SAM-IQ: Instance segmentation for 3D image data of particulate systems

(article starts on next page)

Powder Technology 484 (2026) 122847

Contents lists available at ScienceDirect

Powder Technology

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

SAM-IQ: Instance segmentation for 3D image data of particulate systems

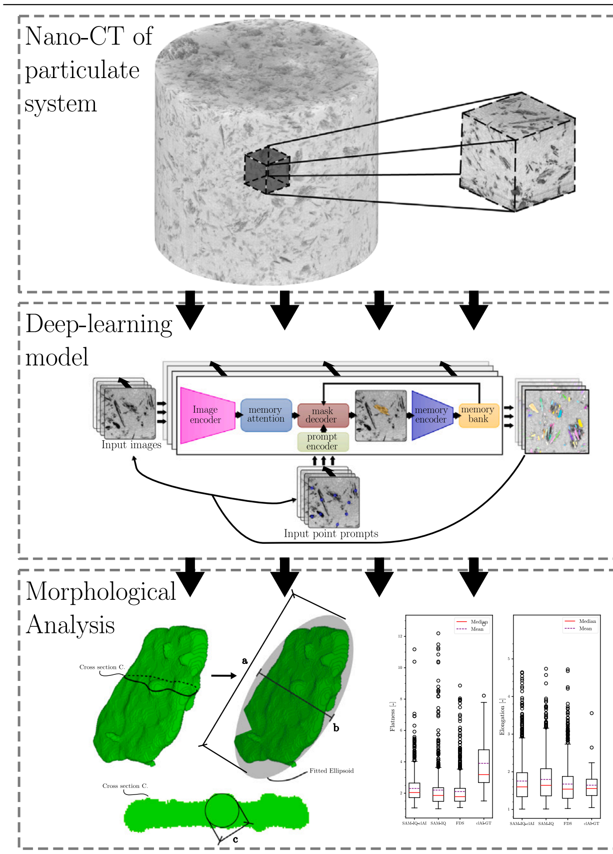
Theo Örtendahl [a,*], Angela Casarella [b], Olga Stamati [c], Jelke Dijkstra [a]

[a] Chalmers University of Technology, Department of Architecture and Civil Engineering, Gothenburg, 43218, Sweden
[b] Imperial College London, Department of Civil and Environmental Engineering, London, SW7 2AZ, UK
[c] Univ. Grenoble Alpes, CNRS, Grenoble INP, 3SR, Grenoble, 38000, France

HIGHLIGHTS

- SAM-IQ provides instance segmentation for 3D particulate systems.
- Morphological descriptors highly depend on the segmentation method.
- Fine-tuning with 2D data did not significantly improve segmentation results.
- Hybrid workflows are highly recommended in data scarce domains.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
X-ray nano-tomography
Kaolinite

ABSTRACT

Accurate 3D instance segmentation is essential for quantitative analysis in engineering and science, especially in systems like clay suspensions where particles are sub-micron, anisotropic, and densely packed. In this study,

* Corresponding author.
E-mail address: theoo@chalmers.se (T. Örtendahl).

https://doi.org/10.1016/j.powtec.2026.122847
Received 22 January 2026; Received in revised form 10 June 2026; Accepted 20 June 2026
Available online 13 July 2026

SAM 2
Morphology
3D instance segmentation

we introduce SAM-IQ: a spatial 3D implementation of Segment Anything Model 2 (SAM2). We then compare it against two other segmentation strategies applied to a synchrotron X-ray nano-CT image of Speswhite kaolin clay in suspension: (i) a traditional image processing pipeline (Filtering Distance map-based Segmentation, FDS) and (iii) SAM-IQ-clAI, a fine-tuned variant of the model trained on a clay-specific dataset of annotated 2D slices (clAI-D). Performance was assessed using voxel-level Dice similarity and particle morphology descriptors such as size, flatness, and elongation. Results reveal that each method yields significantly different particle counts and shape metrics, highlighting that segmentation choices directly influence particle statistics and modelling outcomes and underscoring the need to treat segmentation as an integral part of measurement rather than a neutral preprocessing step. SAM-based methods consistently outperform FDS in dense, low-contrast volumes, with fine-tuning improving detection of sub-micron particles but yielding only marginal gains in morphological accuracy. By combining automated prompt generation, adaptive termination, and minimal training requirements, SAM-IQ provides a scalable, data-efficient framework for 3D segmentation in noisy volumetric images, with applications extending well beyond clay systems.

1. Introduction

X-ray computed tomography (CT) has transformed modern engineering by enabling non-destructive 3D imaging of complex internal structures in materials. Originally developed for visualisation and identification (e.g., in medical diagnostics) CT has increasingly become a tool for quantitative metrology, used to extract geometrical and topological data from material systems [1]. In materials science and geomechanics, X-ray microtomography has been widely applied to coarse-grained materials such as sands and gravels, enabling grain-scale measurements of size, shape, and orientation of grains to support fabric analysis and inform constitutive modelling [e.g. 2–4].

However, imaging and analysing fine-grained or low-contrast particulate materials remains a fundamental challenge. When particle sizes fall below the micrometre scale, or when phase contrast is weak (e.g., in colloids, clays, or biological suspensions) the segmentation of individual particles becomes increasingly ambiguous. This problem is not discipline-specific, but spans diverse domains such as environmental science, soft matter physics, and biomedical imaging, where objects of interest are densely packed, irregularly shaped, or partially overlapping [5–8]. In these contexts, conventional approaches for morphological image processing, such as thresholding, watershedding, or distance maps, often fail to delineate object boundaries, leading to unreliable automated detection and biased quantitative analysis [9].

