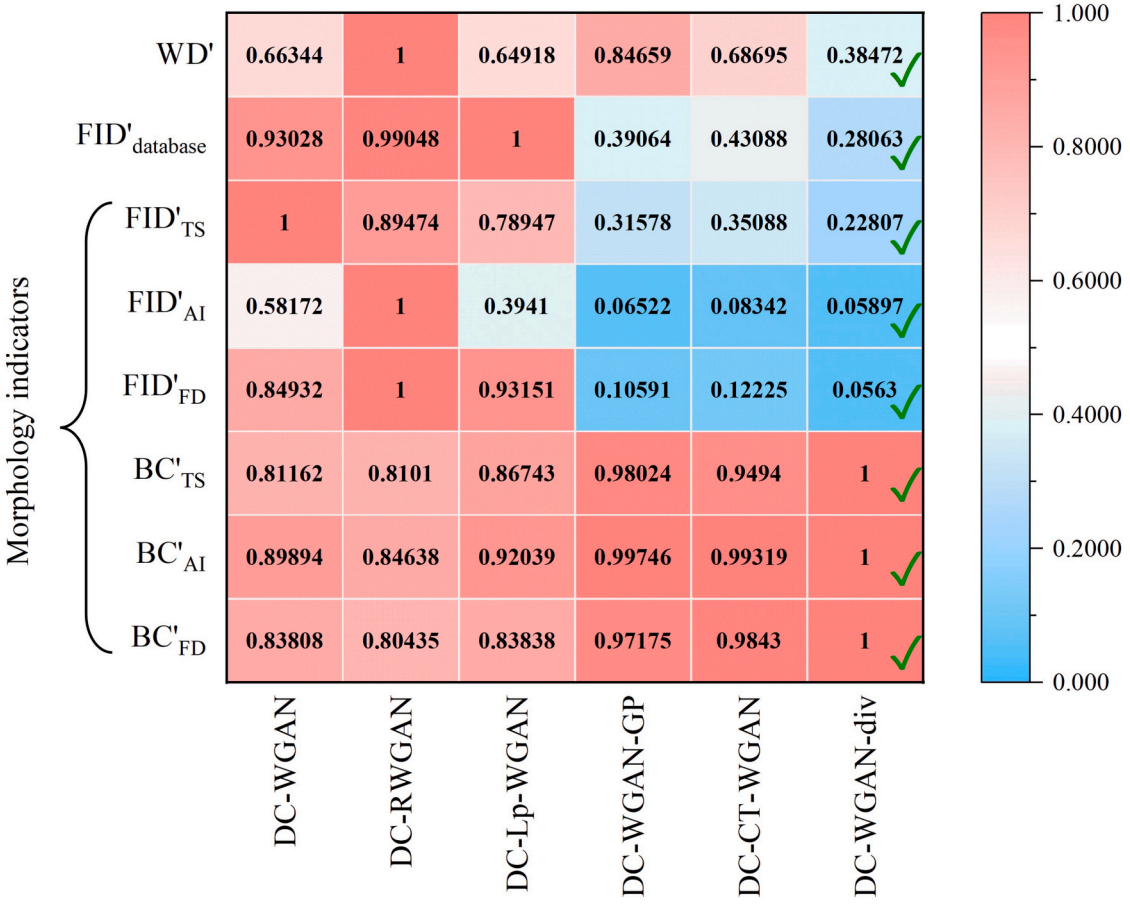


Fig. 18. Comprehensive evaluation for indicators.

