Workflow based on GANs and CNNs towards a digital twin for the 3D morphological characterization of latex aggregates

Léo Théodon, Carole Coufort-Saudejaud, Johan Debayle

Powder Technology 465 (2025) 121286

Contents lists available at ScienceDirect

Powder Technology

journal homepage: www.journals.elsevier.com/powder-technology

Workflow based on GANs and CNNs towards a digital twin for the 3D morphological characterization of latex aggregates

L. Théodon [a], C. Coufort-Saudejaud [b,*], J. Debayle [a]

[a] MINES Saint-Etienne, CNRS, UMR 5307 LGF, Centre SPIN, Saint-Etienne, France
[b] Laboratoire de Génie Chimique, Université de Toulouse, CNRS, INPT, UPS, Toulouse, France

HIGHLIGHTS

- Novel deep learning workflow for 3D characterization from 2D images.
- Precise segmentation and reconstruction of latex aggregates from in-situ imaging.
- Digital twin validated on synthetic data with 5% error on key features.
- Demonstrated feasibility on real industrial latex aggregate data.
- Adaptable framework for real-time particle monitoring in various industries.

ARTICLE INFO

Keywords:
Aggregate
Convolutional Neural Network (CNN)
Deep learning
Generative Adversarial Network (GAN)
Image analysis
Morphological characterization

ABSTRACT

This paper presents a workflow for estimating the 3D morphological characteristics of latex aggregates from 2D in-situ images using deep learning and stochastic geometry models. The method includes automatic image segmentation using a Convolutional Neural Network (CNN), 3D object generation using a Generative Adversarial Network (GAN), and estimation of 3D characteristics. Validation with synthetic datasets shows effective size, shape, and texture characterization, with the Mean Absolute Percentage Error (MAPE) for morphological characteristics of generated objects being around 5% at most. Application to real in-situ images demonstrates feasibility and consistency with experimental observations, successfully generating a digital twin of the latex aggregate population. The method's flexibility and efficiency make it suitable for real-time industrial applications, offering potential for process monitoring and quality control. Future work will focus on enhancing model performance and adapting to different particle types for broader applicability in various industrial settings.

1. Introduction

The morphological characterization of aggregates, agglomerates, and granular objects plays a crucial role in a wide range of scientific and industrial applications. From the pharmaceutical industry, where particle shape and size directly impact drug delivery and efficacy [1,2], to the field of civil engineering, where aggregate morphology influences the strength and durability of construction materials [3], understanding the complex relationship between morphology and material properties is of utmost importance. In the realm of food science, the size, shape, and texture of food particles not only affect the sensory attributes of products but also their processing behavior and shelf life [4,5]. Similarly, in the chemical and petrochemical industries, the morphology of catalysts and adsorbents determines their performance and efficiency in various processes [6]. Beyond traditional applications, recent advances in artificial intelligence have revolutionized materials characterization across diverse domains, from ultra-high-performance concrete composites to granular systems, demonstrating the universal importance of morphological analysis in materials science [7,8].

Traditionally, the morphological characterization of aggregates has relied on manual methods, such as sieve analysis and microscopy, which are time-consuming, labor-intensive, and often subjective [9]. However, with the advent of digital imaging technologies and the increasing computational power, automated image analysis techniques have emerged as a promising alternative [10]. These techniques enable the rapid and objective measurement of various morphological features, such as size, shape, and texture, from two-dimensional (2D) images of aggregates [11,12].

In recent years, machine learning algorithms, particularly deep learning architectures like Convolutional Neural Networks (CNNs), have revolutionized the field of image analysis [13]. This technological

* Corresponding author.
E-mail addresses: leo.theodon@orano.group (L. Théodon), carole.saudejaud@toulouse-inp.fr (C. Coufort-Saudejaud), debayle@emse.fr (J. Debayle).

https://doi.org/10.1016/j.powtec.2025.121286
Received 30 April 2025; Received in revised form 9 June 2025; Accepted 18 June 2025
Available online 1 July 2025

advancement is particularly crucial as industries increasingly demand real-time, non-destructive characterization methods that can be integrated into automated production lines. CNNs have demonstrated remarkable performance in tasks such as object detection, segmentation, and classification, making them well-suited for the morphological characterization of aggregates [14,15]. For instance, Zhang et al. [16] employed a CNN-based approach to segment and classify sand particles from 2D images, achieving high accuracy and efficiency compared to traditional methods. Similarly, Wu et al. [17] used a CNN to estimate the size distribution of iron ore pellets from images captured during the pelletizing process, enabling real-time monitoring and control of the process.

Despite the success of CNN-based methods in characterizing aggregates from 2D images, they often fail to capture the three-dimensional (3D) nature of these objects. The 2D morphological characteristics extracted from images are inherently limited and may not provide a complete understanding of the aggregate's true morphology [18]. To address this limitation, researchers have explored the use of 3D imaging techniques, such as X-ray computed tomography (CT) and laser scanning, to obtain volumetric data of aggregates [19–21]. However, these techniques are often expensive, time-consuming, and not suitable for in-situ characterization of aggregates in industrial settings, although in-situ imaging is crucial since it allows for real-time monitoring of particle aggregation without disrupting the sample, ensuring that the physical characteristics of the aggregates, both in 2D and 3D, are accurately captured without deformation or preferential orientation. This accuracy is vital for understanding and controlling the aggregation process in various industrial applications [22].

An alternative approach to bridging the gap between 2D and 3D morphological characterization is the concept of "digital twins" [23]. A digital twin is a virtual representation of a physical object or system that can be used for simulation, optimization, and prediction purposes. In the context of aggregate characterization, a digital twin would be a 3D model of the aggregate population that accurately captures its morphological characteristics and can be used to estimate the properties and behavior of the real aggregate system [24,25].

This paper proposes a novel workflow to estimate the 3D morphological characteristics of latex aggregates from 2D in-situ images by combining deep learning techniques and stochastic geometry models [26,27]. The process begins with the automatic segmentation of in-situ images using instance segmentation models. Next, 3D voxelized objects are generated from the segmentation masks through a deep learning generative model. Finally, the 3D morphological characteristics of the aggregates are estimated from these generated objects. The ultimate goal is to create a digital twin of the real aggregate population for applications in process monitoring, quality control, and product design.

The proposed workflow is validated using synthetic aggregate populations generated by a stochastic geometrical model and photorealistic synthetic in-situ images of latex aggregates. The synthetic images are used to train a deep learning model for instance segmentation, namely Mask R-CNN [28]. The segmentation model is then applied to the synthetic in-situ images, and the resulting segmentation masks are used as input for a Generative Adversarial Network (GAN) [29] to generate 3D voxelized objects. The morphological characteristics of the generated objects are compared to the ground truth to evaluate the performance of the method.

This work contributes to the field of aggregate characterization by providing a comprehensive framework for estimating the 3D morphological properties of aggregates from 2D images, leveraging the power of deep learning and stochastic geometry models. The proposed method has the potential to be applied in various industrial settings, enabling better process control, optimization, and product development. Furthermore, the concept of digital twins introduced in this paper opens up new possibilities for the virtual exploration and manipulation of aggregate systems, paving the way for more efficient and sustainable manufacturing practices [30,31].

The remainder of this paper is organized as follows. Section 2 describes the experimental setup used for capturing the in-situ images and outlines the various morphological characteristics studied. Section 3 elaborates on the proposed methodology and the overall structure of the study. Section 4 details the segmentation model and the procedure for generating synthetic images. Section 5 covers the training process of the Generative Adversarial Network. Section 6 validates the entire workflow using synthetic datasets. Section 7 demonstrates the application of the workflow to real images, focusing on the creation of a digital twin of a latex aggregate population. Section 8 discusses the advantages and limitations of the proposed method. Finally, Section 9 presents the conclusions and suggests directions for future research.

2. Data acquisition & characterization

2.1. Data acquisition

To study the aggregation process of D-480 latex particles, a series of in-situ images were captured using a specialized experimental setup as shown by Fig. 1. The main components of this setup include a stirred tank, a camera, a calibration grid, and a lighting system. The stirred tank, with a capacity of 20 L, was used to mix the latex suspension and induce aggregation by injecting a salt solution. The tank was equipped with a curved three-blade impeller and a two-finger baffle to ensure proper mixing under turbulent conditions.

Images were acquired using a *pco.panda 26 DS sCMOS* camera with a resolution of 5120×5120 pixels, covering an area of 4.6×4.6 cm^2 with a resolution of 8.9 μm/pixel. The camera was positioned in front of the tank, focusing on a vertical plane located 5.5 cm from the tank wall and centered on the stirring shaft. A 2D calibration grid, placed in the focal plane, was used to convert camera pixels to metric units. Illumination was provided by a continuous intensity 462 nm blue LED light source, ensuring consistent lighting conditions throughout the aggregation process.

Series of 2000 images were captured at different times during the aggregation process. The acquired images, as shown in Fig. 2, provide valuable information about the 2D and 3D characteristics of the aggregates without causing any deformation or introducing preferential orientation. It is assumed that within each series of images, the morphological characteristics of the aggregates do not evolve significantly due to the rapid acquisition of the images.

2.2. Morphological characterization

This article focuses on characterizing the morphology of aggregates at both the 2D level, using in-situ images, and the 3D level, using synthetic objects generated by a generative model. Morphometric characterization involves performing 2D or 3D geometric measurements to define various morphological characteristics, as presented in Table 1. These measurements allow for the quantification of different aspects of the objects' morphology, following nomenclature established by Théodon et al. [32]. The morphological characteristics can be categorized into three main groups:

- Size-related characteristics: These include measurements such as area and volume, which provide information about the overall size of the objects. In 2D, the area is calculated from the segmented images, while in 3D, the volume is determined from the voxelized representation of the objects.
- Shape-related characteristics: These characteristics describe the overall shape of the objects, such as elongation and aspect ratio. Elongation measures the degree to which an object is elongated or stretched, while aspect ratio compares the length of the object to its width. These characteristics are calculated using the principal axes of the objects in both 2D and 3D.

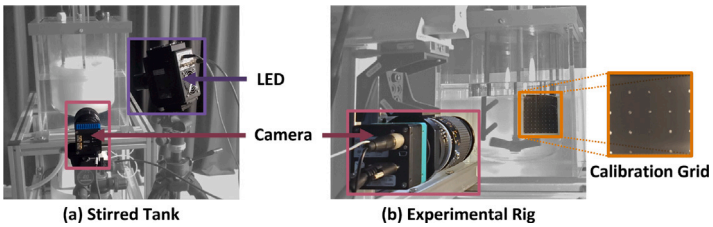


Fig. 1. Illustration of the experimental setup used to capture in-situ images of the D-480 latex aggregation process, showcasing the stirred tank, camera, LED light source, and the calibration target.

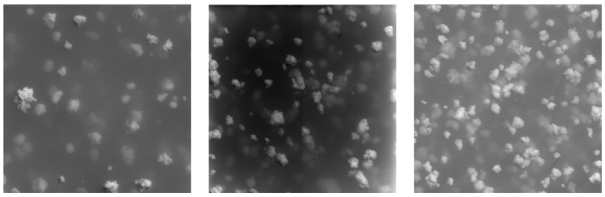


Fig. 2. Examples of in-situ images taken during the aggregation process after 3 h (a), after 7 h (b), and after 10 h (c).

Table 1
List of morphological characteristics used in this paper.

Parameters	Symbol	Definition and equation
2D & 3D parameters		
Feret diameter max.	F_{max}	Longest caliper (Feret) length
Feret diameter min.	F_{min}	Smallest caliper (Feret) length
Aspect ratio	AR	F_{min}/F_{max}
2D parameters		
Projected area	A	Area of the object
Convex area	A_c	Area of the convex hull
Perimeter	P	Length of the object outline
Equivalent circle diameter	ECD	$2\times\sqrt{A/\pi}$
Convexity	Co	A/A_c
Circularity	C	$4\pi\times A/P^2$
3D parameters		
Volume	V	Volume of the object
Convex volume	V_c	Volume of the convex hull
Volume	V_p	Volume of the closed pores
Surface area	S	Area of the object surface
Equivalent sphere diameter	ESD	$2\times\sqrt[3]{3\times V/(4\pi)}$
Elongation	e	Equivalent ellipsoid elongation
Solidity	SLD	V/V_c
Sphericity	Φ_S	$6\pi^2\times V/(\sqrt{\pi}S)^3$

- Angularity and texture-related characteristics: These characteristics, such as circularity and sphericity, provide information about the surface roughness and angularity of the objects. Circularity measures how close the object's shape is to a perfect circle in 2D, while sphericity measures how close the object's shape is to a perfect sphere in 3D. These characteristics are calculated using the perimeter and surface area of the objects in relation to their size.

The morphological characterization of aggregates using these measurements is crucial for understanding the evolution of the aggregation process and for validating the performance of the generative model in creating realistic 3D representations of the aggregates.

3. Methodology

The main objective of this article is to propose a method for estimating the 3D characteristics of latex aggregates from in-situ images. The proposed method consists of three steps:

(a) In-situ images are automatically segmented using instance segmentation deep-learning models.
(b) For each object detected in the in-situ images, a 3D voxelized object is generated from the corresponding segmentation mask using a deep learning generative model.
(c) The 3D morphological characteristics of the aggregates visible in the in-situ images can be estimated from the characteristics measured on the objects generated from the segmentation masks.

Consequently, the following three sections are structured as follows:

1. Two synthetic aggregate populations $\mathcal{A}$ and $\mathcal{B}$ are generated and then used to compose photorealistic synthetic images. These images will be used to train an automatic segmentation model, specifically a Mask R-CNN with a ResNet101 backbone.
2. A GAN-type model is trained on independent data. This model allows the generation of 3D voxelized objects from a 2D grayscale image of an aggregate.
3. The generator is used to generate object populations $\mathcal{A}'$ and $\mathcal{B}'$ from the data obtained by automatic segmentation of the synthetic images. The distributions of morphological characteristics of populations $\mathcal{A}$ and $\mathcal{A}'$ on one hand, and $\mathcal{B}$ and $\mathcal{B}'$ on the other hand, are compared to evaluate the performance of the method.