Recent advances in synchrotron-based X-ray computed tomography now offer the spatial resolution and contrast required, with voxel sizes down to tens of nanometres, enabling direct 3D visualisation of sub-micron particles. For instance, synchrotron-based phase-contrast nano-holo-tomography (X-ray nano-CT) enables the imaging of hydrated clay suspensions without destructive sample preparation, offering a unique opportunity to quantify natural microstructures in their native state [10]. Yet, even with such imaging advances, segmentation remains a major bottleneck to quantitative analysis, as densely packed, anisotropic, and low-contrast particles are difficult to identify reliably in 3D.

Deep-learning methods offer a promising route to address these challenges. In particular, large vision foundation models such as the Segment Anything Model (SAM) and its successor SAM 2 [11,12] have demonstrated remarkable generalisation capabilities across diverse visual domains [13–15]. Unlike conventional Convolutional Neural Networks (CNNs) [16,17], which require extensive domain-specific datasets, SAM and SAM 2 are pre-trained on large, diverse image datasets encompassing millions of annotated objects. This large-scale pretraining enables instance segmentation (object-wise segmentation at the pixel-level) with minimal domain-specific annotation, making these models especially attractive for scientific imaging, where labelled data are scarce [18,19].

However, neither SAM nor SAM 2 are optimised out-of-the-box for 3D high-resolution scientific imaging. Issues such as oversegmentation, false positives, and poor boundary detection persist, particularly in dense or noisy volumes [20]. Furthermore, SAM 2 was designed for video segmentation, treating time as a pseudo-third dimension, and has not been adapted for true spatial 3D segmentation in scientific volumes [21]. Recent adaptations of SAM 2 for medical 3D imaging [21] demonstrate that foundation models can perform well even with limited annotated data. However, in synchrotron-based nano-CT, the challenge is more fundamental: datasets are extremely rare and costly to generate, making large-scale or diverse training practically impossible. This motivates our focus on pre-trained, data-efficient segmentation that minimises dependence on annotation.

In this study, we introduce SAM-IQ (Segment Anything Model for Instance-based Quantification): a novel spatial 3D implementation of SAM 2 designed for quantitative segmentation of complex particulate systems. SAM-IQ extends SAM 2 beyond its original 2D or spatiotemporal design, enabling instance segmentation in noisy, low-contrast, and sub-micron 3D images. Unlike existing 3D SAM adaptations, SAM-IQ automates input-prompt generation, allowing efficient processing of large numbers of objects with minimal user intervention. This makes the framework particularly suitable for high-throughput or labour-intensive scientific imaging tasks. We demonstrate its capabilities using X-ray nano-holotomography of kaolin clay particles in suspension, an archetypal example of a densely packed, low-contrast system that challenges conventional approaches.

To assess the benefits of domain adaptation, we introduce a fine-tuned variant, SAM-IQ-clAI, trained on a curated clay dataset (clAI-D), and evaluate the benefits of domain adaptation by benchmarking its performance against the unmodified SAM 2 (SAM-IQ) and a traditional image segmentation pipeline (Filtering Distance map-based Segmentation, FDS).

Beyond pixel-level accuracy, assessed via the Dice similarity coefficient [22], we assess how segmentation strategies affect particle morphology and derived quantitative descriptors. In fields such as geomechanics, materials science, and colloidal physics, particle morphology metrics (size, aspect ratio, flatness, orientation) govern macroscopic phenomena including fabric anisotropy and mechanical strength. Thus, segmentation accuracy directly impacts the reliability of downstream analyses and particle-scale model calibration (e.g., Discrete Element Method (DEM) or Coarse-Grained Molecular Dynamics (CGMD)).

2. Materials and methodology

2.1. Kaolinite morphology and aggregation behaviour in suspension

Kaolinite is a 1:1 layered silicate mineral composed of alternating silica tetrahedral and alumina octahedral sheets. These layers form thin, platy particles that typically have a thickness ten times smaller than their longest side. In this study, Speswhite kaolin clay is used, a refined monomineral kaolin clay with a particle size ranging from 700 nm to 12 μm.

The small size and plate-like geometry lead to significant anisotropy in surface chemistry and interaction potential. In an aqueous suspension, particularly under acidic conditions, the kaolinite particles exhibit charge heterogeneity: negatively charged basal faces (silica), positively

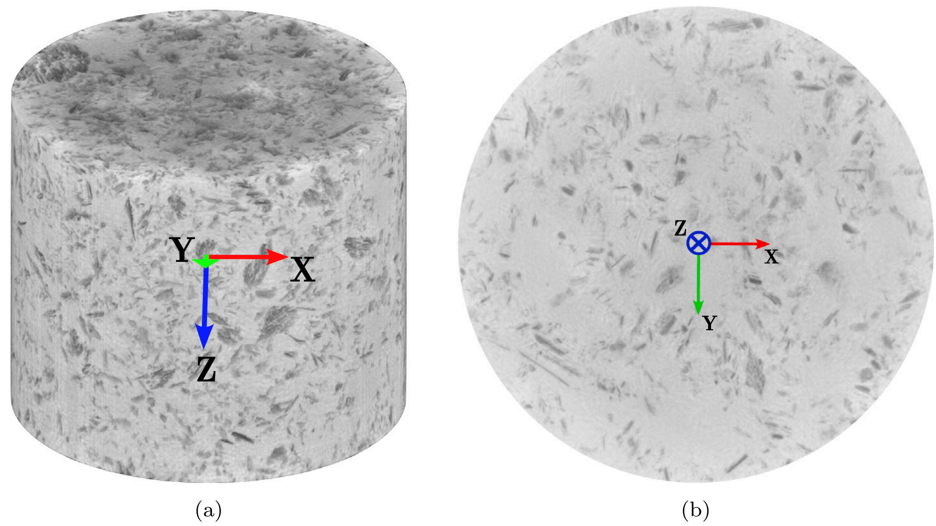


Fig. 1. (a) 3D rendering of the reconstructed kaolin clay nano-CT image at 50 nm voxel size and (b) a representative horizontal slice.