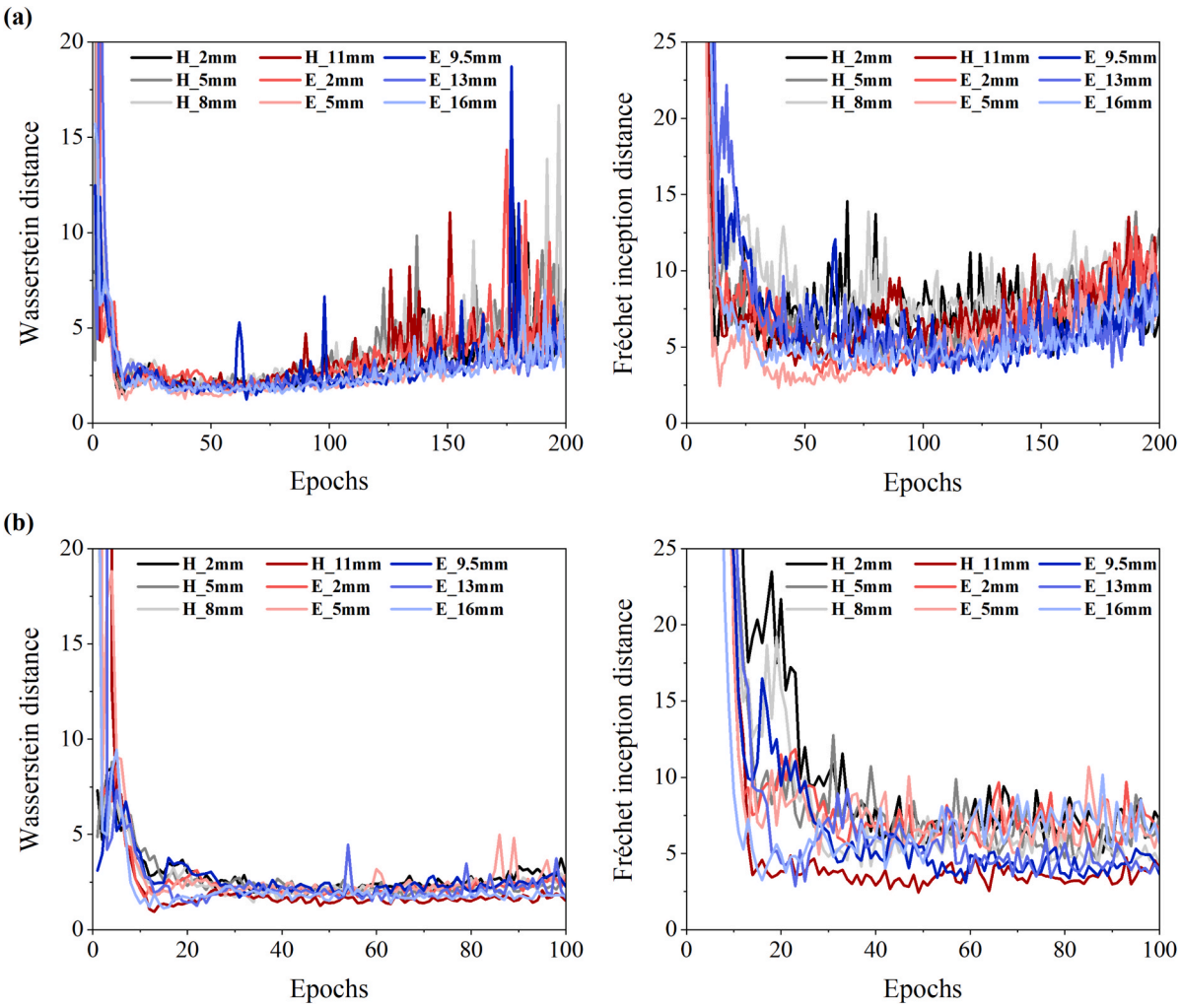


Fig. 19. Wasserstein distance and Fréchet inception distance for the training process of DC-WGAN-div: (a) with 200 training epochs, (b) with 100 training epochs.

(up to 0.9972) indicated superior curve-fitting accuracy. These results implied that the three models not only preserved the overall compaction trend but also better captured the detailed reorientation and densification mechanisms observed in the real mixtures.

In summary, the SGC simulation results validated the effectiveness of database augmentation, with all models yielding realistic compaction responses. Among them, DC-WGAN-GP, DC-CT-WGAN, and DC-WGAN-div demonstrated the strongest consistency with real mixtures, and thus provided the most reliable basis for mechanical response evaluation.

3.1.4. Comprehensive evaluation

Furthermore, to compare the augmentation performance of different DC-WGANs, a comprehensive evaluation was performed using eight indicators, namely, WD (see Eq. (9) & Eq. (10)), FID of database (see Eq. (11)), FID of the true sphericity, angularity index, and fractal dimension (see Eq. (16)), BC of the true sphericity, angularity index, and fractal dimension (see Eq. (17)). The values of WD and FID corresponded to those at the 200th epoch. Fig. 18 shows the values of the above eight indicators for different DC-WGANs, where all values are normalized. For WD and FID, lower values indicate better augmentation quality, while for BC, higher values indicate better quality. For each indicator, the value corresponding to DC-WGAN-div was either the minimum or maximum, indicating that DC-WGAN-div had the best augmentation performance. This is because DC-WGAN-div eliminates the Lipschitz continuity constraint, enabling the model to achieve better performance during training.

3.2. Feasibility analysis of database lightweight

3.2.1. Model training

After investigating data augmentation based on the full database, this section further explored the impact of database lightweight on data augmentation. To ensure comparability of the results, only one of the six DC-WGANs was used to train the model in this section. Based on benchmark analysis, DC-WGAN-div with the best augmentation performance (see Fig. 18) was selected for training on nine different databases (see Fig. 2). Initially, 200 epochs were selected to train the different databases. Fig. 19 (a) shows the trends of the WD and FID. However, after 100 epochs, the respective values tended to increase and oscillate noticeably, likely due to the small amount of data increasing the variability of the data distribution. Therefore, an early stopping strategy was applied by setting the number of epochs to 100, and the model was retrained. As shown in Fig. 19 (b), after reducing the epochs to 100, the values stabilize overall. Then, G was used to generate 160 new data samples for each database, followed by post-processing. In addition, the training time using both the full database and the database lightweight was also tested. For the full database, the training time was 53 min 33 sec (200 epochs). For the database lightweight, the average training time for the nine databases was 4 min 57 sec (100 epochs), roughly one-tenth of the original training time. This substantial reduction in training time has important practical implications for large-scale simulation projects. For example, when generating augmented databases for multiple aggregate types or conducting parametric studies