To ensure data independence and not attribute any potential good results to over-fitting of the generative model, two different models will be used to generate the synthetic images and to train the GAN.

This study specifically addresses compact aggregates characteristic of the D-480 latex, representing a well-defined morphological class relevant to numerous industrial applications, rather than attempting universal coverage of all possible aggregate geometries.

4. Image segmentation

4.1. Overview

This section focuses on the creation of a training dataset consisting of synthetic aggregates, which will be used to generate photorealistic images that closely resemble the in-situ images captured during the real aggregation process. The synthetic images will be used to train and evaluate the performance of the selected image segmentation model, which will be presented in this section.

To create the training dataset, two synthetic aggregate populations, denoted as $\mathcal{A}$ and $\mathcal{B}$, are generated. The synthetic aggregates will then be used to compose photorealistic images that simulate the in-situ imaging conditions, incorporating various challenges present in the real images, such as object overlap, varying degrees of focus, and significant variations in image brightness, contrast, and gray balance.

Several state-of-the-art instance segmentation architectures were considered for the segmentation task, as they are better suited for handling overlapping objects compared to direct segmentation methods or semantic segmentation [33]. The architectures evaluated include:

- Mask R-CNN (He et al. [28,34]) with ResNet50 and ResNet101 backbones.

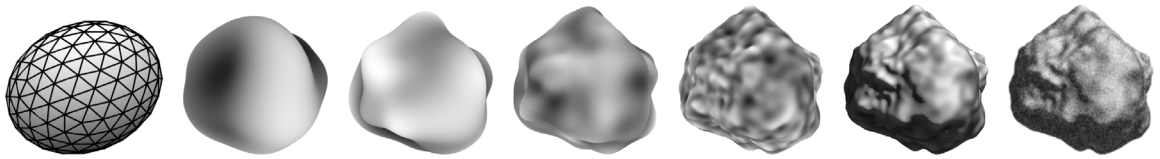


Fig. 3. Sequential generation of an aggregate using an extended version of the SPHERE model: From the basic shape (a) to the final rendering (g) with progressive deformations using 1 to 4 random Gaussian fields (b–e), enhanced by lighting and shading effects (f), and finished with noise and blur effects (g).

- YOLOv8 (Jocher et al. [35]), known for its fast execution and ability to accurately detect objects.
- SAM (Kirillov et al. [36]), a versatile multimodal model developed by Meta AI.

After evaluating the performance of these models on the synthetic dataset, the Mask R-CNN architecture with a ResNet101 backbone was selected as it yielded the best results. The selection was based on both classical machine learning metrics, such as accuracy and precision, and the similarity between the distributions of 2D morphological characteristics measured on the segmentation masks and those of the ground truth. It was further validated through comprehensive comparative analysis [37], which demonstrated that while different segmentation architectures show varying individual performance metrics, their impact on final 3D morphological estimation remains statistically equivalent, confirming the robustness of our approach to segmentation variations.

The following subsections will provide more details on the generation of the synthetic aggregate populations, the creation of photorealistic images, and the training and evaluation of the selected Mask R-CNN model.

4.2. Construction of the training dataset

4.2.1. Synthetic images generation

The SPHERE model [26], a stochastic geometric model that generates aggregates or granular objects by deforming star-shaped objects using random Gaussian fields, is used to create synthetic images that closely resemble the real in-situ images shown in Fig. 2. In this study, an extended version of the SPHERE model is employed, which uses four Gaussian random fields (RGFs) to deform the mesh of an ellipsoid, resulting in more realistic-looking objects. Fig. 3 illustrates the process of creating an aggregate that will be used to compose a synthetic image.

The synthetic images are generated using the following steps:

1. For each image, a number n of aggregates is generated using the four-RGF version of the SPHERE model, where n is determined by a Poisson distribution with parameter λ.
2. The n aggregates are spatially distributed uniformly in the final image, with successive applications of Gaussian blur and increasing brightness to simulate a depth of field effect. The image generator thus follows a Boolean model [38].
3. Post-processing techniques, such as the application of additional Gaussian noise and Gaussian blur, are used to obtain a more realistic rendering.

To create the training set, two synthetic populations of aggregates, denoted as $\mathcal{A}$ and $\mathcal{B}$, are generated using the SPHERE model with different size distributions. Population $\mathcal{A}$ follows a monomodal size distribution, while population $\mathcal{B}$ follows a bimodal size distribution. These populations are then used to generate two subsets of synthetic images, each containing 2000 images: Dataset. 1, which resembles the in-situ images captured after 7 h (Fig. 2(b)), is created using population $\mathcal{A}$ with $\lambda = 80$, and Dataset. 2, which resembles the in-situ images captured after 10 h (Fig. 2(c)), is created using population $\mathcal{B}$ with $\lambda = 130$. Examples of synthetic images with different lighting conditions and aggregate concentrations are presented in Fig. 4. In total, approximately 460,000 aggregates are distributed over 4000 synthetic images, as illustrated in Fig. 5.

Table 2
Comparison of FID scores between synthetic datasets and real image sets captured at different times.

	Dataset. 1 (7 h)	Dataset. 2 (10 h)
FID	83.1	45.3

4.2.2. Validation of the training dataset

To ensure the quality of the training dataset, it is essential to generate synthetic images that closely resemble the real in-situ images. This involves carefully adjusting the parameters of the aggregate generation model and the 3D rendering engine. A metric is needed to evaluate the similarity between the synthetic and real images and to ensure that the results obtained from the training data are transferable to real images.

The Fréchet Inception Distance (FID), a measure commonly used in GAN models (Heusel et al. [39]) to assess the similarity between synthetic and real images, is chosen as the metric. FID can generally be considered a criterion for similarity in terms of human vision. It calculates the Fréchet distance (cf. Dowson and Landau [40]) between high-level features of real and generated images, extracted by a specific layer (the last layer) of the pre-trained Inception-v3 model (Szegedy et al. [41]). A lower FID score indicates a closer similarity between the feature distributions of the generated and real images, suggesting a strong resemblance of the synthetic images to the real ones.

The FID is calculated as follows:

1. All real and generated images are processed by the Inception-v3 model to extract the high-level features contained in the last layer.
2. The means μ and covariance matrices Σ of the features are calculated for both sets of images.
3. The Fréchet distance between the two sets of features is calculated using their means μ_k and covariance matrices Σ_k.

The explicit formula for FID is:

$$\mathrm{FID} = \|\mu_r - \mu_g\|^2 + \mathrm{Tr}\left(\Sigma_r + \Sigma_g - 2\sqrt{\Sigma_r \Sigma_g}\right) \quad (1)$$

where μ_r and μ_g are the mean feature vectors of the real and generated images, respectively, and Σ_r and Σ_g are the covariance matrices of the features of the real and generated images, respectively.

Table 2 shows the FID scores calculated by comparing Dataset. 1 and Dataset. 2 with the real images captured at 7 h (Fig. 2(b)) and 10 h (Fig. 2(c)) of the aggregation process, respectively. Both FID scores are below 100, which is generally considered satisfactory. However, FID scores can vary greatly depending on the context and type of images studied. Therefore, to evaluate the quality of the FID scores more objectively, the sets of real images at 7 h and 10 h were randomly divided into subsets of 1000 images. The average FID calculated between these subsets of 1000 real images at 7 h and 10 h was 66.6 and 26.7, respectively, which serves as a reference point for good FID performance for these two sets of images.

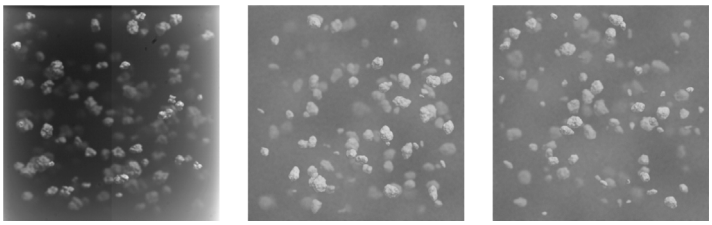
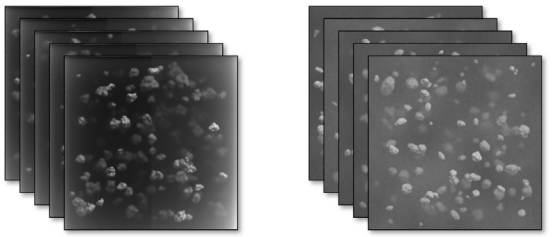


Fig. 4. Examples of synthetic images generated using the version of the SPHERE model with four random Gaussian fields.



(a) Dataset. 1, $\lambda_1 = 80$ (b) Dataset. 2, $\lambda_2 = 130$

Fig. 5. Illustration of the two sets of 2000 synthetic images generated using the SPHERE model.

Table 3
Summary of training parameters.

Parameter	Value
Dataset split	75%/12.5%
Epochs	50
Batch size	32
Learning rate	0.001

Fig. 6 illustrates the evolution of the average FID score obtained by comparing subsets of real images of increasing size. As the sets of real images contain 2000 images each, only subsets of up to 1000 images can be compared. It can be observed that the rate of decrease in the average FID scores slows down significantly as the subset size increases, suggesting that the FID scores are converging towards a stable value. The decreasing FID trend reflects statistical convergence as distribution estimates stabilize with increasing sample size. The FID scores obtained by comparing the sets of 2000 real images with 2000 synthetic images are close to this stable value, indicating a strong resemblance between the synthetic and real images. Thus, the global training set, consisting of a total of 4000 synthetic images (2000 images from Dataset. 1 and 2000 from Dataset. 2), is confirmed to have a strong resemblance to the real images, as evidenced by the low FID scores obtained. The parameters of the SPHERE model and image processing pipeline were systematically calibrated to minimize FID scores between synthetic and real images, with detailed methodology described in [37].

4.3. Model training

The training dataset consists of all images from Dataset. 1 and Dataset. 2, with the resolution reduced from 2048 × 2048 pixels to 1024 × 1024 pixels. This reduction aims to decrease the learning and prediction time while maintaining a high level of detail. The dataset of 4000 images is divided into three parts: 3000 images (75%) for training and 500 images (12.5%) each for the test and validation sets (in the same proportions for Dataset. 1 and Dataset.2). The key parameters are summarized in Table 3.

Furthermore, the segmentation model was initialized with weights from pre-training on the COCO dataset (Lin et al. [42]) to leverage its large diversity of images for improved feature extraction and to ensure good training conditions.

Table 4
Different metrics calculated for the segmentation model on the test sets of Dataset. 1 and Dataset. 2.

	AP	Mean IoU	Accuracy	Precision	Recall	F1 score	ARE
Dataset. 1	0.67	0.807	0.999	0.900	0.917	0.890	12%
Dataset. 2	0.76	0.853	0.999	0.900	0.953	0.918	33%

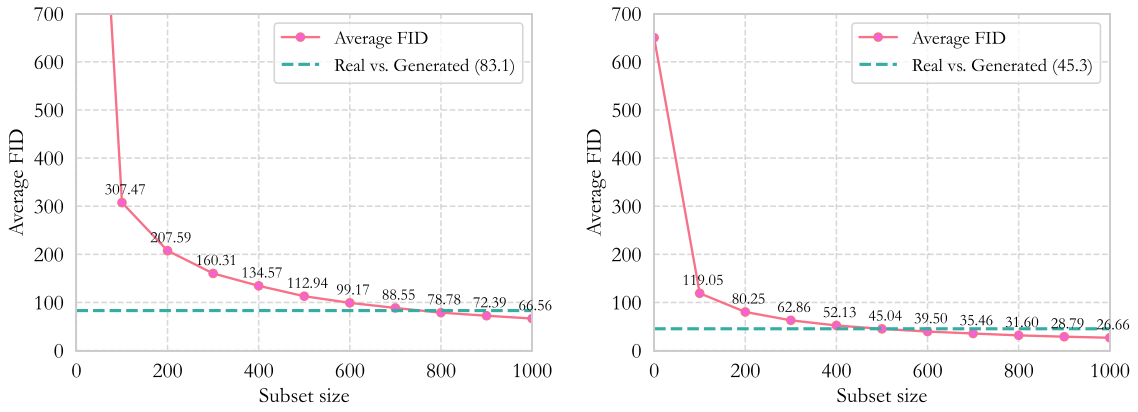
4.4. Performance evaluation

To evaluate the performance of the Mask R-CNN model with the ResNet101 backbone, the global test set is divided based on the origin of the images, distinguishing between Dataset. 1 and Dataset. 2. The model is evaluated separately on the two test datasets due to significant differences in how they were generated and the appearance of the images they contain. Dataset. 1 has an average of 80 objects per image, with very significant variations in brightness and contrast, while Dataset. 2 has an average of 130 objects, with significantly higher overall brightness and blur (Fig. 5). Additionally, the size distribution of aggregates in Dataset. 1, which is composed of the synthetic aggregates referred to as population $\mathcal{A}$, follows a beta law. In contrast, the size distribution of aggregates in Dataset. 2, consisting of the synthetic aggregates from population $\mathcal{B}$, is bimodal, resulting from a sum of beta laws. This distinction allows for evaluating the model's ability to accurately identify objects of different sizes within the same image. The usual performance measures, defined below, are then calculated, with the results detailed in Table 4.

- **Macro metrics**: In the case of single-class instance segmentation, as is the case here where the only class is *aggregate*, macro-metrics (accuracy, precision, recall, F1-score) calculated from a confusion matrix [43] evaluate the performance by considering each detected object individually and then averaging.
- **IoU** (*Intersection over Union*): A measure of the overlap between predicted segmentation masks and real objects, with higher IoU indicating better segmentation accuracy.
- **AP** (*Average Precision*): As defined by Gu et al. [44], this is a popular measure for evaluating the balance between precision and recall across different IoU thresholds, calculated in increments of 0.05, ranging from an IoU of 0.5 to 0.95.
- **ARE** (*Absolute Relative Error*): A measure of the accuracy of object detection by comparing the number of detected objects to the number of real objects, with a lower ARE indicating higher counting accuracy.