charged faces (alumina) and pH-dependent edge charges. This leads to distinct modes of interparticle association, commonly described as face-to-face (FF), edge-to-face (EF), and edge-to-edge (EE).

At low pH, the contrasting surface charges on different parts of the particle promote heteroaggregation, notably FE and EE arrangements of aggregates of FF particles, resulting in the formation of flocs. These structures are stabilised by a balance of electrostatic repulsion/attraction and van der Waals attraction as described by the DLVO theory. The flocculation alters the apparent morphology and measurable dimensions of the particles in suspension, as aggregation causes platelets to overlap or align, often complicating segmentation and analysis in tomographic imaging.

For this study, clay was suspended in demineralised deaired water adjusted to pH 4, to encourage cluster formation. After a short sedimentation period, smaller and hydrodynamically stable clusters remained suspended, providing suitable samples for high-resolution nano-CT imaging. These clusters present realistic challenges for instance segmentation, including overlapping particles, variable orientations, and heterogeneous local densities, making them an ideal test case for evaluating SAM-IQ and traditional segmentation methods.

2.2. Phase contrast X-ray nano-holo-tomography

A 3D tomographic scan of Speswhite kaolin clay suspended in acidic water (mixing fluid at pH = 4 ± 0.1) was acquired at the ID16B beamline at the European Synchrotron Radiation Facility (ESRF) in Grenoble, France [23]. A clay–water mixture at $10\,\mathrm{g\,L^{-1}}$ was pipetted into a quartz capillary tube of $300\,\mu\mathrm{m}$ diameter. After an initial sedimentation phase where the larger clay particles and aggregates settled at the bottom of the capillary tube, the scan was performed once the smaller, hydrodynamically stable clusters had reached equilibrium in suspension (i.e., when no significant differences were observed between two consecutive holotomographies lasting 40 min each).

The imaging technique used in ID16B is phase-contrast in-line holotomography [24], in which a complete acquisition consists of the combination of four tomography scans taken at different sample-detector distances (changing the magnification). This approach resolves improved phase modulation, ultimately increasing the contrast and spatial resolution in the reconstructed image. The measurements used a beam with an energy of $29.6\,\mathrm{keV}$. For each scan, 3023 projections with a pixel size of $50\,\mathrm{nm}$ and an FoV of $\Phi 128 \times 108\,\mu\mathrm{m}^2$ were recorded over $360°$ using a PCO edge 5.5 scientific CMOS camera with an exposure time of $15\,\mathrm{ms}$ per projection.

The final 3D volumes were obtained after a multi-step reconstruction process. First, the individual projections of the four different scans were aligned. Then, the phase retrieval calculation was done using an in-house Octave script [24] through an iterative calculation based on a Paganin-like approach with δ/β of 150, where δ is refractive index increment and β is the absorption index. Finally, the 3D reconstruction of the retrieved phase maps was performed using a filtered back projection algorithm with the ESRF software PyHST2 [25], yielding a final 3D volume of 2160 × 2500 × 2500 px.

The resulting dataset (Fig. 1) features a dense, heterogeneous arrangement of clay platelets (dark objects) floating in a light grey background (acidic water). Many platelets partially overlap and/or aggregate into clusters. Individual particles are generally thin and platy, but vary in size and orientation, making boundary detection and instance separation particularly challenging [26].

Features approaching the voxel size of the image cannot be accurately resolved [27]. According to the Nyquist–Shannon sampling theorem [28], at least $2\,\mathrm{px}$ are required to resolve a spatial frequency, although in practice more are needed to compensate for noise and imaging artefacts. Even when using a high-quality binarisation approach, Kerckhofs et al. [27] showed that only about 89% of voxels in micro-CT images correctly represent the underlying pore or material phase, while the remaining 11% do not overlap with the ground truth, reflecting intrinsic limits in spatial resolution and reconstruction fidelity. Following Wiebicke et al. [29], a minimum size threshold of $10\,\mathrm{px}$ is applied along the intermediate axis of the particle, which was chosen to account for the platy morphology of kaolinite platelets. In fact, the smallest axis may approach or fall below the voxel resolution, while the largest axis can be inflated by particle overlap or aggregation. This criterion ensures that only particles sufficiently resolved in the 3D volume are included in quantitative analyses.

2.3. Constructing clAI-D

To train and evaluate the SA-models on clay images, we created a specialised training dataset, denoted clAI-D. Given the challenges and time cost of generating fully labelled 3D volumes, the training set consists of annotated 2D slices. All 2D images used for training were fully labelled to avoid introducing bias from incomplete annotations.