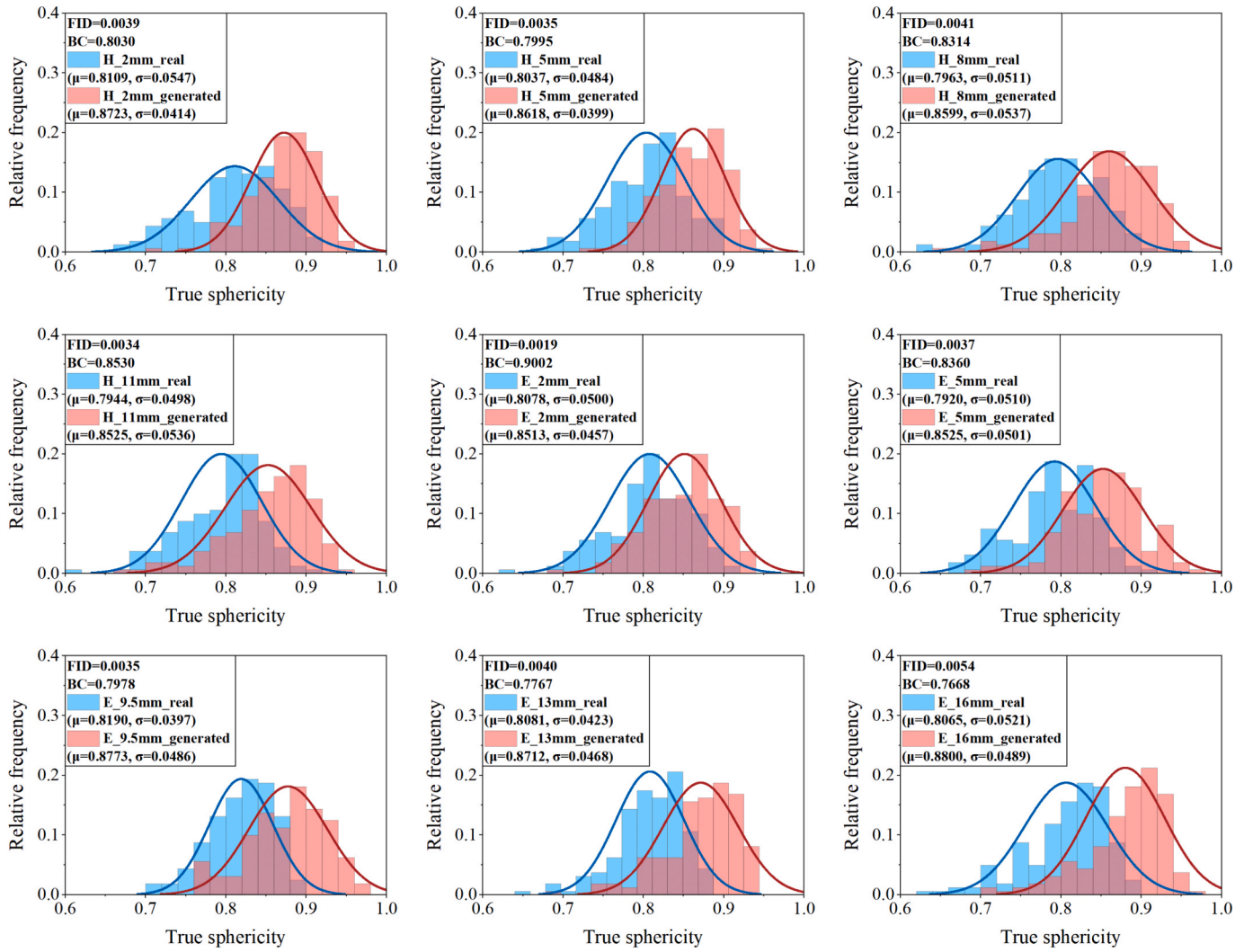


Fig. 20. Comparison of true sphericity for virtual aggregates from different grain size databases.

involving hundreds of simulations, using the lightweight database can save hundreds of hours of computational time. Such savings translate directly into reduced computational costs and faster project turnaround, enabling engineers and researchers to efficiently explore a wider range of aggregate morphologies and simulation scenarios without compromising the overall reliability of the augmented data.

3.2.2. *Morphological comparison of augmented virtual aggregates*

Fig. 20, Fig. 21, and Fig. 22 show the comparison of true sphericity, angularity index, and fractal dimension for virtual aggregates from different grain size databases. Compared with aggregates augmented by DC-WGAN-div using the full database (cf. Section 3.1.2), the quality of aggregates generated based on the database lightweight was generally lower. Based on the current results, the impact of the amount of original data on different morphological indicators can be further investigated.

For true sphericity, the augmented data distributions showed an overall decrease in overlap with the real data distributions. Compared with the augmented data distribution generated by DC-WGAN-div using the full database (cf. the last plot in Fig. 13), the corresponding augmented data distributions and mean values exhibited an overall rightward shift. This indicates that as the amount of original data decreases, the augmented aggregate shapes tend to become more spherical. However, the standard deviations of the augmented and real data distributions differ insignificantly, as reflected by the similarity between their contours. These results indicate that the training model can capture the data augmentation patterns well, but the amount of data limits the augmentation quality of the samples in terms of sphericity.

For angularity index, the differences in means between the two distributions were not significant, but the standard deviations of the augmented data distributions were larger, indicating that the augmented data exhibited greater variability. In addition, as the amount of original data decreased, the augmented data distributions did not show a noticeable trend of shifting to the left or right (cf. the last plot in Fig. 14). This indicates that the reduction in the amount of original data does not significantly affect the mean of the distribution. Overall, the overlaps between the augmented and real data distributions were relatively large, indicating that the reduction in the amount of original data had little effect on the augmentation quality of the samples in terms of angularity.

For fractal dimension, each augmented data distribution was generally shifted to the right of the real data distribution, but the differences were acceptable. A systematic rightward shift was observed in the augmented data distributions as the amount of original data decreased (cf. the last plot in Fig. 15), indicating that the augmented samples exhibited partial distortion at the fine scale. The differences in standard deviation between the augmented and real data distributions were relatively small, indicating that the augmented samples also exhibited relatively low variability in texture. Overall, the reduction in the amount of original data had an intermediate effect on the fractal dimension compared with the other two indicators.