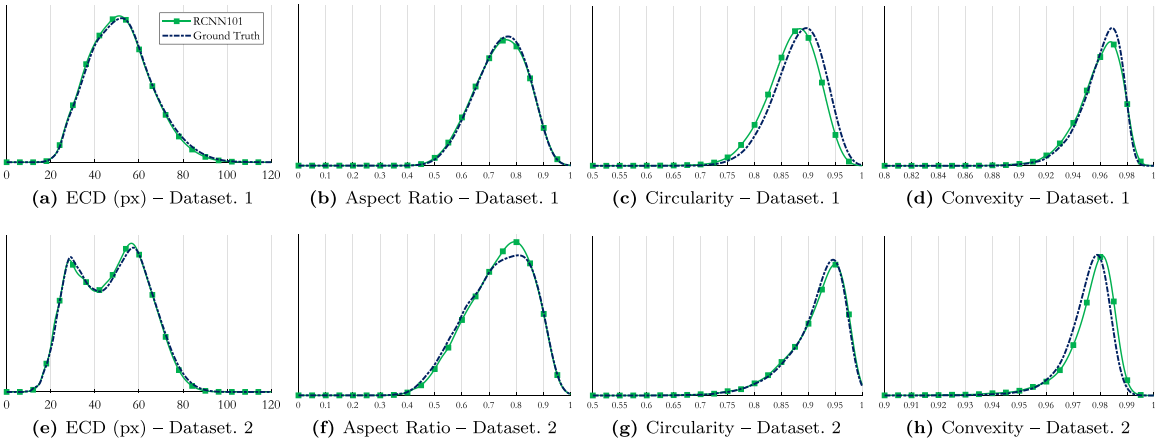
The results presented in Table 4 show that the segmentation model performs better on Dataset. 1 than on Dataset. 2. This difference can be attributed to the lower contrast and higher level of blur in the images of Dataset. 2, as well as the higher object density and overlap rate. Moreover, the absolute relative error (ARE) shows a lower detection rate for Dataset. 2, precisely due to the large number of objects overlapping each other and significant lens blur. Overall, the ARE and Fig. 8 highlight the effective object detection capabilities of the segmentation model.

In general, given the imbalance between the proportion of surface covered by objects and the background, accuracy proves to be a less



(a) Dataset. 1 (Population $\mathcal{A}$) (b) Dataset. 2 (Population $\mathcal{B}$)

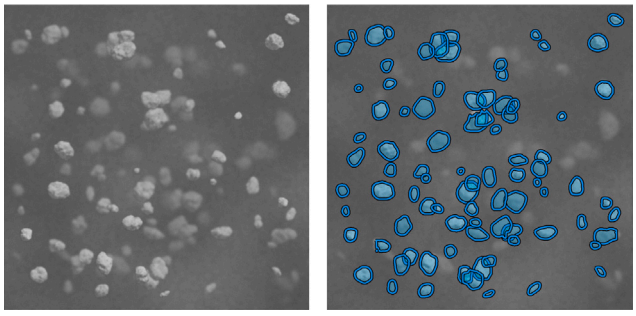
Fig. 6. Illustration of the evolution of mean FID values as a function of the size of real image subsets. The dotted line indicates the score obtained by comparing real images with synthetic images.



(a) ECD (px) – Dataset. 1 (b) Aspect Ratio – Dataset. 1 (c) Circularity – Dataset. 1 (d) Convexity – Dataset. 1

(e) ECD (px) – Dataset. 2 (f) Aspect Ratio – Dataset. 2 (g) Circularity – Dataset. 2 (h) Convexity – Dataset. 2

Fig. 7. Size (Equivalent Circular Diameter), shape (Aspect Ratio), and angularity or texture (Circularity and Convexity) distributions determined by the segmentation model on the test sets of Dataset. 1 and Dataset. 2.



(a) Original image (b) ResNet101

Fig. 8. Example of segmentation of an image from the test set of Dataset. 2 by the Mask R-CNN ResNet101 model, with object detection rates similar to those presented in the last column (ARE) of Table 4.

suitable measure for evaluating model performance. Nevertheless, the segmentation model turns out to be reliable when considering recall measures, which are more appropriate in this context. These metrics are significant because they are directly related to the quality and accuracy of the predicted segmentation masks.

However, in the context of morphological characterization of aggregates, the usual metrics, although essential in the field of deep learning (cf. Tian et al. [45]), are not necessarily suitable for evaluating the performance of the model. This is because these metrics do not directly assess the accuracy of the estimated morphological characteristics. For this reason, the distributions of morphological characteristics of the detected objects are calculated and compared to the ground truth.

The results presented in Fig. 7 demonstrate that the segmentation model performs reasonably well, particularly in terms of the estimated 2D morphological characteristic distributions, which closely match the ground truth for size (ECD) and shape (AR). However, some differences are observed for angularity and texture characteristics such as circularity and convexity. These discrepancies can be attributed to the intentionally low quality of the images and the relatively small size of the detected objects. In such cases, even a few pixels of difference in the segmentation masks can lead to noticeable variations in these characteristics [46,47].

Despite these minor limitations, the overall performance of the segmentation model is satisfactory, providing a solid foundation for the next stage of the analysis. Indeed, the instance segmentation architecture handles overlapping particles, and spatial homogeneity of aggregate distributions ensures that out-of-focus detections do not bias population-level characterization [37]. Building upon these results, the focus now shifts to the development of a generative model capable of transforming the segmentation masks into 3D voxelized objects. This approach aims to bridge the gap between the 2D segmentation outputs and the desired 3D morphological characterization of the aggregates, enabling a more comprehensive understanding of their properties and behavior.

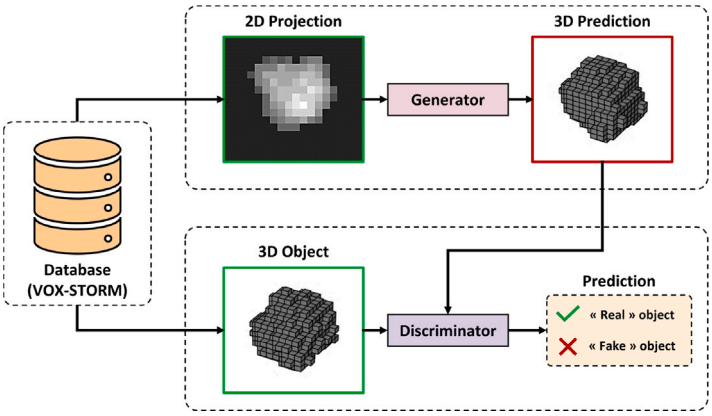


Fig. 9. Illustration of the working principle of the GAN used to generate 3D voxelized structures from grayscale 2D images. The training data is generated by the VOX-STORM model [27].

5. 3D object generation from 2D images

Having successfully segmented the 2D images and obtained satisfactory results in terms of morphological characterization, the next challenge is to generate 3D representations of the aggregates from the 2D segmentation masks. This step is crucial for achieving a more comprehensive understanding of the aggregates' properties and behavior, as 3D morphological characteristics provide valuable insights that cannot be fully captured by 2D analysis alone.

Recent advancements in deep learning have paved the way for powerful generative models that can learn to create realistic 3D objects from 2D input data. These models, such as Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs), have shown remarkable success in various domains [48], including computer vision, graphics, and medical imaging [49–51]. Several approaches have been proposed for generating 3D structures from 2D slices or projections, such as 3D-R2N2 [52], GAN-based dimensionality expansion [53], voxel-based 3D detection and reconstruction from a single image [54], transformer-based single image voxel reconstruction [55], and image-to-voxel model translation with conditional adversarial networks [56].

Many of these methods are designed for everyday objects such as cars, bicycles, furniture, or humans, and may utilize multiple images of the same object to achieve a visually realistic reconstruction. In cases where the objects are similar to the aggregates studied in this work, these methods often rely on binary images as input. For instance, Giannis et al. [57] propose using Convolutional Neural Networks (CNNs) with multiple binary images to reconstruct 3D objects, and they briefly discuss the potential of GANs for single binary image reconstruction without providing results.

However, acquiring multiple images of a single aggregate may not be feasible in industrial settings, and binary images may not provide sufficient information for accurate morphological characterization. In contrast, this section explores the application of GANs for generating 3D voxelized objects from a single grayscale 2D image. By leveraging the ability of GANs to learn the underlying distribution of 3D shapes and textures from a single grayscale projection, the aim is to bridge the gap between the 2D segmentation results and the desired 3D morphological characterization of the aggregates, focusing on the accurate representation of their morphological properties.

5.1. GANs for 3D object generation

Reconstructing a 3D object from a single 2D projection is a challenging problem, as multiple 3D objects can produce the same 2D projected image. This is because the conversion of a 3D object to a 2D image results in a loss of information, particularly depth, making the problem ill-posed [58] and difficult to solve [59]. Hadamard's criteria for well-posed problems, namely the existence, uniqueness, and stability of the solution (i.e., the solution depends continuously on the initial data of the problem), are not satisfied in this case. While multi-view projections could theoretically reduce this ambiguity, the originality and practical relevance of our approach lies precisely in addressing the single-view constraint imposed by industrial in-situ monitoring systems. In real latex aggregation processes, as in most industrial settings, equipment accessibility, reactor geometry, and real-time processing requirements limit imaging to single-view configurations. The proposed method specifically targets this industrial reality, providing a practical solution for real-world applications where multiple camera angles are not feasible. Given this inherent non-uniqueness, the proposed approach focuses on generating statistically probable 3D objects for population-level morphological characterization rather than attempting impossible unique geometric reconstruction. This probabilistic strategy is more aligned with industrial monitoring needs, where representative population behavior provides actionable insights for process control.

To mitigate the loss of information during the 2D projection, grayscale images are used instead of binary images [59]. This choice not only aligns with the real-world conditions and the way in-situ images are captured but also aims to preserve as much information as possible that would otherwise be lost during binarization.

Given the ill-posed nature of the problem, the focus shifts from faithfully reconstructing the geometry of a 3D object from multiple 2D projections to generating the most probable 3D object based on a single known 2D projection. To address this challenge, a Generative Adversarial Network (GAN) structure, proposed by Goodfellow et al. [29], is employed to train the generative model. GANs consist of two models trained in parallel: a generator and a discriminator.

GANs have proven to be highly effective in generating new data from complex distributions, such as those encountered in image analysis tasks. These distributions are often characterized by high variability, diversity, sophisticated internal structures, and non-linear relationships. In the context of this work, the generator takes a grayscale 2D projected image as input and outputs a 3D voxelized object. The discriminator, on the other hand, takes a 3D voxelized object as input and returns a probability $p \in [0; 1]$, indicating whether the object is from the training set ($p \to 1$) or generated by the generator ($p \to 0$). Fig. 9 illustrates the working principle of these two networks.

Generator architecture. The generator is designed to transform a 2D projected image of size 64×64 into a 3D voxelized structure of size $32 \times 32 \times 32$. Its architecture consists of several 3D convolutional layers, each followed by non-linear activation functions, typically ReLU and LeakyReLU [60], and batch normalization layers. The final layer uses a sigmoid activation function to produce an output in the range $[0; 1]$, representing the probability of voxel presence or absence. The detailed architecture of the generator is shown in Fig. 10.

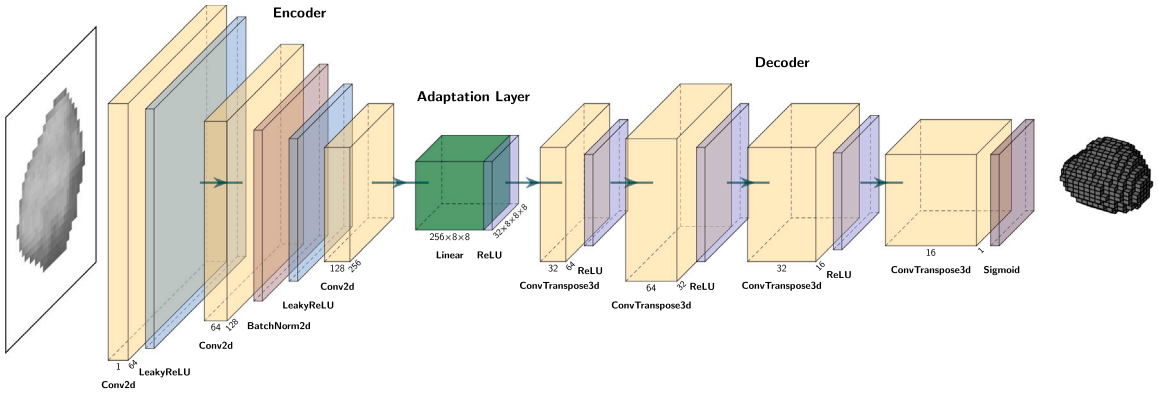


Fig. 10. Detailed diagram of the generator architecture used in the GAN for generating 3D voxelized structures from 2D projected images.

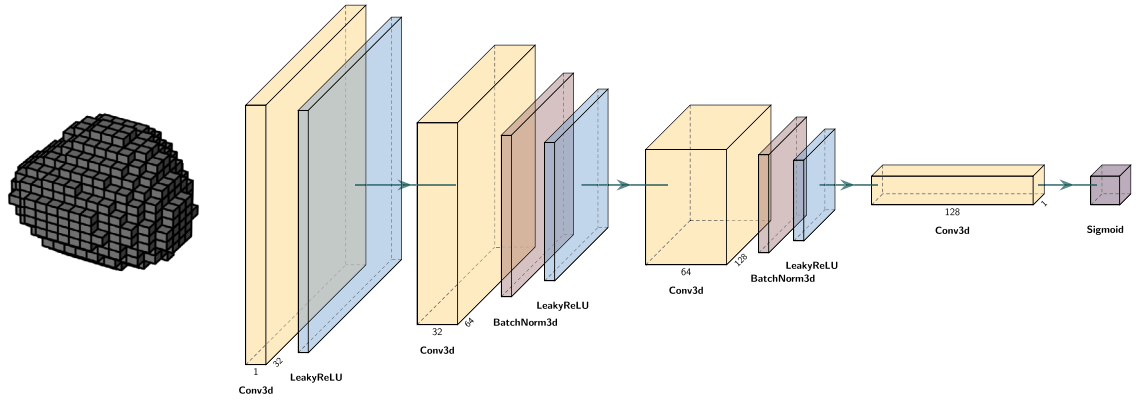


Fig. 11. Detailed diagram of the discriminator architecture used in the GAN to evaluate the resemblance of generated 3D voxelized structures from 2D projected images with 3D voxelized structures from the training set.