The 2D training masks were generated using *SAMJ-IJ*, Garcia-Lopez-de Haro et al. [30], a SAM-plugin for ImageJ [31]. From randomly selected locations in the clay image stack, 100 cropped slices (500 × 500 px each) were prepared and manually labelled using SAM and bounding box input prompts, followed by minor edits to ensure accurate particle masks (see Fig. 2). To augment the dataset, each slice was rotated and flipped, effectively increasing the dataset size eightfold

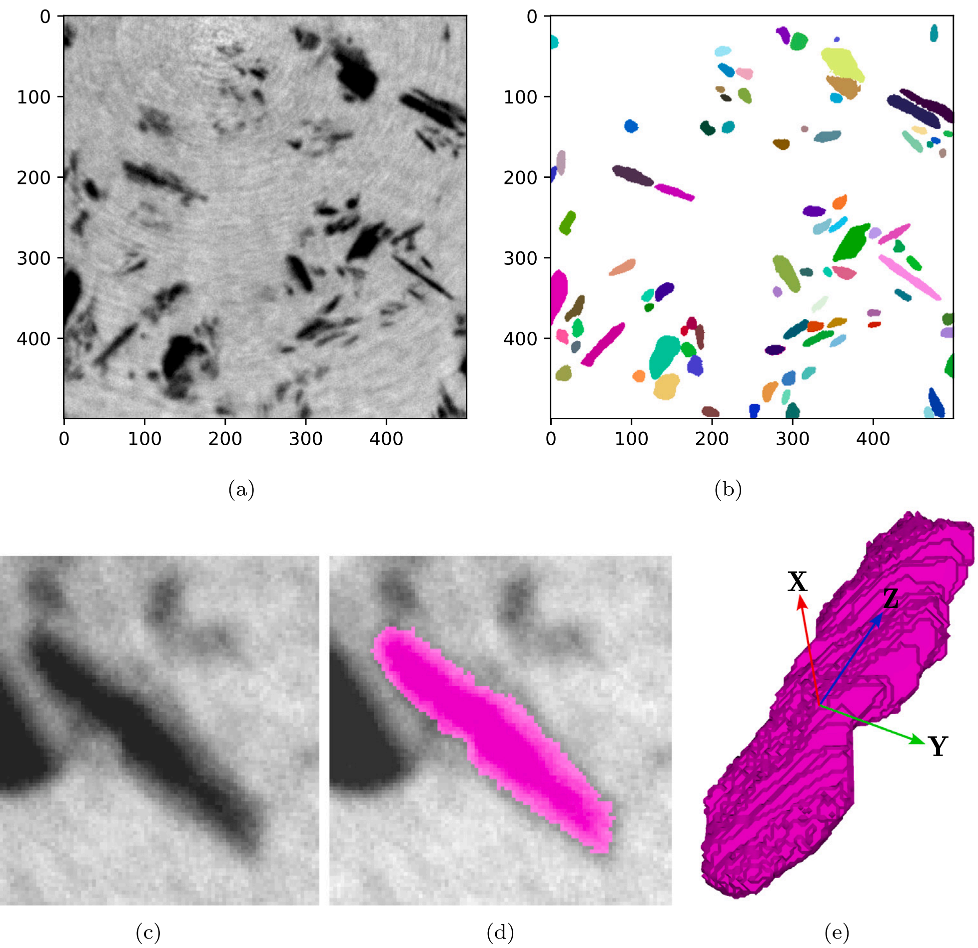


Fig. 2. (a) and (b): Representative slice from the clAI-D training set. (c), (d), and (e): Particle from the clAI-GT 3D Ground Truth set.

and resulting in a total of 58216 individual clay particle masks. Fig. 3 displays the number of labelled objects per 500 × 500 px image. Typical slices in clAI-D contain between 50 and 90 labelled objects.

2.4. Constructing clAI-GT

For evaluation, a smaller set of manually segmented 3D particles or "masklets" was used. The Ground Truth 3D masklets (hereby designated as clAI-GT) were generated using *QuPath extension SAM* [32], a SAM 2 plugin for QuPath. Each particle was initially enclosed in a 3D bounding box, and the resulting masks were manually processed to remove segmentation artefacts and incorrectly labelled regions. A critical step in the processing was the volume rotation, which ensured that each particle was accurately segmented along multiple dimensions. This procedure yielded 32 labelled 3D particles, which served as a ground truth for assessing segmentation accuracy and particle morphology metrics, providing a reference to compare SAM-IQ and FDS. Importantly, clAI-GT is not a ground truth for the entire sample's properties, but rather for the selected individual objects. One object from clAI-GT is presented in Fig. 3c, d, and e.

2.5. Instance segmentation

2.5.1. FDS

To provide a baseline approach without deep learning, we developed a conventional image processing pipeline named Filtering Distance map-based Segmentation (FDS).

The FDS workflow was selected as the benchmark after exploratory evaluation of several alternative instance segmentation strategies within the practical constraints of the dataset used in the current study. Due to the strong data scarcity in this field and the absence of established segmentation workflows for platelet-shaped clay particles, conventional deep-learning approaches such as U-Net [33] or DeepLabV3+ [34] were not feasible within the scope of this work because they require extensive annotated training datasets. The performance of flood-filling segmentation methods, combined thresholding approaches, and semi-supervised machine-learning segmentation, and the proposed FDS workflow were investigated. Among these methods, FDS consistently provided the best overall segmentation quality and particle separation performance and was therefore selected as the strongest conventional baseline for comparison with SAM-based segmentation. See Fig. A.14 for a visual overlay comparison of the segmentation results.