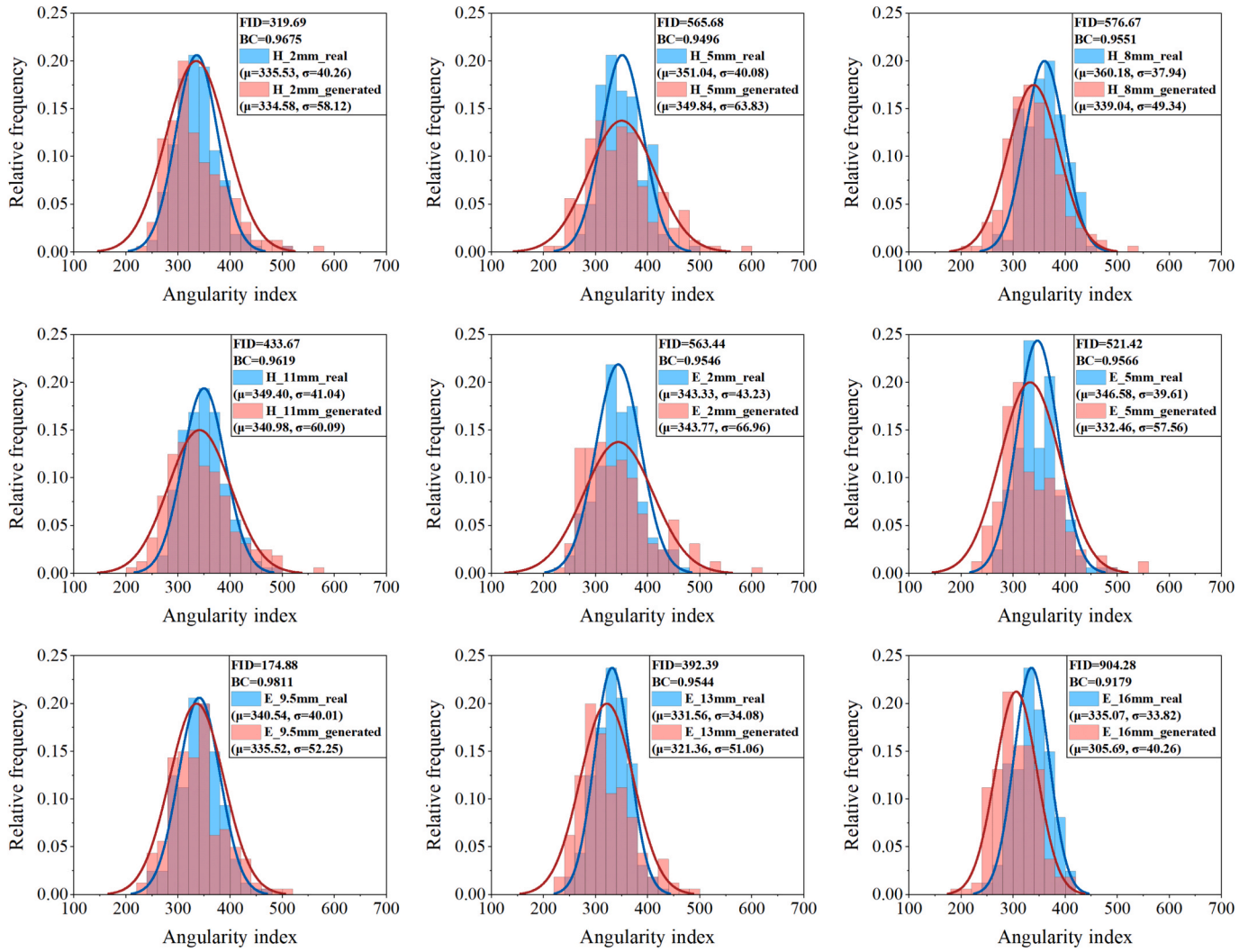


Fig. 21. Comparison of angularity index for virtual aggregates from different grain size databases.

Furthermore, the impact of the database lightweight on the quality of aggregate augmentation was quantitatively compared based on the BC values. Using the BC values from the original augmentation results of DC-WGAN-div (see last plot in Fig. 13, Fig. 14, and Fig. 15) as the benchmark, the percentage differences in BC values for the database lightweight are shown in Table 9. The database lightweight had the greatest impact on the true sphericity. For the nine different cases, the maximum difference in BC values was 16.74 %, and the minimum was 2.26 %. The database lightweight had the least impact on the angularity index. For different cases, the maximum difference in BC value was 6.68 %, and the minimum difference was 0.25 %. The database lightweight had an intermediate effect on the fractal dimension. For different cases, the maximum difference in BC value was 14.67 %, and the minimum difference was 1.00 %.

The above results indicated that changes in the amount of original data had the greatest impact on true sphericity, followed by the fractal dimension, and the least impact on the angularity index. Although the database lightweight had a certain negative impact on the quality of augmented aggregates, it offered significant advantages in training time. Therefore, after comprehensive consideration, it is valuable to use the database lightweight. Furthermore, to improve the quality of the augmented aggregates based on the database lightweight, a potential solution is to use larger voxels, which improves augmentation quality by increasing the resolution of the aggregate model.

4. Conclusions and outlook

In this study, the 3D augmented virtual aggregate databases were developed based on stereo vision and DC-WGANs. To comprehensively evaluate the quality of the databases augmented by different DC-WGANs, a multi-criteria evaluation approach was proposed. After identifying the optimal DC-WGAN model, the feasibility of database lightweight augmentation was further investigated. The main conclusions are as follows:

- The DC-WGAN variants exhibited different training performance. Overall, DC-WGAN, DC-RWGAN, and DC-Lp-WGAN demonstrated relatively low training accuracy, while DC-WGAN-GP, DC-CT-WGAN, and DC-WGAN-div showed higher training accuracy. Among all variants, DC-WGAN-div achieved the lowest WD and FID values at epoch 200, indicating the best overall training performance.
- The diversity of virtual aggregate augmentation was verified. The results demonstrated that the generative models exhibited good generalization ability and could produce virtual aggregates with a variety of distinct morphologies. Visually, the virtual aggregates augmented using DC-WGAN-GP, DC-CT-WGAN, and DC-WGAN-div had morphological features closer to those of the real aggregates. However, each variant occasionally generated some visually implausible virtual aggregates. Furthermore, the differences in morphological indicators between real aggregates and augmented