Discriminator architecture. The discriminator is responsible for distinguishing between the 3D structures of size $32 \times 32 \times 32$ generated by the generator and the real 3D structures from the training set, which are generated by the VOX-STORM model. It employs a convolutional neural network architecture similar to that of the generator, but with 2D convolutional layers adapted to the projected images and activation functions such as LeakyReLU to capture discriminative features. The final layer uses a sigmoid activation function to produce an output in the range $[0; 1]$, representing the probability that the object is real or fake. The detailed architecture of the discriminator is shown in Fig. 11.

The training, testing, and validation datasets are generated using the VOX-STORM model, a stochastic geometric model proposed by Théodon et al. [27]. This model allows for the rapid generation of synthetic aggregates or granular objects in the form of 3D voxelized structures. It has been successfully employed to estimate the 3D morphological characteristics of latex aggregates from ex-situ images, demonstrating its ability to generate representative datasets for the objects studied in this article.

5.2. Training

The principle of GANs is to simultaneously train the discriminator and the generator within the same training loop, with a total loss function depending on the predictions of both networks. The main idea is that the discriminator should be able to easily distinguish between real objects and those produced by the generator, helping the latter to improve and generate objects that increasingly resemble the training data.

5.2.1. Loss function

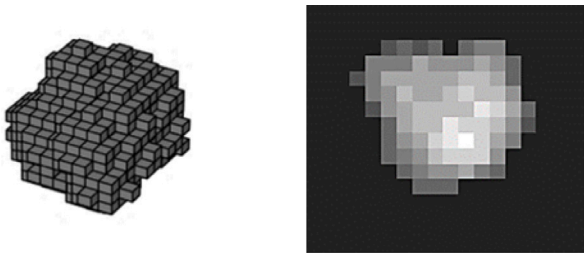
The loss function used in this work is a combination of the classical GAN loss (minimax loss) and the binary cross-entropy loss (BCE) to ensure that the generated 3D structures are realistic and close to the real structures. The GAN loss is defined as follows:

$$\mathcal{L}_{\text{GAN}}(D, G) = \mathbb{E}_{\mathbf{x}\sim p_{data}(\mathbf{x})}[\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z}\sim p_{\mathbf{z}}(\mathbf{z})}[\log(1 - D(G(\mathbf{z})))] \tag{2}$$

In this equation, $\mathcal{L}_{\text{GAN}}(D, G)$ represents the total GAN loss, D and G denote the discriminator and generator, respectively, $\mathbf{x}$ represents the real 3D images from the data distribution $p_{data}(\mathbf{x})$, and $\mathbf{z}$ represents the 2D images from the distribution $p_{\mathbf{z}}(\mathbf{z})$, which serves as the input to the generator G. The term $\log D(\mathbf{x})$ represents the probability that the real data is recognized as real by the discriminator, while $\log(1 - D(G(\mathbf{z})))$ represents the probability that the generated data is recognized as fake by the discriminator. The combination of the GAN loss and BCE loss helps to balance the training of the generator and discriminator, ensuring that the generated 3D structures are both realistic and closely match the real structures from the training set.

5.2.2. Training data

The training data consists of 40,000 voxelized 3D objects of size $32 \times 32 \times 32$ generated by the VOX-STORM model, each associated with a grayscale 2D projected image of size 64×64. The parameters used to generate this dataset were chosen to obtain a good diversity of morphological characteristics. Fig. 12 shows an example of a 3D object from the dataset, along with its associated grayscale 2D projection. To prevent overfitting, different light intensities and noise levels were applied to degrade the data.

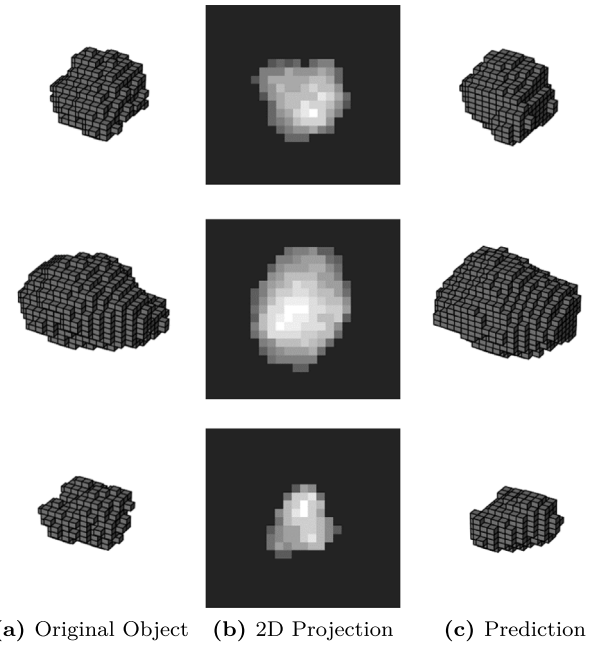


(a) 3D Object (b) 2D Projection

Fig. 12. Example of a 3D object of size 32 × 32 × 32 generated by the VOX-STORM model [27] with its associated 2D projection of size 64 × 64 (grayscale image).

Table 5
Summary of optimization and data configuration parameters.

Parameter	Generator	Discriminator
Optimizer	Adam	Adam
Learning rate	2×10^{-6}	1×10^{-6}
Betas	(0.5, 0.99)	(0.5, 0.99)
Number of epochs	5000	
Dataset size	40,000	
Training proportion	70%	
Validation proportion	15%	
Test proportion	15%	
Batch size	2048	



(a) Original Object (b) 2D Projection (c) Prediction

Fig. 13. Comparison of 3D objects produced by the generator from 2D images from the test set with the original 3D objects.

5.2.3. Training loop

The GAN training loop involves several steps and uses specific parameters, as shown in Table 5. The dataset of 40,000 elements is divided into training (70%), validation (15%), and test (15%) sets, with a batch size of 2048. The training is performed for 5000 epochs using the Adam optimizer [61] with the parameters specified in Table 5. During each epoch, the discriminator is trained by generating fake data samples from the grayscale 2D images using the generator, calculating the discriminator's predictions for both real and generated data, and updating its parameters based on the GAN loss defined by Eq. (2). The generator is then trained by generating fake data samples from the 2D images, calculating the discriminator's predictions for these samples, and updating its parameters based on the BCE loss between the predictions and the real labels. At the end of the training loop, the generator is capable of generating 3D objects from grayscale 2D images that closely resemble the training data generated by the VOX-STORM model.

5.3. Validation

To validate the method, independent data from the test set is used. These data points are unknown to the generator, as only the training and validation sets are used during the learning phase. Fig. 13 shows some examples of 3D objects produced by the generator from grayscale 2D images from the test set. It can be observed that these generated objects, while not entirely identical to the original 3D objects, still exhibit strong similarities.

It is important to note that the usual metrics (accuracy, recall, etc.) are not used for validation in this case, as the focus is solely on estimating 3D morphological characteristics rather than reconstructing objects. Instead, the distributions of 3D morphological characteristics are compared by calculating their Mean Absolute Percentage Error (MAPE), which is a measure of accuracy often used to evaluate the predictions of deep learning models.

The MAPE is defined as the average of the absolute errors in percentage of the actual values and is expressed by the following formula:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100 \tag{3}$$

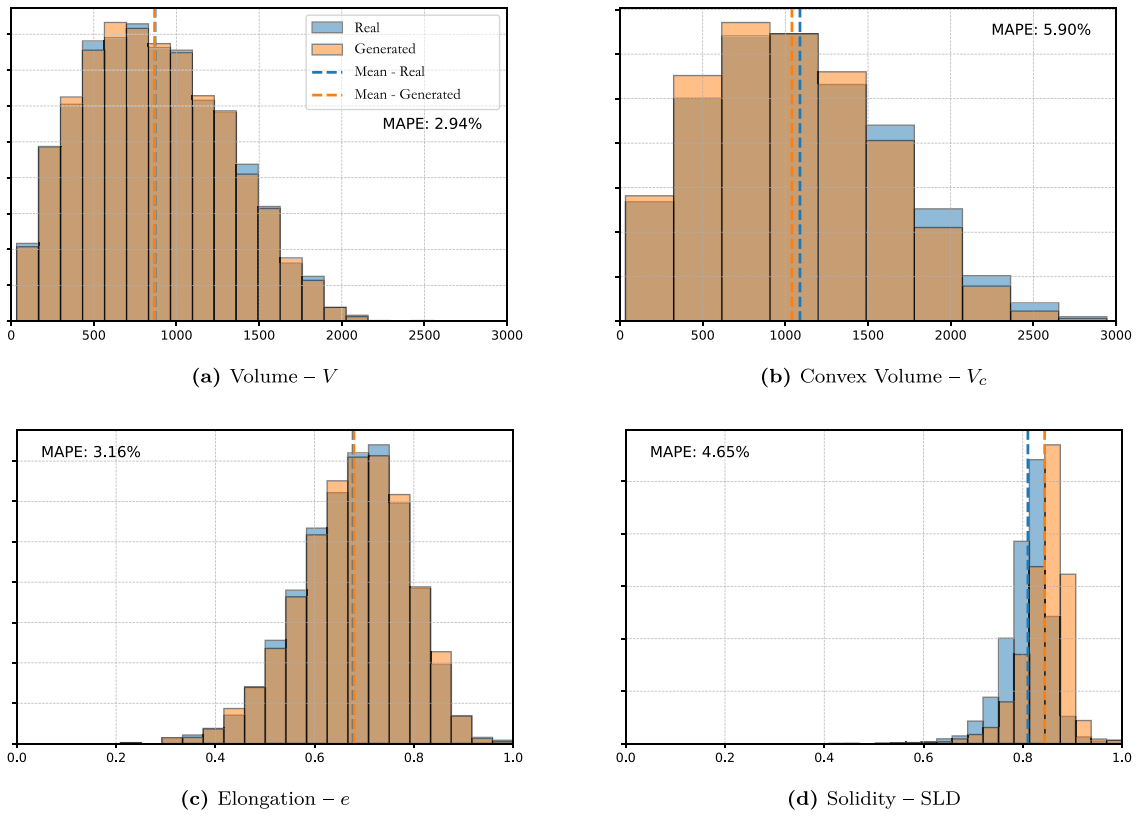
where:

- n is the size of the dataset,
- y_i is the actual value of the ith observation,
- $\hat{y}_i$ is the predicted value of the ith observation.

In the context of evaluating predictions of 3D morphological characteristics, the MAPE quantifies the average difference between the characteristics of the test set population and the characteristics of the population produced by the generator.

Fig. 14 compares the 3D morphological characteristics of the test set objects generated by the VOX-STORM model with those of the objects produced by the generator from grayscale 2D images using histogram overlays. The mean values and MAPE values are indicated for volume V, convex volume V_c, solidity SLD, and elongation of the equivalent ellipsoid e. The results obtained are particularly satisfactory, with relatively low MAPE values (around 5% at most), and the distributions of 3D morphological characteristics correspond strongly. A clear difference is visible in the convex volume, which is underestimated by the generator, and consequently, the solidity is overestimated. This can be easily explained by the fact that we are working from a single projection, and many concave areas may be hidden. Similarly, the elongation of the equivalent ellipsoid is also slightly overestimated for the same reason. In a sense, the fact that the distributions do not match perfectly is rather reassuring, as it means that the good results obtained cannot be attributed to overfitting of the model. The systematic underestimation of convex volume (MAPE: 5.90%) and overestimation of solidity (MAPE: 4.65%) can be attributed to the inherent information loss in single-view projection, where concave regions may be occluded. This bias is consistent across the dataset, suggesting systematic rather than random errors, which can be partially compensated through statistical calibration in industrial applications.

In the previous section, an image segmentation model was trained using the SPHERE model. Now, a generative model has been trained to generate 3D objects from 2D images. In the following section, the segmentation masks will be used as input to the generative model to generate synthetic object populations $\mathcal{A}'$ and $\mathcal{B}'$, and their 3D morphological characteristics will be compared to those of the original



(a) Volume – V (b) Convex Volume – V_c

(c) Elongation – e (d) Solidity – SLD

Fig. 14. Comparison of 3D morphological characteristics distributions between the 6000 test set objects generated by the VOX-STORM model and the 6000 objects produced by the generator from grayscale 2D projected images.

populations $\mathcal{A}$ and $\mathcal{B}$. It is also worth mentioning that the GAN is trained with VOX-STORM and the segmentation model with SPHERE to ensure independent validation and to avoid attributing any potential good results to overfitting.

6. Complete workflow validation

6.1. Overview

The results obtained by the generative model on the test set in the previous section demonstrate its potential for generating 3D objects from 2D images. However, to further validate the model's performance and robustness, it is essential to test it under conditions that better approximate real-world applications [62]. This section introduces a more advanced validation process that utilizes synthetic images representative of in-situ images as input data.

The motivation behind this additional validation is based on two main considerations. First, although the test data used in the previous section was not employed during model training, it was generated using the same VOX-STORM model and parameters as the training data. To eliminate any potential bias related to the data generation process [63], this validation uses synthetic data created by a modified version of the SPHERE model that incorporates 4 random Gaussian fields (see Section 4.2).

Second, in real-world scenarios, the input data for the generative model will be obtained from the segmentation of in-situ images, which may be of poor quality and introduce significant bias. By generating photorealistic synthetic images that mimic in-situ images and using the segmented data as input to the generator, we can evaluate the method's robustness under conditions that closely resemble the application to real in-situ images.