The FDS method combines standard filters and distance maps to segment particles from a binarised 3D image. The method consists of four key steps. Firstly, a difference of the Gaussian filter (DoG) is applied to enhance high-gradient and high-intensity regions, improving particle-edge contrast [35]. An intersection of a global and local thresholding method is used: (i) global Otsu thresholding effectively removes large background areas [36]; (ii) the local Bernsen method separates particles locally (i.e., within clusters and closely neighbouring particles) [37]. A distance map is computed from the binarised image, followed by a second global Otsu threshold to refine separation. Lastly, a simple connected components algorithm is applied, which segments and labels the binarised 3D image. The total computational time is 37 358 ms when running the FDS pipeline on a 500^2 pixel image using an Intel Core i9-10900 CPU at 2.80 GHz and an NVIDIA GeForce RTX 2070 SUPER. Full details on the implementation are provided in Appendix A.

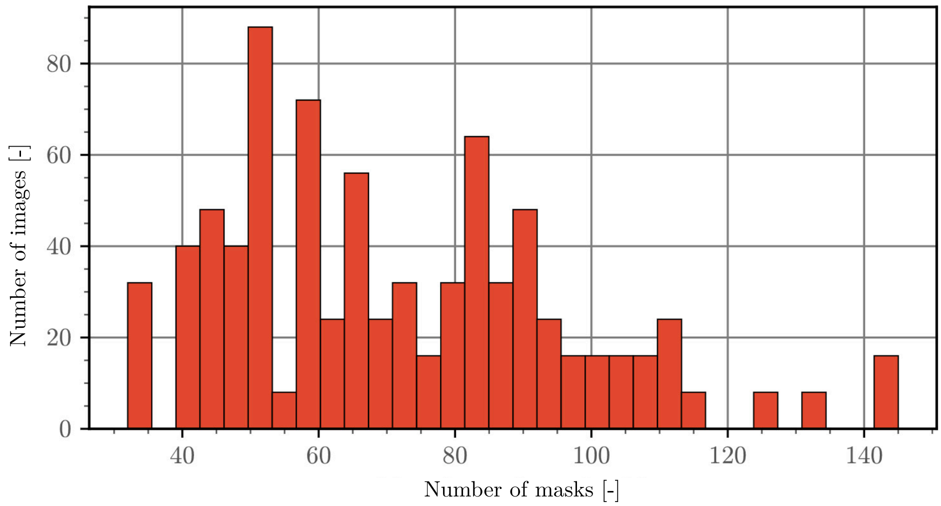


Fig. 3. The number of objects in each training 2D image of clAI-D.

2.5.2. SAM-IQ implementation

The Segment Anything Model 2 (SAM 2) provides instance segmentation via deep learning, without traditional pixel-level algorithms [11, 12]. Under ideal conditions, SAM 2 is capable of generating accurate particle masks and masklets using automated input prompt generation; however, complex datasets such as the one in this study often require guided input.

As represented in Fig. 4 (a), the SAM 2 architecture consists of three main components: an *image encoder*, a *prompt encoder*, and a *mask decoder*. The image encoder processes the input images by extracting important features, such as edges and shapes [12]. The prompt encoder takes prompts as input to guide segmentation, which can consist of points or bounding boxes (sparse input) or an initial mask (dense input). The frame embedding created by the image encoder and tokens from the prompt encoder are passed to the mask decoder, which outputs a segmentation mask for the given image.

For 3D and video applications, SAM 2 incorporates *memory attention*, a *memory bank*, and a *memory encoder* to retain contextual information across frames/slices. This allows the model to propagate learned features across frames/slices, improving consistency in sequential segmentation and preserving spatial structure. However, SAM 2 is trained on video data (2D + temporal dimension) and does not perform "true" 3D segmentation. Its performance depends on the direction of propagation, which can lead to orientation-dependent artefacts, a key limitation in volumetric scientific imaging.

To overcome this, we developed a novel spatial 3D implementation of SAM 2 (SAM-IQ), enabling volumetric segmentation of 3D scans containing numerous discrete objects. In SAM-IQ, positive point prompts are generated automatically based on intensity minima in each slice. Segmentation proceeds slice by slice, with SAM propagating masks through the volume while maintaining 3D spatial coherence. The core novelty of SAM-IQ lies in its adaptive termination criteria, which dynamically control the propagation of masks to minimise over-segmentation, z-axis artefacts, and cumulative errors during sequential segmentation. These termination conditions are governed by three user-defined thresholds that can be adjusted for different imaging contexts and particle morphologies: maximum 2D object area (area threshold), minimum object thickness (minimum axis), and maximum object overlap (overlap threshold). Segmentation terminates when any of the following conditions are met:

(i) Excessive particle drift (centroid movement > 2× equivalent disc diameter);
(ii) Excessive particle size (object area > area threshold);
(iii) Excessive overlap with previous object masks (overlap > overlap threshold);
(iv) No valid particle mask is detected.

When termination occurs, the resulting mask is stored in the master volume with labels, and the algorithm proceeds to the next input prompt. If a prompt corresponds to an already segmented particle, segmentation is skipped to prevent redundancy. This iterative process introduces error propagation, meaning that the quality of new particle masks is affected by the quality of previous masks.

After segmentation, the masks undergo post-processing to remove segmentation artefacts and increase segmentation quality. An anisotropic morphological opening and closing is applied to correct artefacts specific to the z-direction (Fig. 5), followed by an oversegmentation correction script [38]. These steps significantly reduce false boundaries and improve inter-slice continuity. However, it should still be noted that spatial discrepancies are still present, resulting in partly anisotropic segmentation results. For details of these steps, see Appendix B. Fig. 4 illustrates the overall SAM 2 architecture along with the iterative loop-back feature of finished masks implemented in SAM-IQ for 3D segmentation.