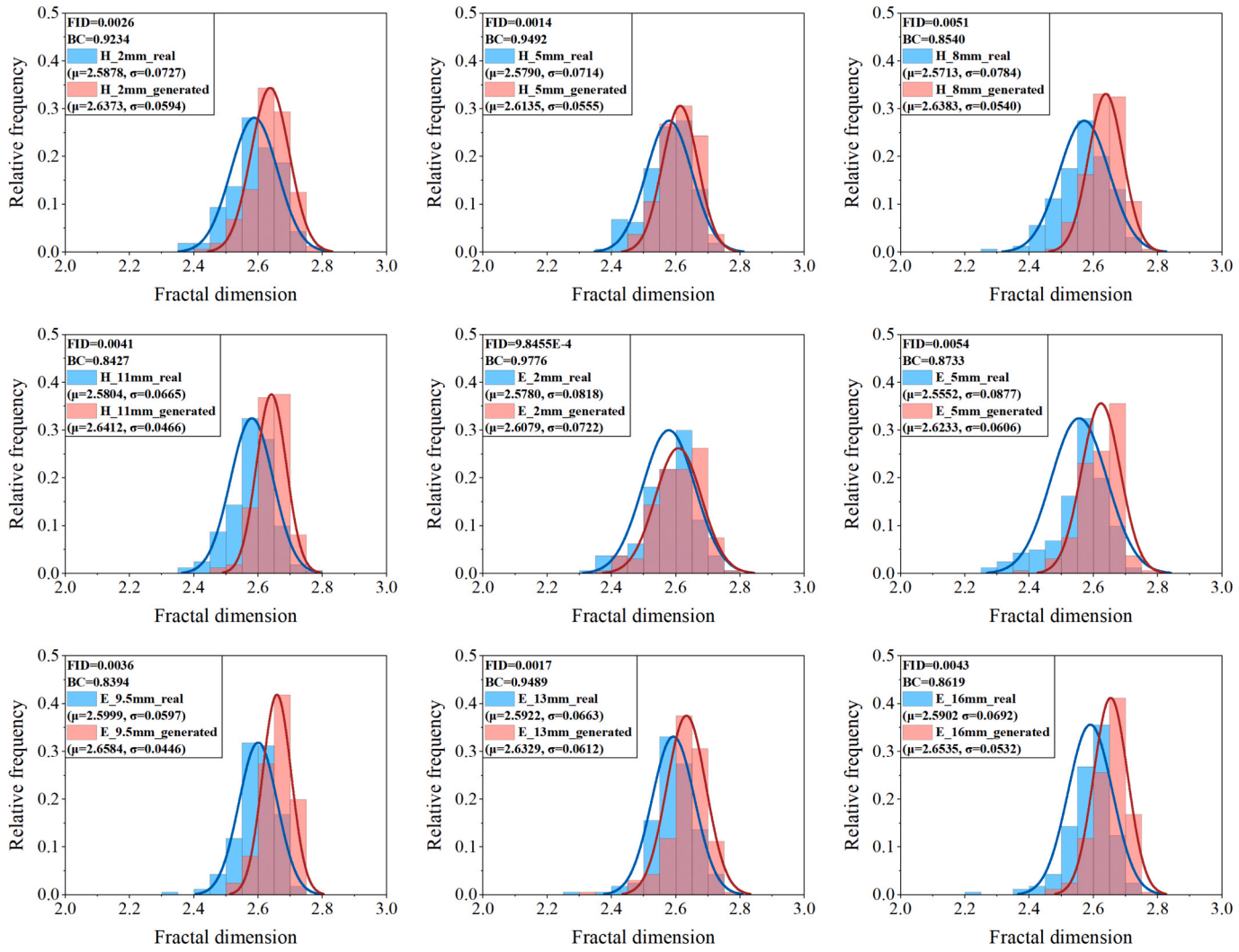


Fig. 22. Comparison of fractal dimension for virtual aggregates from different grain size databases.

Table 9
Difference in Bhattacharyya coefficient for database lightweight.

	H_2mm	H_5mm	H_8mm	H_11mm	E_2mm	E_5mm	E_9.5 mm	E_13mm	E_16mm
True sphericity	12.81 %	13.19 %	9.73 %	7.38 %	2.26 %	9.22 %	13.38 %	15.67 %	16.74 %
Angularity index	1.64 %	3.46 %	2.90 %	2.21 %	2.95 %	2.75 %	0.25 %	2.97 %	6.68 %
Fractal dimension	6.49 %	3.88 %	13.52 %	14.67 %	1.00 %	11.57 %	15.00 %	3.91 %	12.72 %

aggregates were compared using FID and BC. The aggregates augmented using DC-WGAN-div had the best quality in terms of true sphericity, angularity index, and fractal dimension. The FID values for the three indicators were 0.0013, 217.51, and 4.1097e-4, while the BC values were 0.9210, 0.9836, and 0.9875.

- The aggregate packing and SGC simulations demonstrated that the augmented aggregate databases reproduced the mechanical responses of asphalt mixtures with high fidelity. Packing density values from the augmented mixtures deviated by less than 2 % from those of the real mixtures, while the gyration–height fittings achieved coefficients of determination (R^2) above 0.992, with compaction exponents ranging narrowly from –0.0285 to –0.0298. Among the augmentation methods, DC-WGAN-GP, DC-CT-WGAN, and DC-WGAN-div exhibited the strongest consistency with the real mixtures, providing the most reliable basis for DEM-based mechanical response evaluation.
- The database lightweight had a certain impact on the quality of the augmented aggregates. The change in data volume affected true sphericity the most, followed by fractal dimension, and had the least effect on the angularity index. Across different cases, the maximum and minimum differences in BC values were 16.74 % and 2.26 % for true sphericity, 14.67 % and 1.00 % for fractal dimension, and 6.68 % and 0.25 % for the angularity index, respectively. However, the database lightweight offered a significant advantage in model training time, requiring approximately one-tenth of the time needed for the full database.