The proposed validation process consists of several key steps.

1. Firstly, two populations $\mathcal{A}$ and $\mathcal{B}$ of synthetic objects are generated using the enhanced SPHERE model with 4 random Gaussian fields. These objects are then rendered in 3D, and photorealistic synthetic images are created to resemble captured in-situ images, incorporating factors such as blur, noise, variable luminosity, and object overlap.
2. Secondly, the synthetic images are segmented using instance segmentation models based on convolutional neural networks. For each detected object, a thumbnail containing the segmentation mask and the segmented object is isolated, resized to match the generator's input size (64 × 64 px), and normalized in terms of grayscale levels.
3. The generator, which was trained on data produced by the VOX-STORM model in the previous section, is then applied to each thumbnail, generating a corresponding 3D object. This process allows for the constitution of two estimators $\mathcal{A}'$ and $\mathcal{B}'$ of the original populations $\mathcal{A}$ and $\mathcal{B}$.
4. To refine the generated 3D objects and prepare them for morphological characterization, a post-processing step is applied. This involves converting the voxelized objects into a point cloud representation, applying the *Marching Cubes* algorithm [64], and performing Laplacian smoothing [65]. This post-processing step aims to eliminate artifacts and isolated voxels while preserving the objects' morphology and enabling surface-based analysis.
5. Lastly, the distributions of 3D morphological characteristics of the objects in the original populations $\mathcal{A}$ and $\mathcal{B}$ are compared to those of the objects in the estimated populations $\mathcal{A}'$ and $\mathcal{B}'$. This comparison allows for the assessment of the generative model's performance in capturing the essential morphological properties of the objects based on their 2D projections.

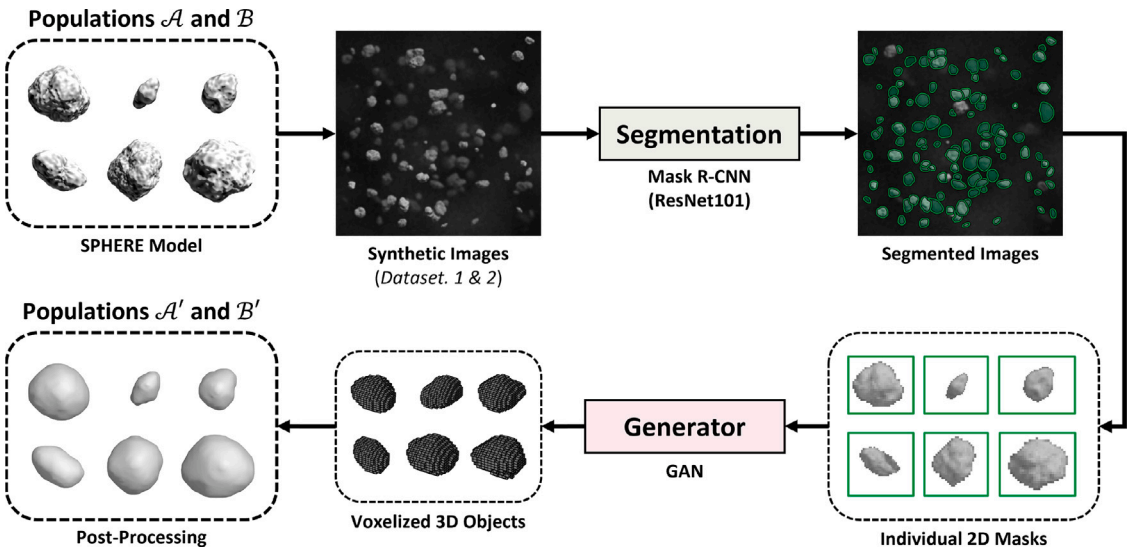


Fig. 15. Illustration of the process to validate the proposed method for estimating the 3D morphological characteristics of an aggregate population from in-situ images. Two populations of synthetic objects are generated by the SPHERE model, photorealistic images are created and segmented, and the generator trained using the VOX-STORM model produces a population of 3D objects from the individual projected images resulting from the segmentation.

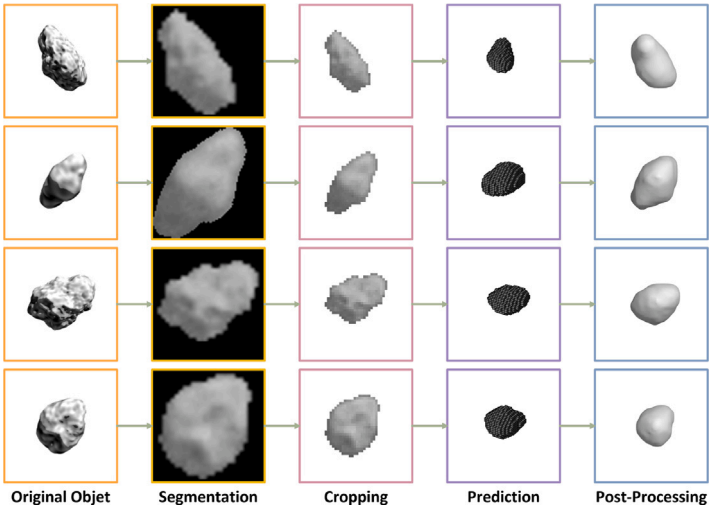


Fig. 16. Example of generator predictions from sub-images extracted by segmentation of photo-realistic synthetic in-situ images.

Fig. 15 provides a visual overview of the complete validation workflow, from the generation of synthetic objects and photorealistic images to the segmentation, 3D object generation, and morphological characterization steps.

6.2. Results

After segmenting the synthetic images from Section 4.2, the GAN's generative model is used to generate the estimated synthetic populations $\mathcal{A}'$ and $\mathcal{B}'$. Specifically, it generates 3D objects from 64 × 64 px thumbnails, which have been cropped, scaled, and had their grayscale levels equalized. Fig. 16 visualizes the complete process through a few examples and compares the morphology of the original object generated by the SPHERE model with the generator's prediction, after post-processing, which corresponds to Laplacian smoothing.

To evaluate the performance of the proposed method, we compare the distributions of 3D morphological characteristics of the objects predicted by the generator with those of the ground truth objects generated by the SPHERE model. The morphological characteristics considered in this study are volume V, equivalent sphere diameter ESD, surface area S, elongation of the equivalent ellipsoid e, solidity SLD, and sphericity Φ_S. Figs. 17 and 18 show the distributions and histograms of these characteristics for both the predicted objects and the ground truth.

To quantify the similarity between the predicted and ground truth distributions, we use the Total Variation Distance (TVD) [66], which is defined as follows:

$$TV(p,q) = \frac{1}{2}\int_{-\infty}^{\infty} |p(x) - q(x)|dx \tag{4}$$

where p and q represent the probability density functions of the compared continuous distributions. The Total Variation Distance ranges from 0 (indicating identical distributions) to 1 (denoting completely distinct distributions). A lower TVD value indicates a higher similarity between the distributions.

Table 6 presents the TVD values between the predicted and ground truth distributions for each morphological characteristic and dataset. The relatively low TVD values for most characteristics demonstrate that the generator can accurately capture the 3D morphology of the objects from their 2D projections, even when the input data is derived from the

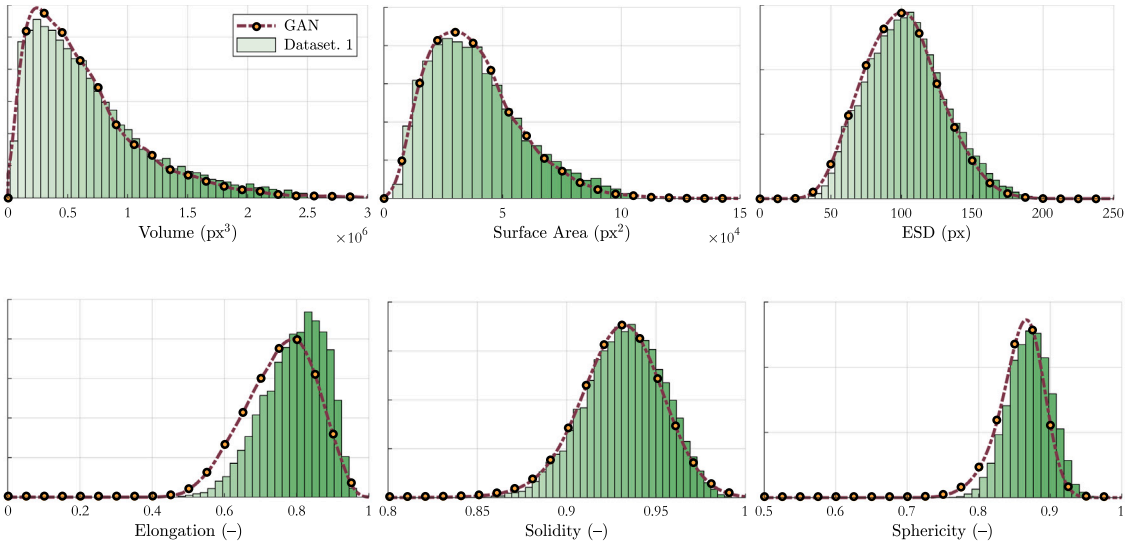


Fig. 17. Comparison of the distributions of 3D morphological characteristics predicted by the generator (GAN) and the VOX-STORM model with the *ground truth* corresponding to the synthetic objects of *Dataset. 1* (population $\mathcal{A}$) generated by the SPHERE model. The distributions and histograms are expressed in terms of frequency of counts and are normalized.

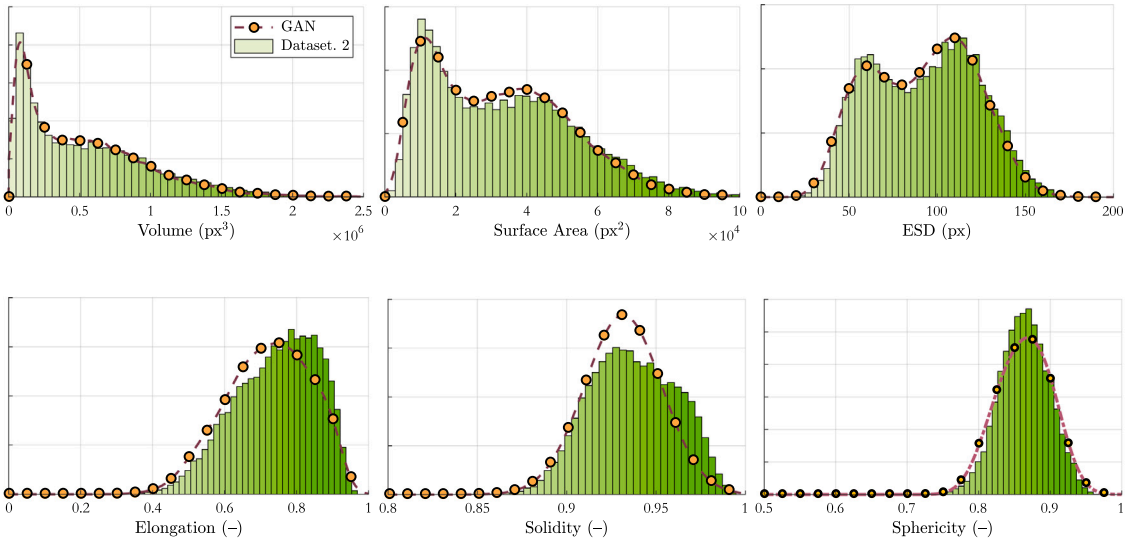


Fig. 18. Comparison of the distributions of 3D morphological characteristics predicted by the generator (GAN) and the VOX-STORM model with the *ground truth* corresponding to the synthetic objects of *Dataset. 2* (population $\mathcal{B}$) generated by the SPHERE model. The distributions and histograms are expressed in terms of frequency of counts and are normalized.

Table 6
Total variation distance between the distributions of 3D morphological characteristics predicted by the GAN and the corresponding *ground truths* representing the distributions of 3D morphological characteristics of the objects generated by the SPHERE model. The results are rounded to the nearest hundredth.

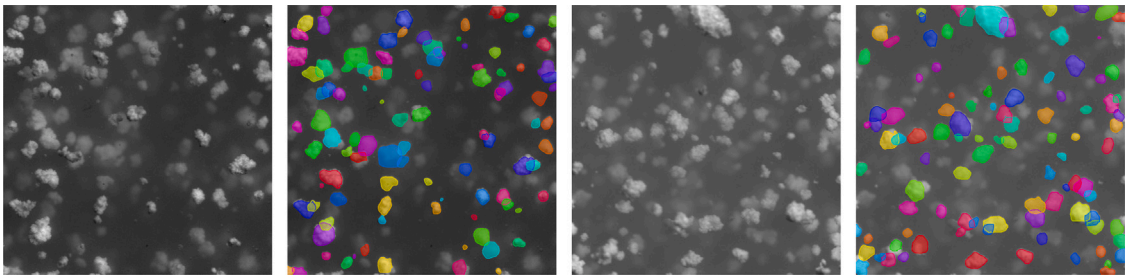
Dataset	V	S	ESD	e	SLD	Φ_S
Dataset. 1	0.05	0.04	0.05	0.17	0.13	0.15
Dataset. 2	0.03	0.04	0.03	0.18	0.18	0.17

segmentation of photorealistic synthetic images. This is particularly remarkable considering that the generator was trained on data produced by a different model (VOX-STORM) and that the segmentation process introduces its own bias and imperfections.

The proposed method offers several advantages in terms of flexibility and efficiency. Unlike other approaches that may require extensive retraining or fine-tuning for each new dataset, the generator can quickly generate 3D objects from 2D images without any additional training. This makes the method suitable for real-world applications where the characteristics of the objects may vary significantly between different datasets or over time.

Moreover, the generator's ability to learn and generalize from a diverse set of training data allows it to handle a wide range of object morphologies and imaging conditions. This robustness is essential for practical applications, where the quality and characteristics of the input images may not always be optimal or consistent.