The inference speed of SAM-IQ on a 3D image of Speshwhite kaolinite particles with a size of 500^3 voxels is 2 s per object, which equals to 3 h to 4 h to segment the entire subvolume.

2.5.3. clAI-D fine-tuning

To assess whether domain-specific adaptation improves segmentation performance in complex, high-density granular materials, SAM-IQ-clAI was fine-tuned on the clAI-D training set. Apart from this fine-tuning step, the architectures of SAM-IQ and SAM-IQ-clAI are identical. The image encoder, mask decoder, and prompt encoder were jointly optimised on annotated 2D slices, using positive point prompts scattered across each mask. At each iteration of the fine-tuning process, an annotated 2D slice was randomly selected from the dataset. [12].

Training was performed on an Nvidia A100 (40 GB VRAM) for 80 000 iterations, achieving a predicted mean Intersection of Union (IoU) of 0.725 after 4 h. AdamW, a modification of the previous Adam optimiser, was used during training [39]. The learning rate and weight decay coefficient were set to 10^{-5} and $4 \cdot 10^{-5}$ respectively. Mixed precision training was used to speed up the process [40]. The predicted mean IoU values for all iterations are plotted in Fig. 6.

Linear regression between 40000 and 80000 iterations indicated an IoU gain of 0.0025 per 10000 iterations, suggesting that additional training would only yield marginal benefits. Full details of the loss function and training scripts are provided in Appendix C.

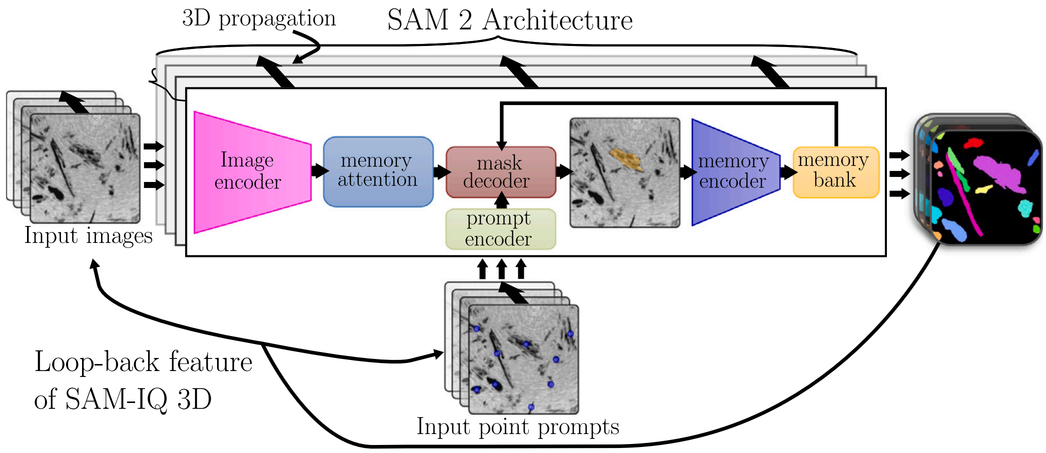


Fig. 4. Flowchart showing the main components of the SAM architecture.

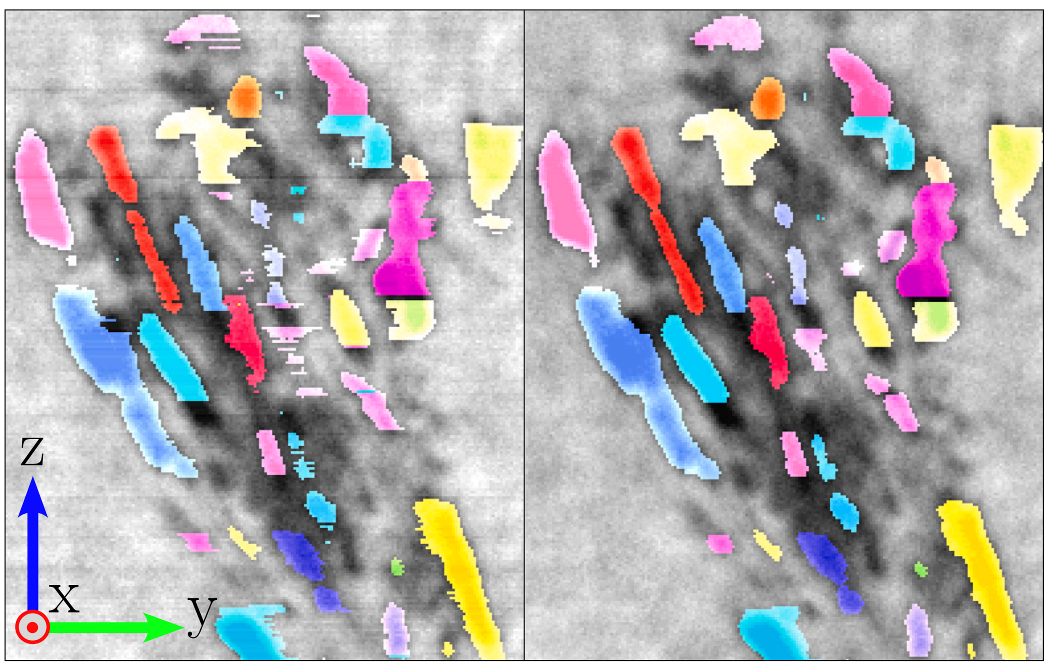


Fig. 5. SAM-IQ labelled objects before and after morphological binary operations are applied for correction of z-axis artefacts.