The above results showed that, with the assistance of stereo vision and DC-WGANs, it is feasible to establish and augment virtual aggregate databases that exhibit realistic and diverse morphological characteristics. However, some limitations remained in this study. Although the database lightweight approach can speed up model training, it also sacrifices the quality of the augmented data. Therefore, balancing model

training speed and the quality of augmented data remains a worthwhile challenge. In addition, this study focused solely on database lightweight and employed only WGAN variants. Future research will explore two important directions: the effect of model structure lightweight on augmentation performance, and the applicability of other GAN architectures, such as BigGAN [55] and f-GAN [56], to further improve the accuracy and generalizability of the generated models. Furthermore, by providing a more comprehensive and morphologically diverse virtual aggregates, the augmented database can be directly integrated into numerical simulations of asphalt mixtures, improving the predictive accuracy of pavement mechanical behavior and durability performance, thereby enhancing the practical utility of the research for design and maintenance decision-making. Beyond asphalt mixture applications, the proposed framework offers a flexible and scalable approach that can be generalized to other engineering contexts requiring augmented yet realistic 3D representations of heterogeneous materials. This highlights the broader potential of AI-based data augmentation to support high-fidelity modeling, accelerate simulation workflows, and facilitate data-driven decision-making across various civil and materials engineering applications.

CRediT authorship contribution statement

Dong Feng: Writing – original draft, Visualization, Methodology, Investigation, Conceptualization. **Jinchao Guan:** Writing – review & editing, Validation. **Chaoliang Fu:** Writing – review & editing, Writing – original draft, Validation, Formal analysis, Supervision. **Yuanyuan Hu:** Writing – review & editing. **Frederic Otto:** Writing – review & editing, Methodology. **Alvaro García Hernandez:** Writing – review & editing, Validation, Supervision. **Pengfei Liu:** Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research is supported by the China Scholarship Council (CSC, No. 202308080042) and the German Research Foundation (LI 3613/3-1, Project ID 528307766).

Appendix A

In this study, sample data, code, and training models are provided in an open-source format and can be accessed via the following link: https://github.com/ffd960131/DC-WGANs-aggregate. Fig. A.1 shows the project structure of this study. Among them, main.py is used to call model training and data generation; model.py contains the generator and discriminator structures; param.py is used to set hyperparameters; tester.py is used to test the model; trainer.py is used to train the model; utils.py contains some utility functions.

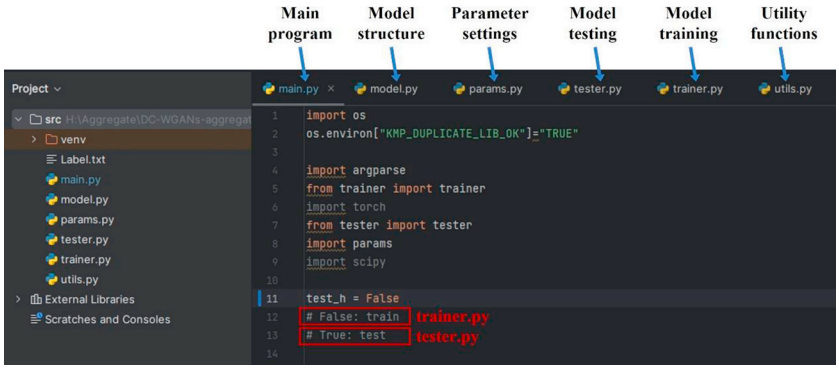


Fig. A1. Project structure.

Data availability

Data will be made available on request.

References

[1] Y. Sun, C. Du, H. Gong, Y. Li, J. Chen, Effect of temperature field on damage initiation in asphalt pavement: a microstructure-based multiscale finite element method, Mech. Mater. 144 (2020) 103367, https://doi.org/10.1016/j.mechmat.2020.103367.

[2] M. Nitka, J. Tejchman, Comparative DEM calculations of fracture process in concrete considering real angular and artificial spherical aggregates, Eng. Fract. Mech. 239 (2020) 107309, https://doi.org/10.1016/j.engfracmech.2020.107309.

[3] J. Suchorzewski, M. Nitka, Size effect at aggregate level in microCT scans and DEM simulation – splitting tensile test of concrete, Eng. Fract. Mech. 264 (2022) 108357, https://doi.org/10.1016/j.engfracmech.2022.108357.

[4] Y. Liu, J. Xie, Y. Zhang, D. Wei, K. Li, J. Song, Z. Dai, Study on the micromechanical response during permanent deformation of asphalt mixtures by discrete element modeling with real aggregate morphology, Constr. Build. Mater. 392 (2023) 131778, https://doi.org/10.1016/j.conbuildmat.2023.131778.

[5] X. Wei, Y. Sun, H. Gong, J. Chen, Determining appropriate level of detail of 3D aggregate model for accurate and efficient mesoscopic simulation of asphalt concrete, Road Mater. Pavement Des. 25 (2024) 2729–2746, https://doi.org/10.1080/14680629.2024.2329457.

[6] M. Wang, X. Yang, W. Wang, Establishing a 3D aggregates database from X-ray CT scans of bulk concrete, Constr. Build. Mater. 315 (2022) 125740, https://doi.org/10.1016/j.conbuildmat.2021.125740.

[7] C. Jin, X. Yang, Z. You, K. Liu, Aggregate shape characterization using virtual measurement of three-dimensional solid models constructed from X-Ray CT images of aggregates, J. Mater. Civ. Eng. 30 (2018) 4018026, https://doi.org/10.1061/(ASCE)MT.1943-5533.0002210.

[8] H.I. Ozturk, I. Rashidzade, A photogrammetry based method for determination of 3D morphological indices of coarse aggregates, Constr. Build. Mater. 262 (2020) 120794, https://doi.org/10.1016/j.conbuildmat.2020.120794.