The results presented in this section demonstrate the effectiveness of the proposed method in estimating the 3D morphological characteristics of objects from their 2D projections. The generator's performance, as measured by the Total Variation Distance between the predicted and ground truth distributions, highlights its potential for accurate and efficient 3D characterization of objects in real-world settings.



(a) Original Image (7 h) (b) Segmented Image (c) Original Image (10 h) (d) Segmented Image

Fig. 19. Illustration of segmentation masks predicted by the Mask R-CNN model for real in-situ images.

6.3. Discussion

The results presented in the previous section demonstrate that the proposed method, combining instance segmentation and a GAN-based generative model, can effectively estimate the 3D morphological characteristics of synthetic aggregate populations from their 2D projections. The distributions of size and shape characteristics are particularly well estimated, with relatively low Total Variation Distances between the predicted and ground truth distributions. The propagation of segmentation errors to 3D estimation was assessed through a Total Variation Distance analysis, showing that population-level distributions remain robust despite individual particle segmentation imperfections (ARE: 12–33%).

However, the performance is slightly less accurate for angularity and texture characteristics, such as sphericity and solidity, as well as shape characteristics, such as elongation. This can be attributed to several factors:

- **Input quality**: The input to the generative model is based on data obtained from the segmentation of synthetic images. To resemble real in-situ images, these synthetic images are heavily degraded, which can lead to segmentation errors [67] and impact the generator's predictions [68–70].
- **Low resolution**: Although the generated images have a resolution of 2048 × 2048 px, the aggregates themselves are very small, which does not allow for accurate capture of their contours and texture. Moreover, the resolution of the input images to the generator (64 × 64) and the output (32 × 32 × 32) is extremely low (albeit in line with the state-of-the-art [57]), which limits the generation of objects with particularly complex angularity or texture.
- **Single projection**: It is important to keep in mind that the method relies on a single projection of each object, which does not provide a complete appreciation of the object's geometry [59]. This is particularly relevant when considering solidity, as concave areas may be hidden, and especially elongation.

Despite these limitations, the proposed approach can produce results within seconds on a machine equipped with an Nvidia GeForce RTX 3090 GPU with 24 GB of GDDR6X memory. Furthermore, the process can be fully automated, integrating the segmentation step, and used as-is in any situation, for any dataset, making it suitable for integration into an industrial process. The real-time processing capability (15 ms per particle on RTX 3090 GPU) addresses the specific constraints of continuous industrial monitoring from single in-situ projections, representing a distinct application domain from laboratory-based 3D reconstruction methods.

The proposed approach specifically targets the morphological characteristics of compact latex aggregates with relatively high 3D fractal dimensions [71], as typically observed in the industrial D-480 latex aggregation process studied here. While real aggregates can theoretically exhibit complex fractal or dendritic structures, the industrial observations indicate that the aggregates formed in this specific system remain within the morphological range captured by the models. The VOX-STORM model used for GAN training is actually capable of generating significantly more complex structures (higher fractal dimensions, more intricate concavities) than those observed in this industrial context, providing robust training coverage. Conversely, the SPHERE model used for synthetic image generation has proven sufficient to represent the morphological variability of aggregates encountered in this specific industrial application. Extending to dramatically different aggregate types (e.g., highly dendritic or fractal structures) would constitute a different research scope and is beyond the intended application domain of this work.

The ability of the GAN to learn and generalize from a diverse set of training data, generated by the VOX-STORM model, allows it to handle a wide range of object morphologies and imaging conditions. This robustness is essential for practical applications, where the quality and characteristics of the input images may vary significantly. Building upon the results obtained on synthetic datasets, the next section will focus on the application of the proposed workflow to real in-situ images, as mentioned in Section 2.1.

7. Application

This section focuses on the application of the proposed workflow to real in-situ images taken during the aggregation process of D-480 latex particles, as described in Section 2.1, based on the results obtained on synthetic datasets. The goal is to apply the method to real-world scenarios where the challenges of image quality, object overlap, and morphological diversity are more pronounced.

7.1. Segmentation of real images

The Mask R-CNN (ResNet101) model is used to perform automatic segmentation of the in-situ images. Fig. 19 shows examples of segmented real images, and Fig. 20 presents the distributions of 2D morphological characteristics obtained. It is evident that only the object density evolves over time, which is consistent with the observations made during the experimental protocol. The evolution of aggregate size, illustrated by Fig. 20(a) (number distribution) and Fig. 20(b) (volume distribution), is extremely limited, with a very slight increase in the number of large objects over time. These results can be explained in two ways:

1. As mentioned previously, it was observed during the experimental process that the evolution of the morphological characteristics of the aggregates, particularly their size, was very limited. This was primarily due to the type of latex used.
2. Image analysis generally only allows the observation of the largest objects, typically beyond 100 μm. The camera used offers a resolution of approximately 10 μm/pixel, which does not allow for the detection of smaller objects. Unfortunately, measurements performed in parallel by laser diffraction through successive sampling show that a large proportion (in number)

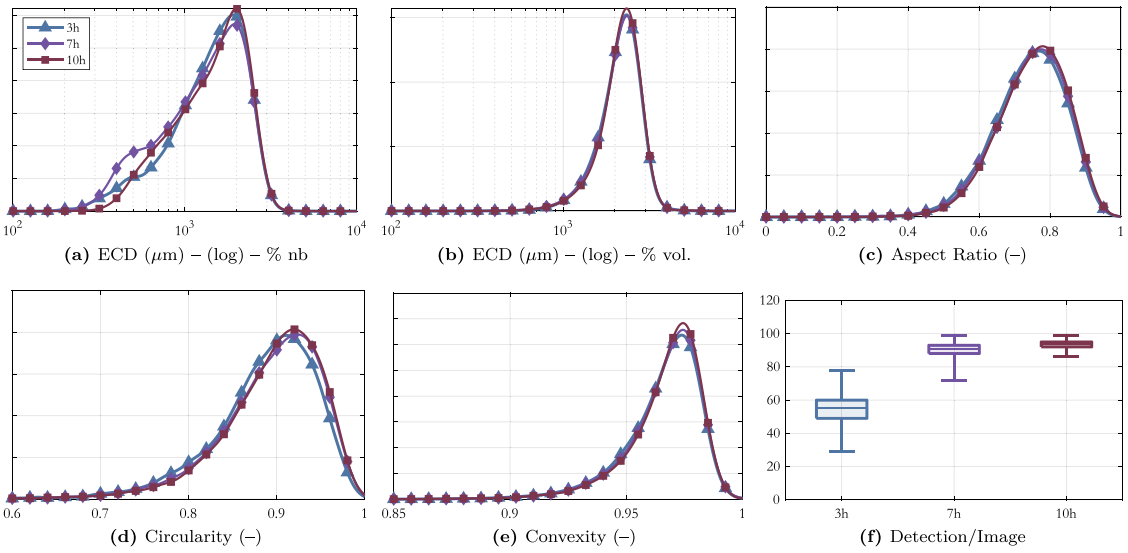


Fig. 20. Distributions of 2D morphological characteristics obtained by segmentation of in-situ images and number of objects detected per image (f).

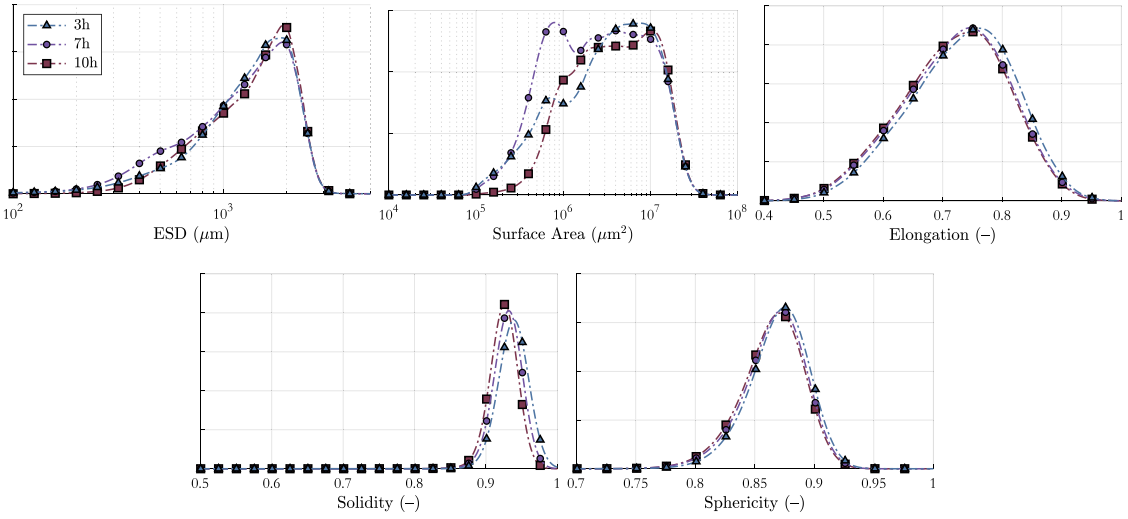


Fig. 21. Distributions of 3D morphological characteristics measured on the digital twin of the latex aggregate population.

of the aggregate population is between 0.1 and 1 μm. However, these measurements do not allow for the observation of objects beyond 1700 μm, which means that the combination of these two types of measurements is necessary to have a global overview of the evolution of the population morphology.

Therefore, although the distributions of morphological characteristics do not change significantly over time, it is reassuring to note that these results are consistent with other observations and measurements performed in parallel with this experiment. In particular, the increasing number of objects detected per image over time is consistent with the increase in the proportion of objects larger than 100 μm, i.e., objects detectable by the optical device.

7.2. Generation of a digital twin

The pre-trained GAN generator, trained on synthetic data, is used to generate a population of voxelized objects from the data obtained by segmenting the in-situ images. This population of synthetic objects can be considered as a digital twin of the real population of latex aggregates.

The results obtained are presented in Fig. 21. The same trends observed for the 2D distributions can be seen in the 3D morphological characteristics. There is little evolution over time, with perhaps slightly larger objects being generated and thus detected.

7.3. Discussion

The application of the proposed workflow to real in-situ images of latex aggregates demonstrates the technical feasibility of estimating the 3D morphological characteristics of the aggregate population from 2D images. Although the ground truth is not accessible in this real scenario, the results obtained from the segmentation of in-situ images and the subsequent generation of 3D objects using pre-trained GAN are consistent with observations made during the experimental process. The slow evolution of morphological characteristics over time, observed in both 2D and 3D distributions, is in line with expectations given the type of latex used and the limitations of image analysis in detecting small objects. These results suggest that the proposed method is capable of capturing the essential characteristics of the aggregate population in real-life scenarios, although direct quantitative validation is not possible in the absence of ground truth.

The generation of a digital twin of the aggregate population provides several benefits for industrial applications. It enables the virtual exploration and optimization of the aggregation process, reducing the need for physical experiments and leading to improved product quality and process efficiency [31,72]. Real-time monitoring and prediction of aggregate population characteristics become possible by continuously updating the digital twin with new in-situ images. This allows for the early detection of deviations and the implementation of timely adjustments to process parameters, ensuring consistent product quality [30,73].

The successful application of the proposed workflow to real in-situ images demonstrates its effectiveness in estimating the 3D morphological characteristics of latex aggregates in real-world scenarios. The generation of a digital twin opens up new opportunities for process optimization, control, and monitoring in industrial settings [23,74].

8. Advantages and limitations

The proposed workflow for estimating the 3D morphological characteristics of latex aggregates from 2D in-situ images offers several advantages over existing methods.

1. The combination of deep learning techniques, such as CNNs for image segmentation and GANs for 3D object generation, with stochastic geometry models for generating synthetic training data, provides a comprehensive and efficient solution for characterizing complex granular systems compared to traditional 2D characterization methods [14,75] or methods based solely on stochastic geometry [76–78]. The use of synthetic data generated by the SPHERE model for training the segmentation model and the VOX-STORM model for training the generative model ensures dataset independence and reduces the risk of overfitting.
2. The method is highly flexible and can be adapted to various types of particles or aggregates with different morphological properties. The stochastic geometry models used in this study, SPHERE and VOX-STORM, have been shown to generate realistic and representative 3D structures of granular objects [27,32], making them suitable for a wide range of applications. Moreover, the use of grayscale images instead of binary images as input for the generative model preserves more information about the object's texture and angularity [59], enabling a more accurate reconstruction of the 3D morphology.
3. The proposed workflow is computationally efficient and can generate 3D objects from 2D images within seconds using a GPU. This makes the method suitable for real-time monitoring and control of industrial processes, where the rapid characterization of particle morphology is crucial for optimizing product quality and process efficiency [22,79].

However, the proposed method also has some limitations that should be acknowledged.

1. One of the main challenges is the limited resolution of the input images and the generated 3D objects. The current state-of-the-art for generating 3D objects from 2D images using GANs or CNNs typically involves input images of size 64×64 pixels and output objects of size $32 \times 32 \times 32$ voxels [52,54,57]. This low resolution may not capture the fine details of the object's surface texture and angularity, leading to a less accurate characterization of the 3D morphology.
2. Another limitation is the reliance on a single 2D projection of the object for generating the 3D structure. While the use of grayscale images provides more information than binary images, it still does not capture the full 3D geometry of the object [59]. This can lead to inaccuracies in the estimation of certain morphological characteristics, such as elongation and solidity, where the presence of concave regions or hidden surfaces can significantly affect the measurements.
3. Furthermore, the performance of the proposed method depends on the quality of the input images and the accuracy of the segmentation model. In real-world scenarios, in-situ images may suffer from various issues, such as noise, blur, and uneven illumination, which can affect the segmentation results and, consequently, the quality of the generated 3D objects [68–70]. While the use of synthetic images for training the segmentation model can help mitigate these issues to some extent, there may still be a domain gap between the synthetic and real images that could impact the performance of the method [80].

Despite these limitations, the proposed workflow advances the characterization of 3D morphological properties of particles and aggregates derived from 2D images. It shows promising results on synthetic validation data and demonstrates the potential for straightforward application to real images, facilitating deployment in an industrial context.