2.6. Evaluation methods

Segmentation performance is evaluated using two complementary approaches:

1. Voxel-level agreement, quantified by the Dice similarity coefficient (DSC) [22]. The DSC measures volumetric overlap between predicted and clAI-GT masklets, where a higher DSC indicates closer alignment.
2. Morphological accuracy, assessed through particle shape descriptors such as flatness and elongation. The morphological descriptors in this case quantify how well the segmented geometry reproduces the anisotropic features of clay platelets.

To ensure a consistent basis for comparison, the clAI-GT masklets were paired with the corresponding predicted label from SAM-IQ-clAI, SAM-IQ, and FDS. Metrics were computed on a per-particle basis and aggregated to obtain method-level statistics. Additionally, the total number of particles detected for each method was used as a performance indicator.

The selection of appropriate shape description is not trivial, leading to a multitude of potential descriptors [41]. The morphological characterisation of the kaolin clay platelets was performed using the Software for Practical Analysis of Materials (SPAM) [38] and the ImageJ [31] plugins, 3D Suite [42] and MorphoLibJ [43].

The particle size was derived from the segmented volume or the equivalent spherical radius, while the shape was described by the flatness and elongation of the fitted ellipsoid. The ellipsoid fitting was performed using a 3D-moments-based method found in 3D Suite [42]. Following one variation presented by Angelidakis et al. [44], flatness is defined as the ratio of the intermediate to the shortest ellipsoid axis b/c, and elongation as the ratio of the longest to intermediate axis a/b, where $c \leq b \leq a$. A sketch of the best fitting ellipsoid for one of the particles in clAI-GT is given in Fig. 7.

As shown in Fig. 7, the smallest axis, c, was computed as the diameter of the largest inscribed sphere fully contained within each particle, derived from a distance map of the binary volume. This approach mitigates the tendency of conventional ellipsoid fitting to overestimate particle thickness in platy geometries and provides a more accurate local measure of clay platelet thickness. The larger axes, a and

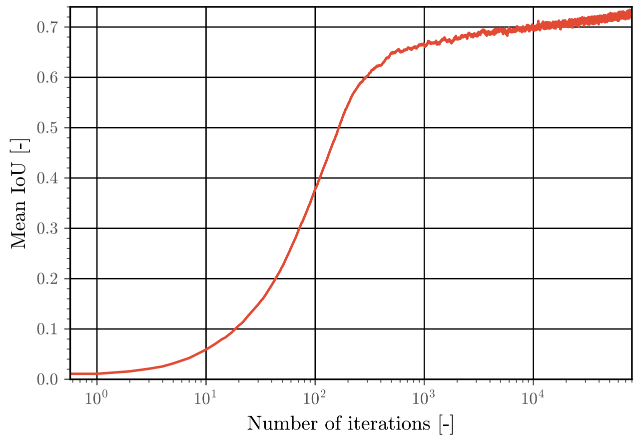


Fig. 6. Mean intersection of union (IoU) over number of iterations during SAM-IQ-clAI training on the clAI-D set.

b were taken as the major and intermediate axes of the best-fit ellipsoid encapsulating each particle. All ellipsoid fitting and axis calculations were performed using 3D Suite and MorphoLibJ [42,43].

Importantly, it should be mentioned that shape descriptors that are dependent on boundary condition (e.g. flatness and elongation) are highly affected by morphological operations, image resolution, and thresholding technique [45]. Moreover, for low-density pixel environment, this fact holds even stronger [45].

The analysis was repeated on a volume that was rotated 90° around the x-axis, to assess the impact of the direction of propagation on the segmentation results and the measured particle properties. Each corresponding particle was matched using the centroid distance, where any two particles with centroids closer than 5 pixels are considered matching particles. The measured particle properties, i.e., volume, flatness, and elongation, were cross-plotted in Fig. 13.

3. Results

3.1. Visual comparison of the segmentation approaches

A visual comparison provides an intuitive overview of the segmentation task and highlights the challenges inherent to nano-CT imaging of clay suspensions. Fig. 8 shows representative slices from a 3D nano-holotomography scan of Speswhite kaolin clay in an acidic suspension, overlaid with segmentation results obtained using SAM-IQ-clAI, SAM-IQ, and the Filtering Distance map-based Segmentation (FDS) pipeline.

Each segmented particle is coloured uniquely for clarity, although colours do not correspond across methods. The result highlights the differences in how each technique interprets and partitions the same particle field. For example, a region showing multiple colours (labels) for one method but a single colour (label) for another indicates oversegmentation or undersegmentation, respectively.

Visually, it is clear that at the current state all methods fail to identify all particles, and in some cases label background pixels erroneously. The three segmentation approaches exhibit distinct behaviours. SAM-IQ-clAI, shows improved coverage, detecting slightly more particles overall. This is likely due to its ability to separate closely spaced particle pairs and groups (see the three dense regions in the first row of Fig. 8). Both versions of SAM produce similar visual outputs in general and outperform the traditional FDS method in particle detection and continuity. This can also be appreciated in the slices on the second row of Fig. 8 where the correspondent z-y slices perpendicular to the x–y slices in the first row are shown.

A few limitations of the SA-models can be identified from visual inspection. Firstly, they do not always follow particle boundaries precisely, producing a “blobbing” effect where neighbouring particles merge or appear overly rounded. This behaviour is visible even in the xy-plane, for instance in the beige and yellow particles in the top left corner of the SAM overlays in the first row of Fig. 8.