[9] L. Zhao, S. Zhang, D. Huang, X. Wang, Y. Zhang, 3D shape quantification and random packing simulation of rock aggregates using photogrammetry-based reconstruction and discrete element method, Constr. Build. Mater. 262 (2020) 119986, https://doi.org/10.1016/j.conbuildmat.2020.119986.

[10] H. Wang, C. Wang, Y. Bu, Z. You, X. Yang, M. Oeser, Correlate aggregate angularity characteristics to the skid resistance of asphalt pavement based on image analysis technology, Constr. Build. Mater. 242 (2020) 118150, https://doi.org/10.1016/j.conbuildmat.2020.118150.

[11] B. Rajan, D. Singh, Understanding influence of crushers on shape characteristics of fine aggregates based on digital image and conventional techniques, Constr. Build. Mater. 150 (2017) 833–843, https://doi.org/10.1016/j.conbuildmat.2017.06.058.
[12] K. Li, Q. Liu, Y. Tian, C. Du, Z. Xu, The consequences of dimension reduction for open graded friction course (OGFC) asphalt mixtures: morphological characteristics and finite element model (FEM) simulation, Buildings 14 (2024) 545, https://doi.org/10.3390/buildings14020545.
[13] W. Wang, L. Wang, Review on design, characterization, and prediction of performance for asphalt materials and asphalt pavement using multi-scale numerical simulation, Materials 17 (2024) 778. doi: 10.3390/ma17040778.
[14] B. Suhr, W.A. Skipper, R. Lewis, K. Six, Shape analysis of railway ballast stones: curvature-based calculation of particle angularity, Sci. Rep. 10 (2020) 6045, https://doi.org/10.1038/s41598-020-62827-w.
[15] Y. Liu, F. Gong, Z. You, H. Wang, Aggregate morphological characterization with 3D optical scanner versus X-ray computed tomography, J. Mater. Civ. Eng. 30 (2018) 4017248, https://doi.org/10.1061/(ASCE)MT.1943-5533.0002091.
[16] C. Wang, F. Li, Q. Liu, H. Wang, P. Benmoussa, S. Jeschke, M. Oeser, Establishment and extension of digital aggregate database using auxiliary classifier Wasserstein GAN with gradient penalty, Constr. Build. Mater. 300 (2021) 124217, https://doi.org/10.1016/j.conbuildmat.2021.124217.
[17] J. Taheri-Shakib, A. Al-Mayah, A review of microstructure characterization of asphalt mixtures using computed tomography imaging: prospects for properties and phase determination, Constr. Build. Mater. 385 (2023) 131419, https://doi.org/10.1016/j.conbuildmat.2023.131419.
[18] M. Alawneh, H. Soliman, Using imaging techniques to analyze the microstructure of asphalt concrete mixtures: literature review, Appl. Sci. 13 (2023) 7813, https://doi.org/10.3390/app13137813.
[19] J. Liu, F. Yu, T. Funkhouser, Interactive 3D Modeling with a Generative Adversarial Network, in: 2017 International Conference on 3D Vision (3DV), 2017, pp. 126–134. Doi: 10.1109/3DV.2017.00024.
[20] J. Wu, C. Zhang, T. Xue, W.T. Freeman, J.B. Tenenbaum, Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling, in: NIPS'16 Proceedings of the 30th International Conference on Neural Information Processing Systems, 2016, pp. 82–90. Doi: 10.48550/arXiv.1610.07584.
[21] V.V. Kniaz, F. Remondino, V.A. Knyaz, Generative adversarial networks for single photo 3D reconstruction, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLII-2/W9 (2019) 403–408. Doi: 10.5194/isprs-archives-xlii-2-w9-403-2019.
[22] H. Kato, T. Harada, Learning View Priors for Single-View 3D Reconstruction, 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2019) 9778–9787. Doi: 10.48550/arXiv.1811.10719.
[23] I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative ADVERSARIAL networks, Adv. Neural Inf. Proces. Syst. 27 (2014) 2672–2680, https://doi.org/10.48550/arXiv.1406.2661.
[24] Q.-Y. Zhou, J. Park, V. Koltun, Fast Global Registration, in: Computer Vision – ECCV 2016, 2016, pp. 766–782. Doi: 10.1007/978-3-319-46475-6_47.
[25] D. Marius Muja, Lowe, fast approximate nearest neighbors with automatic algorithm configuration, Int. Conf. Comput. Vision Theory Appl. (2009), https://doi.org/10.5220/0001787803310340.
[26] L. Weng, From GAN to WGAN, arXiv preprint (2019). Doi: 10.48550/arXiv.1904.08994.
[27] M. Arjovsky, S. Chintala, L. Bottou, Wasserstein generative adversarial networks, in: Proceedings of the 34th International Conference on Machine Learning, 2017, pp. 214–223. Doi: 10.48550/arXiv.1701.07875.
[28] C. Villani, Optimal transport: old and new, Springer Science & Business Media, Berlin, 2008. ISBN: 978-3-540-71049-3.
[29] X. Guo, J. Hong, T. Lin, N. Yang, Relaxed Wasserstein with Applications to GANs, in: ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2021. Doi: 10.1109/icassp39728.2021.9414454.
[30] C. Zhou, J. Zhang, J. Liu, Lp-WGAN: Using Lp-norm normalization to stabilize Wasserstein generative adversarial networks, Knowl.-Based Syst. 161 (2018) 415–424. Doi: 10.1016/j.knosys.2018.08.004.
[31] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, A. Courville, Improved training of wasserstein GANs, Adv. Neural Inf. Proces. Syst. 30 (2017), https://doi.org/10.48550/arXiv.1704.00028.