9. Conclusion

This article presents a complete workflow for estimating 3D morphological characteristics of latex aggregates from 2D in-situ images, combining CNNs for segmentation, GANs for 3D generation, and stochastic geometry models (SPHERE and VOX-STORM) for synthetic training data.

The three-step workflow encompasses: (1) automatic image segmentation, (2) 3D object generation from segmentation masks, and (3) morphological characteristic estimation. Validation using synthetic populations $\mathcal{A}$ and $\mathcal{B}$ demonstrated accurate estimation with low Total Variation Distance values, showing that estimated populations $\mathcal{A}'$ and $\mathcal{B}'$ closely matched ground truth distributions for size, shape, and angularity characteristics.

Application to real in-situ images confirmed technical feasibility and consistency with experimental observations, successfully generating a digital twin of the latex aggregate population. This enables virtual process exploration, real-time monitoring, and predictive capabilities for industrial applications.

The method's flexibility, efficiency, and robustness make it suitable for various industrial settings requiring 3D morphological characterization from single 2D projections. Future work could focus on improving segmentation and generative model performance, adapting the method to other particle types, and validating on additional industrial datasets. Enhancements could include higher-resolution inputs via deep learning-based super-resolution techniques [81,82], diffusion models as GAN alternatives [83–86], and integration with other characterization techniques [87] such as rheology or chemical analysis [71] for comprehensive understanding of morphology-property relationships in granular systems [88].

CRediT authorship contribution statement

L. Théodon: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Conceptualization. **C. Coufort-Saudejaud:** Writing – review & editing, Supervision.

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

The author(s) acknowledge(s) the support of the French Agence Nationale de la Recherche (ANR), under grant ANR-20-CE07-0025 (MORPHING project).

Data availability

The authors do not have permission to share the workflow itself but the 2 models mentioned in the paper (SPHERE and VOX STORM) are available. The links are provided in the cited papers.

References

[1] A. Mehle, B. Likar, D. Tomaževič, In-line recognition of agglomerated pharmaceutical pellets with density-based clustering and convolutional neural network, in: 2017 Fifteenth IAPR International Conference on Machine Vision Applications, MVA, 2017, pp. 9–12, http://dx.doi.org/10.23919/MVA.2017.7986760.

[2] Z.M. Lu, L. Zhang, D.M. Fan, N.M. Yao, C.X. Zhang, Crystal texture recognition system based on image analysis for the analysis of agglomerates, Chemometr. Intell. Lab. Syst. 200 (2020) 103985, http://dx.doi.org/10.1016/j.chemolab.2020.103985, URL: https://www.sciencedirect.com/science/article/pii/S0169743919307695.

[3] J. Huyan, W. Li, S. Tighe, Y. Zhang, B. Yue, Image-based coarse-aggregate angularity analysis and evaluation, J. Mater. Civ. Eng. 32 (6) (2020) 04020140, http://dx.doi.org/10.1061/(ASCE)MT.1943-5533.0003150.

[4] J. Adamcik, J.-M. Jung, J. Flakowski, P. De Los Rios, G. Dietler, R. Mezzenga, Understanding amyloid aggregation by statistical analysis of atomic force microscopy images, Nature Nanotechnology 5 (6) (2010) 423–428, http://dx.doi.org/10.1038/nnano.2010.59.

[5] F.Z. Vissotto, R.C. Giarola, L.C. Jorge, G.T. Makita, G.M.B.Q. Cardozo, M.I. Rodrigues, F.C. Menegalli, Morphological characterization with image analysis of cocoa beverage powder agglomerated with steam, Food Sci. Technol. 34 (Food Sci. Technol, 2014 34(4)) (2014) http://dx.doi.org/10.1590/1678-457X.6246.

[6] S. Lowell, J. Shields, M. Thomas, M. Thommes, Characterization of porous solids and powders: Surface area, pore size and density, Particle Technology Series, Springer Netherlands, 2012, URL: https://books.google.fr/books?id=IwvSBwAAQBAJ.

[7] W. Xu, L. Zhang, D. Fan, L. Xu, K. Liu, E. Dong, T. Yin, R. Yu, AI-infused characteristics prediction and multi-objective design of ultra-high performance concrete (UHPC): From pore structures to macro-performance, J. Build. Eng. 98 (2024) 111170, http://dx.doi.org/10.1016/j.jobe.2024.111170, URL: https://www.sciencedirect.com/science/article/pii/S2352710224027384.

[8] D. Fan, Z. Chen, Y. Cao, K. Liu, T. Yin, X.-S. Lv, J.-X. Lu, A. Zhou, C. Sun Poon, R. Yu, Intelligent predicting and monitoring of ultra-high-performance fiber reinforced concrete composites - a review, Compos. Part A: Appl. Sci. Manuf. 188 (2025) 108555, http://dx.doi.org/10.1016/j.compositesa.2024.108555, URL: https://www.sciencedirect.com/science/article/pii/S1359835X24005530.

[9] L. Allen Cooley Jr., R.S. James, Micro-deval testing of aggregates in the southeast, Transp. Res. Rec. 1837 (1) (2003) 73–79, http://dx.doi.org/10.3141/1837-08, arXiv:https://doi.org/10.3141/1837-08.

[10] O. Furat, M. Wang, M. Neumann, L. Petrich, M. Weber, C.E. Krill, V. Schmidt, Machine learning techniques for the segmentation of tomographic image data of functional materials, Front. Mater. 6 (2019) http://dx.doi.org/10.3389/fmats.2019.00145, URL: https://www.frontiersin.org/journals/materials/articles/10.3389/fmats.2019.00145.

[11] E.W. Olson, Particle shape factors and their use in image analysis part ii: practical applications, 2011, URL: https://api.semanticscholar.org/CorpusID:123562853.

[12] P. Bosilj, I. Gould, T. Duckett, G. Cielniak, Estimating soil aggregate size distribution from images using pattern spectra, Biosyst. Eng. 198 (2020) 63–77, http://dx.doi.org/10.1016/j.biosystemseng.2020.07.012, URL: https://www.sciencedirect.com/science/article/pii/S1537511020302051.

[13] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444, http://dx.doi.org/10.1038/nature14539.

[14] M. Frei, F. Kruis, Image-based size analysis of agglomerated and partially sintered particles via convolutional neural networks, Powder Technol. 360 (2020) 324–336, http://dx.doi.org/10.1016/j.powtec.2019.10.020, URL: https://www.sciencedirect.com/science/article/pii/S003259101930854X.

[15] P. Monchot, L.c. Coquelin, K. Guerroudj, N. Feltin, A. Delvallée, L.c. Crouzier, N. Fischer, Deep learning based instance segmentation of titanium dioxide particles in the form of agglomerates in scanning electron microscopy, Nanomaterials 11 (4) (2021) http://dx.doi.org/10.3390/nano11040968, URL: https://www.mdpi.com/2079-4991/11/4/968.

[16] L. Zhang, Q. Zhang, Z. Li, H. Wang, Test on compaction reinforcement effect of sand, Adv. Mater. Sci. Eng. 2020 (1) (2020) 3685619, http://dx.doi.org/10.1155/2020/3685619, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1155/2020/3685619, URL: https://onlinelibrary.wiley.com/doi/abs/10.1155/2020/3685619.

[17] X. Wu, X. Liu, J. Duan, Online size distribution measurement of dense iron green pellets using an efficient and multiscale nested U-net method, Powder Technol. 387 (2021) 584–600, http://dx.doi.org/10.1016/j.powtec.2021.04.045, URL: https://www.sciencedirect.com/science/article/pii/S0032591021003260.

[18] J. Song, X. Zhou, R. Jiang, Research on characterization of 3D morphology of coarse aggregate based on laser scanning, Buildings 13 (4) (2023) http://dx.doi.org/10.3390/buildings13041029, URL: https://www.mdpi.com/2075-5309/13/4/1029.

[19] S. Raynaud, G. Vasseur, R. Soliva, In vivo CT X-ray observations of porosity evolution during triaxial deformation of a calcarenite, Int. J. Rock Mech. Min. Sci. 56 (2012) 161–170, http://dx.doi.org/10.1016/j.ijrmms.2012.07.020, URL: https://www.sciencedirect.com/science/article/pii/S1365160912001578.

[20] R. Pashminehazar, A. Kharaghani, E. Tsotsas, Three dimensional characterization of morphology and internal structure of soft material agglomerates produced in spray fluidized bed by X-ray tomography, Powder Technol. 300 (2016) 46–60, http://dx.doi.org/10.1016/j.powtec.2016.03.053, URL: https://www.sciencedirect.com/science/article/pii/S0032591016301565, 7th International Granulation Workshop 2015: Granulation across the length scales.

[21] R. Pashminehazar, S.J. Ahmed, A. Kharaghani, E. Tsotsas, Spatial morphology of maltodextrin agglomerates from X-ray microtomographic data: Real structure evaluation vs. spherical primary particle model, Powder Technol. 331 (2018) 204–217, http://dx.doi.org/10.1016/j.powtec.2018.03.008, URL: https://www.sciencedirect.com/science/article/pii/S0032591018301918.

[22] M.G. Rasteiro, A. Koponen, Monitoring aggregation processes in multiphase systems: A review, Powders 3 (1) (2024) 77–110, http://dx.doi.org/10.3390/powders3010007, URL: https://www.mdpi.com/2674-0516/3/1/7.

[23] F. Tao, J. Cheng, Q. Qi, M. Zhang, H. Zhang, F. Sui, Digital twin-driven product design, manufacturing and service with big data, Int. J. Adv. Manuf. Technol. 94 (9) (2018) 3563–3576, http://dx.doi.org/10.1007/s00170-017-0233-1.

[24] S. Grottel, M. Krone, C. Müller, G. Reina, T. Ertl, Megamol—A prototyping framework for particle-based visualization, IEEE Trans. Vis. Comput. Graphics 21 (2) (2015) 201–214, http://dx.doi.org/10.1109/TVCG.2014.2350479.

[25] J. Staib, S. Grottel, S. Gumhold, Visualization of particle-based data with transparency and ambient occlusion, Comput. Graph. Forum 34 (3) (2015) 151–160, http://dx.doi.org/10.1111/cgf.12627, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.12627, URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/cgf.12627.

[26] L. Théodon, C. Coufort-Saudejaud, J. Debayle, A stochastic model based on Gaussian random fields to characterize the morphology of granular objects, Pattern Recognit. 149 (2024) 110255, http://dx.doi.org/10.1016/j.patcog.2024.110255, URL: https://www.sciencedirect.com/science/article/pii/S0031320324000062.

[27] L. Théodon, J. Debayle, C. Coufort-Saudejaud, VOX-STORM: A stochastic 3D model based on a dual voxel-mesh architecture for the morphological characterization of aggregates, Powder Technol. 444 (2024) 119983, http://dx.doi.org/10.1016/j.powtec.2024.119983, URL: https://www.sciencedirect.com/science/article/pii/S0032591024006260.

[28] K. He, G. Gkioxari, P. Dollár, R. Girshick, Mask R-CNN, 2018, arXiv:1703.06870.

[29] I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial networks, 2014, arXiv:1406.2661.

[30] E. Negri, L. Fumagalli, M. Macchi, A review of the roles of digital twin in CPS-based production systems, Procedia Manuf. 11 (2017) 939–948, http://dx.doi.org/10.1016/j.promfg.2017.07.198, URL: https://www.sciencedirect.com/science/article/pii/S2351978917304067, 27th International Conference on Flexible Automation and Intelligent Manufacturing, FAIM2017, 27-30 June 2017, Modena, Italy.

[31] W. Kritzinger, M. Karner, G. Traar, J. Henjes, W. Sihn, Digital twin in manufacturing: A categorical literature review and classification, IFAC- Pap. 51 (11) (2018) 1016–1022, http://dx.doi.org/10.1016/j.ifacol.2018.08.474, URL: https://www.sciencedirect.com/science/article/pii/S2405896318316021, 16th IFAC Symposium on Information Control Problems in Manufacturing INCOM 2018.

[32] L. Théodon, J. Debayle, C. Coufort-Saudejaud, Morphological characterization of aggregates and agglomerates by image analysis: A systematic literature review, Powder Technol. 430 (2023) 119033, http://dx.doi.org/10.1016/j.powtec.2023.119033, URL: https://www.sciencedirect.com/science/article/pii/S0032591023008161.

[33] E. Shelhamer, J. Long, T. Darrell, Fully convolutional networks for semantic segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 39 (4) (2017) 640–651, http://dx.doi.org/10.1109/TPAMI.2016.2572683.

[34] K. He, G. Gkioxari, P. Dollár, R. Girshick, Mask R-CNN, in: 2017 IEEE International Conference on Computer Vision, ICCV, 2017, pp. 2980–2988, http://dx.doi.org/10.1109/ICCV.2017.322.

[35] G. Jocher, A. Chaurasia, J. Qiu, Ultralytics YOLO, 2023, URL: https://github.com/ultralytics/ultralytics.

[36] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A.C. Berg, W.-Y. Lo, P. Dollár, R. Girshick, Segment anything, 2023, arXiv:2304.02643.

[37] L. Théodon, C. Coufort-Saudejaud, J. Debayle, Deep learning-based instance segmentation to characterize the morphology of compact aggregates through image analysis, in: 2024 14th International Conference on Pattern Recognition Systems, ICPRS, 2024, pp. 1–7, http://dx.doi.org/10.1109/ICPRS62101.2024.10677841.

[38] B.D. Ripley, Stochastic geometry and its applications, J. Roy. Statist. Soc. Ser. A 151 (1) (1988) 239–240, http://dx.doi.org/10.2307/2982216, URL: https://ideas.repec.org/a/bla/jorssa/v151y1988i1p239-240.html.

[39] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, G. Klambauer, S. Hochreiter, GANs trained by a two time-scale update rule converge to a Nash equilibrium, 2017, CoRR abs/1706.08500, URL: http://arxiv.org/abs/1706.08500, arXiv:1706.08500.

[40] D. Dowson, B. Landau, The Fréchet distance between multivariate normal distributions, J. Multivariate Anal. 12 (3) (1982) 450–455, http://dx.doi.org/10.1016/0047-259X(82)90077-X.

[41] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, 2015, CoRR abs/1512.00567, URL: http://arxiv.org/abs/1512.00567, arXiv:1512.00567.

[42] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, C.L. Zitnick, Microsoft COCO: Common objects in context, in: D. Fleet, T. Pajdla, B. Schiele, T. Tuytelaars (Eds.), Computer Vision – ECCV 2014, Springer International Publishing, Cham, 2014, pp. 740–755.

[43] T. Fawcett, An introduction to ROC analysis, Pattern Recognit. Lett. 27 (8) (2006) 861–874, http://dx.doi.org/10.1016/j.patrec.2005.10.010, ROC Analysis in Pattern Recognition.

[44] W. Gu, S. Bai, L. Kong, A review on 2D instance segmentation based on deep neural networks, Image Vis. Comput. 120 (2022) 104401, http://dx.doi.org/10.1016/j.imavis.2022.104401.

[45] Y. Tian, D. Su, S. Lauria, X. Liu, Recent advances on loss functions in deep learning for computer vision, Neurocomputing 497 (2022) 129–158, http://dx.doi.org/10.1016/j.neucom.2022.04.127.

[46] Z. Kulpa, Area and perimeter measurement of blobs in discrete binary pictures, Comput. Graph. Image Process. 6 (5) (1977) 434–451, http://dx.doi.org/10.1016/S0146-664X(77)80021-X, URL: https://www.sciencedirect.com/science/article/pii/S0146664X7780021X.

[47] E. Pirard, G. Dislaire, Robustness of planar shape descriptors of particles, in: Proceedings Mathematical Geology Congress, 2005.

[48] S. Bond-Taylor, A. Leach, Y. Long, C.G. Willcocks, Deep generative modelling: A comparative review of VAEs, GANs, normalizing flows, energy-based and autoregressive models, IEEE Trans. Pattern Anal. Mach. Intell. 44 (11) (2022) 7327–7347, http://dx.doi.org/10.1109/TPAMI.2021.3116668.

[49] Y. Wang, B. Yu, L. Wang, C. Zu, D.S. Lalush, W. Lin, X. Wu, J. Zhou, D. Shen, L. Zhou, 3D conditional generative adversarial networks for high-quality pet image estimation at low dose, NeuroImage 174 (2018) 550–562, http://dx.doi.org/10.1016/j.neuroimage.2018.03.045, URL: https://www.sciencedirect.com/science/article/pii/S1053811918302507.

[50] A. Brock, J. Donahue, K. Simonyan, Large scale GAN training for high fidelity natural image synthesis, 2019, arXiv:1809.11096, URL: https://arxiv.org/abs/1809.11096.

[51] Y. Skandarani, P.-M. Jodoin, A. Lalande, GANs for medical image synthesis: An empirical study, J. Imaging 9 (3) (2023) http://dx.doi.org/10.3390/jimaging9030069, URL: https://www.mdpi.com/2313-433X/9/3/69.

[52] C.B. Choy, D. Xu, J. Gwak, K. Chen, S. Savarese, 3D-R2N2: A unified approach for single and multi-view 3D object reconstruction, 2016, ArXiv arXiv:1604.00449, URL: https://api.semanticscholar.org/CorpusID:6325059.

[53] S. Kench, S.J. Cooper, Generating three-dimensional structures from a two-dimensional slice with generative adversarial network-based dimensionality expansion, Nat. Mach. Intell. 3 (4) (2021) 299–305, http://dx.doi.org/10.1038/s42256-021-00322-1.

[54] F. Liu, X. Liu, Voxel-based 3D detection and reconstruction of multiple objects from a single image, in: Proceedings of the 35th International Conference on Neural Information Processing Systems, NIPS '21, Curran Associates Inc., Red Hook, NY, USA, 2024.

[55] J.J. Lee, B. Benes, Snakevoxformer: Transformer-based single image\\voxel reconstruction with run length encoding, 2023, http://dx.doi.org/10.48550/ARXIV.2303.16293, CoRR arXiv:2303.16293, arXiv:2303.16293.

[56] V.A. Knyaz, V.V. Kniaz, F. Remondino, Image-to-voxel model translation with conditional adversarial networks, in: L. Leal-Taixé, S. Roth (Eds.), Computer Vision – ECCV 2018 Workshops, Springer International Publishing, Cham, 2019, pp. 601–618.

[57] K. Giannis, C. Thon, G. Yang, A. Kwade, C. Schilde, Predicting 3D particles shapes based on 2D images by using convolutional neural network, Powder Technol. 432 (2024) 119122, http://dx.doi.org/10.1016/j.powtec.2023.119122, URL: https://www.sciencedirect.com/science/article/pii/S0032591023009051.

[58] K. Fu, J. Peng, Q. He, H. Zhang, Single image 3D object reconstruction based on deep learning: A review, Multimedia Tools Appl. 80 (2020) 463–498, URL: https://api.semanticscholar.org/CorpusID:225273174.

[59] Y. Yang, J. Han, D. Zhang, D. Cheng, Disentangling deep network for reconstructing 3D object shapes from single 2D images, in: H. Ma, L. Wang, C. Zhang, F. Wu, T. Tan, Y. Wang, J. Lai, Y. Zhao (Eds.), Pattern Recognition and Computer Vision, Springer International Publishing, Cham, 2021, pp. 153–166.

[60] A.K. Dubey, V. Jain, Comparative study of convolution neural network's relu and leaky-relu activation functions, in: S. Mishra, Y.R. Sood, A. Tomar (Eds.), Applications of Computing, Automation and Wireless Systems in Electrical Engineering, Springer Singapore, Singapore, 2019, pp. 873–880.

[61] D.P. Kingma, J. Ba, Adam: A method for Stochastic optimization, 2017, arXiv:1412.6980.

[62] E. Barbierato, M.L.D. Vedova, D. Tessera, D. Toti, N. Vanoli, A methodology for controlling bias and fairness in synthetic data generation, Appl. Sci. 12 (9) (2022) http://dx.doi.org/10.3390/app12094619, URL: https://www.mdpi.com/2076-3417/12/9/4619.

[63] E. Ferrara, Fairness and bias in artificial intelligence: A brief survey of sources, impacts, and mitigation strategies, Sci 6 (1) (2024) http://dx.doi.org/10.3390/sci6010003, URL: https://www.mdpi.com/2413-4155/6/1/3.

[64] W.E. Lorensen, H.E. Cline, Marching cubes: A high resolution 3D surface construction algorithm, in: Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '87, Association for Computing Machinery, New York, NY, USA, 1987, pp. 163–169, http://dx.doi.org/10.1145/37401.37422.

[65] O. Sorkine, D. Cohen-Or, Y. Lipman, M. Alexa, C. Rössl, H.-P. Seidel, Laplacian surface editing, in: Proceedings of the 2004 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing, SGP '04, Association for Computing Machinery, New York, NY, USA, 2004, pp. 175–184, http://dx.doi.org/10.1145/1057432.1057456.

[66] A.B. Tsybakov, Introduction to Nonparametric Estimation, first ed., Springer Publishing Company, Incorporated, 2008.

[67] M. Haris, G. Shakhnarovich, N. Ukita, Task-driven super resolution: Object detection in low-resolution images, in: T. Mantoro, M. Lee, M.A. Ayu, K.W. Wong, A.N. Hidayanto (Eds.), Neural Information Processing, Springer International Publishing, Cham, 2021, pp. 387–395.

[68] S. Dodge, L. Karam, Understanding how image quality affects deep neural networks, 2016, arXiv:1604.04004, URL: https://arxiv.org/abs/1604.04004.

[69] I. Vasiljevic, A. Chakrabarti, G. Shakhnarovich, Examining the impact of blur on recognition by convolutional networks, 2017, arXiv:1611.05760, URL: https://arxiv.org/abs/1611.05760.

[70] M. Lébl, F. Šroubek, J. Flusser, Impact of image blur on classification and augmentation of deep convolutional networks, in: R. Gade, M. Felsberg, J.-K. Kämäräinen (Eds.), Image Analysis, Springer Nature Switzerland, Cham, 2023, pp. 108–117.

[71] A. Hamieh, C. Coufort-Saudejaud, A. Couffin, A. Liné, C. Frances, Temperature influence on MBS latex aggregate morphology, Colloids Surf. A: Physicochem. Eng. Asp. 676 (2023) 132139, http://dx.doi.org/10.1016/j.colsurfa.2023.132139, URL: https://www.sciencedirect.com/science/article/pii/S0927775723012232.

[72] F. Tao, M. Zhang, Y. Liu, A. Nee, Digital twin driven prognostics and health management for complex equipment, CIRP Ann 67 (1) (2018) 169–172, http://dx.doi.org/10.1016/j.cirp.2018.04.055, URL: https://www.sciencedirect.com/science/article/pii/S0007850618300799.

[73] T.H.-J. Uhlemann, C. Lehmann, R. Steinhilper, The digital twin: Realizing the cyber-physical production system for industry 4.0, Procedia CIRP 61 (2017) 335–340, http://dx.doi.org/10.1016/j.procir.2016.11.152, URL: https://www.sciencedirect.com/science/article/pii/S2212827116313129, The 24th CIRP Conference on Life Cycle Engineering.

[74] R. Rosen, G. von Wichert, G. Lo, K.D. Bettenhausen, About the importance of autonomy and digital twins for the future of manufacturing, IFAC- Pap. 48 (3) (2015) 567–572, http://dx.doi.org/10.1016/j.ifacol.2015.06.141, URL: https://www.sciencedirect.com/science/article/pii/S2405896315003808, 15th IFAC Symposium onInformation Control Problems inManufacturing.

[75] B. Rühle, J.F. Krumrey, V.-D. Hodoroaba, Workflow towards automated segmentation of agglomerated, non-spherical particles from electron microscopy images using artificial neural networks, Sci. Rep. 11 (1) (2021) 4942, http://dx.doi.org/10.1038/s41598-021-84287-6.

[76] L. Theodon, C. Coufort-Saudejaud, J. Debayle, GRAPE: A Stochastic geometrical 3D model for aggregates of particles with tunable 2D morphological projected properties, Image Anal. Stereol. 42 (1) (2023) 1–16, http://dx.doi.org/10.5566/ias.2875, URL: https://www.ias-iss.org/ojs/IAS/article/view/2875.

[77] L. Théodon, C. Coufort-Saudejaud, A. Hamieh, J. Debayle, Morphological characterization of compact aggregates using image analysis and a geometrical stochastic 3D model, in: 2023 IEEE 13th International Conference on Pattern Recognition Systems, ICPRS, 2023, pp. 1–7, http://dx.doi.org/10.1109/ICPRS58416.2023.10179036.

[78] L. Théodon, C. Coufort-Saudejaud, J. Debayle, A stochastic 3D model based on random graphs to characterize the morphology of compact aggregates using image analysis, in: 2024 IEEE 12th International Symposium on Signal, Image, Video and Communications, ISIVC, 2024, pp. 1–6, http://dx.doi.org/10.1109/ISIVC61350.2024.10577874.

[79] J. Cardona, C. Ferreira, J. McGinty, A. Hamilton, O.S. Agimelen, A. Cleary, R. Atkinson, C. Michie, S. Marshall, Y.-C. Chen, J. Sefcik, I. Andonovic, C. Tachtatzis, Image analysis framework with focus evaluation for in situ characterisation of particle size and shape attributes, Chem. Eng. Sci. 191 (2018) 208–231, http://dx.doi.org/10.1016/j.ces.2018.06.067, URL: https://www.sciencedirect.com/science/article/pii/S0009250918304469.

[80] Y. Zou, Z. Yu, B.V.K.V. Kumar, J. Wang, Domain adaptation for semantic segmentation via class-balanced self-training, 2018, arXiv:1810.07911, URL: https://arxiv.org/abs/1810.07911.

[81] C. Dong, C.C. Loy, K. He, X. Tang, Image super-resolution using deep convolutional networks, IEEE Trans. Pattern Anal. Mach. Intell. 38 (2) (2016) 295–307, http://dx.doi.org/10.1109/TPAMI.2015.2439281.

[82] Z. Wang, J. Chen, S.C.H. Hoi, Deep learning for image super-resolution: A survey, IEEE Trans. Pattern Anal. Mach. Intell. 43 (10) (2021) 3365–3387, http://dx.doi.org/10.1109/TPAMI.2020.2982166.

[83] F.-A. Croitoru, V. Hondru, R.T. Ionescu, M. Shah, Diffusion models in vision: A survey, IEEE Trans. Pattern Anal. Mach. Intell. 45 (9) (2023) 10850–10869, http://dx.doi.org/10.1109/TPAMI.2023.3261988.

[84] Z. Chen, Y. Wang, F. Wang, Z. Wang, H. Liu, V3D: Video diffusion models are effective 3D generators, 2024, arXiv:2403.06738.

[85] P.T. Junlin Han, VFusion3D: Learning scalable 3D generative models from video diffusion models, 2024, arXiv preprint arXiv:2403.12034.

[86] P. Dhariwal, A. Nichol, Diffusion models beat GANs on image synthesis, in: M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, J.W. Vaughan (Eds.), Advances in Neural Information Processing Systems, Vol. 34, Curran Associates, Inc., 2021, pp. 8780–8794, URL: https://proceedings.neurips.cc/paper_files/paper/2021/file/49ad23d1ec9fa4bd8d77d02681df5cfa-Paper.pdf.

[87] E. Ortiz Ortega, H. Hosseinian, M.J. Rosales López, A. Rodríguez Vera, S. Hosseini, Characterization techniques for morphology analysis, in: Material Characterization Techniques and Applications, Springer Singapore, Singapore, 2022, pp. 1–45, http://dx.doi.org/10.1007/978-981-16-9569-8_1.

[88] M.V. Glazoff, 3D morphological analysis and synthesis of industrial materials surfaces, Integr. Mater. Manuf. Innov. (2019) http://dx.doi.org/10.1007/s40192-019-00142-7.