[32] X. Wei, B. Gong, Z. Liu, W. Lu, L. Wang, Improving the improved training of Wasserstein GANs: a consistency term and its dual effect, International Conference on Learning Representations (ICLR) (2018), https://doi.org/10.48550/arXiv.1803.01541.
[33] J. Wu, Z. Huang, J. Thoma, D. Acharya, L. van Gool, Wasserstein Divergence for GANs, in: Proceedings of the European conference on computer vision (ECCV) (2018) 653–668, Doi: 10.48550/arXiv.1712.01026.
[34] A. Radford, L. Metz, S. Chintala, Unsupervised representation learning with deep convolutional generative adversarial networks, International Conference on Learning Representations (ICLR) (2016), https://doi.org/10.48550/arXiv.1511.06434.
[35] M. Mirza, S. Osindero, Conditional Generative Adversarial Nets, arXiv preprint, 2014, Doi: 10.48550/arXiv.1411.1784.
[36] T. Karras, S. Laine, T. Aila, A style-based generator architecture for generative adversarial networks, IEEE Trans. Pattern Anal. Mach. Intell. 43 (2021) 4217–4228, https://doi.org/10.1109/TPAMI.2020.2970919.
[37] S. Ioffe, C. Szegedy, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, in: ICML'15: Proceedings of the 32nd International Conference on International Conference on Machine Learning, 2015, pp. 448–456. Doi: 10.48550/arXiv.1502.03167.
[38] T. Tieleman, G. Hinton, Lecture 6.5-rmsprop: divide the gradient by a running average of its recent magnitude. COURSERA: Neural Networks for, Mach. Learn. 4 (2012) 26–31. https://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf.
[39] D.P. Kingma, J. Ba, Adam: a method for stochastic optimization, International Conference on Learning Representations (ICLR) (2014), https://doi.org/10.48550/arXiv.1412.6980.
[40] D. Dowson, B. Landau, The Fréchet distance between multivariate normal distributions, J. Multivar. Anal. 12 (1982) 450–455, https://doi.org/10.1016/0047-259X(82)90077-X.
[41] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, S. Hochreiter, GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium, in: 31st Conference on Neural Information Processing Systems (NIPS 2017), 2017. Doi: 10.48550/arXiv.1706.08500.
[42] W. Huang, B. Lai, W. Xu, Z. Tu, 3D Volumetric Modeling with Introspective Neural Networks, AAAI 33 (2019) 8481–8488, https://doi.org/10.1609/aaai.v33i01.33018481.
[43] H. Wadell, Volume, shape, and roundness of quartz particles, J. Geol. 43 (1935) 250–280, https://doi.org/10.1086/624298.
[44] C. Rao, E. Tutumluer, I.T. Kim, Quantification of coarse aggregate angularity based on image analysis, transportation research record, J. Transp. Res. Board 1787 (2002) 117–124, https://doi.org/10.3141/1787-13.
[45] F. Hausdorff, Dimension und äußeres Maß, Math. Ann. 79 (1918) 157–179, https://doi.org/10.1007/BF01457179.
[46] A. Bhattacharyya, On a Measure of Divergence between Two Multinomial Populations, Sankhyā: The Indian Journal of Statistics (1933-1960) 7 (1946) 401–406.
[47] P.A. Cundall, O.D.L. Strack, A discrete numerical model for granular assemblies, Geotechnique 29 (1979) 47–65, https://doi.org/10.1680/geot.1979.29.1.47.
[48] P.A. Cundall, O. STRACK, Modeling of Microscopic Mechanisms in Granular Material 7 137–149. Doi: 10.1016/B978-0-444-42192-0.50018-9.
[49] R.D. Mindlin, H. Deresiewicz, Elastic spheres in contact under varying oblique forces, J. Appl. Mech. 20 (1953) 327–344, https://doi.org/10.1115/1.4010702.
[50] X. Zhou, S. Chen, D. Ge, D. Jin, Z. You, Investigation of asphalt mixture internal structure consistency in accelerated discrete element models, Constr. Build. Mater. 244 (2020) 118272, https://doi.org/10.1016/j.conbuildmat.2020.118272.
[51] K. Zhang, W. Yu, D. Li, D. Zou, S. Zhang, Measurement and simulation validation of numerical model parameters of fresh concrete, Sci. Eng. Compos. Mater. 28 (2021) 437–452, https://doi.org/10.1515/secm-2021-0042.
[52] L. Gao, Y. Zhang, Y. Liu, Z. Wang, X. Ji, Study on the cracking behavior of asphalt mixture by discrete element modeling with real aggregate morphology, Constr. Build. Mater. 368 (2023) 130406, https://doi.org/10.1016/j.conbuildmat.2023.130406.
[53] Y. Liu, Y. Zhou, X. Liu, F. Dong, C. Wang, Z. Wang, Wasserstein GAN-based small-sample augmentation for new-generation artificial intelligence: a case study of cancer-staging data in biology, Engineering 5 (2019) 156–163, https://doi.org/10.1016/j.eng.2018.11.018.
[54] Y. Wei, S. Xu, B. Kang, S. Hoque, Generating training images with different angles by GAN for improving grocery product image recognition, Neurocomputing 488 (2022) 694–705, https://doi.org/10.1016/j.neucom.2021.11.080.
[55] A. Brock, J. Donahue, K. Simonyan, Large scale GAN Training for high fidelity natural image synthesis, ICLR (2019), https://doi.org/10.48550/arXiv.1809.11096.
[56] S. Nowozin, B. Cseke, R. Tomioka, f-GAN: Training generative neural samplers using variational divergence minimization, in: NIPS'16: Proceedings of the 30th International Conference on Neural Information Processing Systems, 2016, pp. 271–279, Doi: 10.48550/arXiv.1606.00709.