A second limitation arises from the slice-wise propagation, which leaves visible artefacts in 3D reconstructions. In the perpendicular z–y slices (Fig. 8, second row), this manifests as straight edges along the top and bottom of particle masks, indicating direction-dependent segmentation. Although morphological post-processing, such as anisotropic opening and closing (see Fig. 5), mitigates this effect, residual discontinuities remain (for example, the straight edges along the upper and lower boundary of many particle masks). In contrast, FDS, being a true 3D pipeline, produces more isotropic results, though at the expense of segmentation precision in dense particle regions.

While the traditional image processing pipeline (FDS) is a true 3D-method, it has some issues. The masks preserve the overall shape of isolated particles relatively well, but the method struggles more than the SA-methods in high-density regions where thresholding and distance-map operations fail to resolve boundaries. The local Bernsen threshold enables partial segmentation in such regions, but often erodes large particles (see the light blue particle in row 2 for FDS in Fig. 8).

Overall, these qualitative differences underscore the need for a systematic quantitative comparison. The following sections evaluate the three approaches in terms of voxel-level accuracy, morphological fidelity, and overall agreement with manually labelled reference data.

3.2. Instance segmentation performance

The Dice Similarity Coefficient (DSC) between predicted masks and the 32 clAI-GT particles was computed for quantitative assessment of the accuracy of each segmentation method. The results of the three methods are shown in Fig. 9.

Both SA-models achieved higher mean and median Dice coefficients (DSC) than the FDS method, indicating closer agreement with clAI-GT. All three approaches successfully identified all particles in the reference dataset, but the SA-models produced more accurate delineations overall. Their performance was nearly identical, with mean DSC values of 0.71 and median values of 0.81 (SAM-IQ-clAI) and 0.82 (SAM-IQ-base), respectively. In comparison, the FDS yielded slightly lower but more consistent results, reflected by a narrower interquartile range and mean and median DSCs of 0.66 and 0.70, respectively.

A minor note of caution is warranted: because the clAI-GT labels were partially generated using a SAM-based tool before manual correction, the DSC values may be slightly biased in favour of the SAM-IQ methods. Nonetheless, the trends in mask quality observed visually (Fig. 8) are fully consistent with the DSC results.

3.3. Morphological results

Although the Dice coefficient provides a useful metric of segmentation accuracy at the voxel level, it does not capture how differences in segmentation propagate to measured particle properties. The latter is essential for quantitative studies of materials based on 3D image data. Therefore, the impact of each segmentation method on subsequent morphological descriptors will be investigated, including comparison against the manually segmented reference 3D set (clAI-GT). The same particles are analysed for each metric.

3.3.1. Size

Table 1 summarises the statistics on the volume properties for each segmentation method, complemented by the size distributions shown in Fig. 10. The volume analysed that is cropped from the original data is of size 500 × 500 × 500 px, corresponding to 15 625 μm^3. In terms of the total number of objects, SAM-IQ-clAI detects 1270 objects,

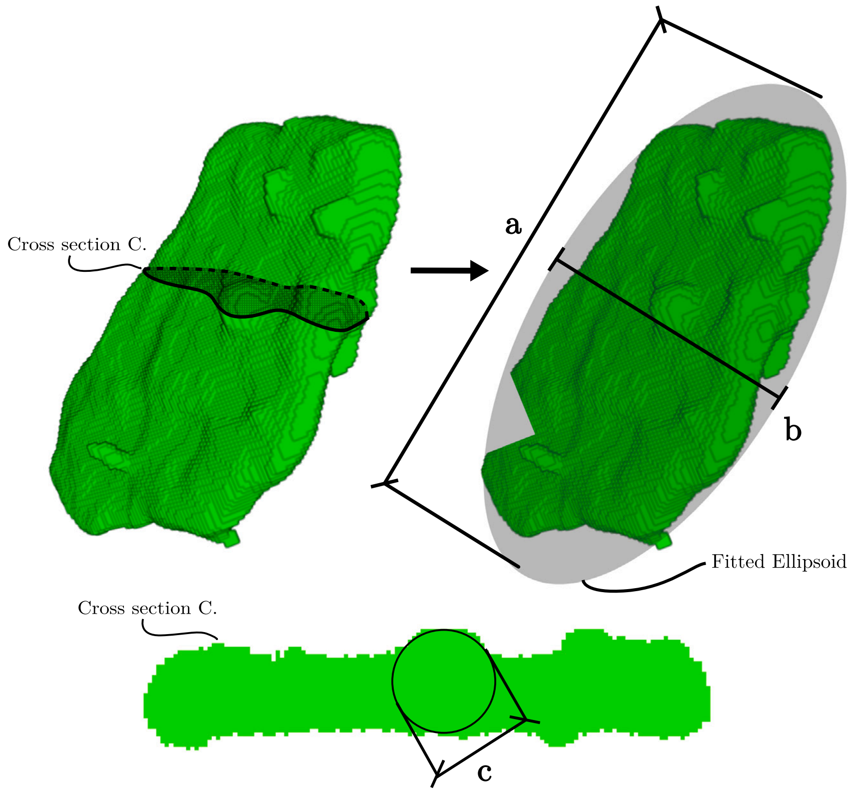


Fig. 7. Segmented particle (green volume) with fitted equivalent ellipsoid (grey transparent volume). The two major axes of the ellipsoid (a and b) are indicated, while the diameter of the largest inscribed sphere (c) represents the particle thickness.

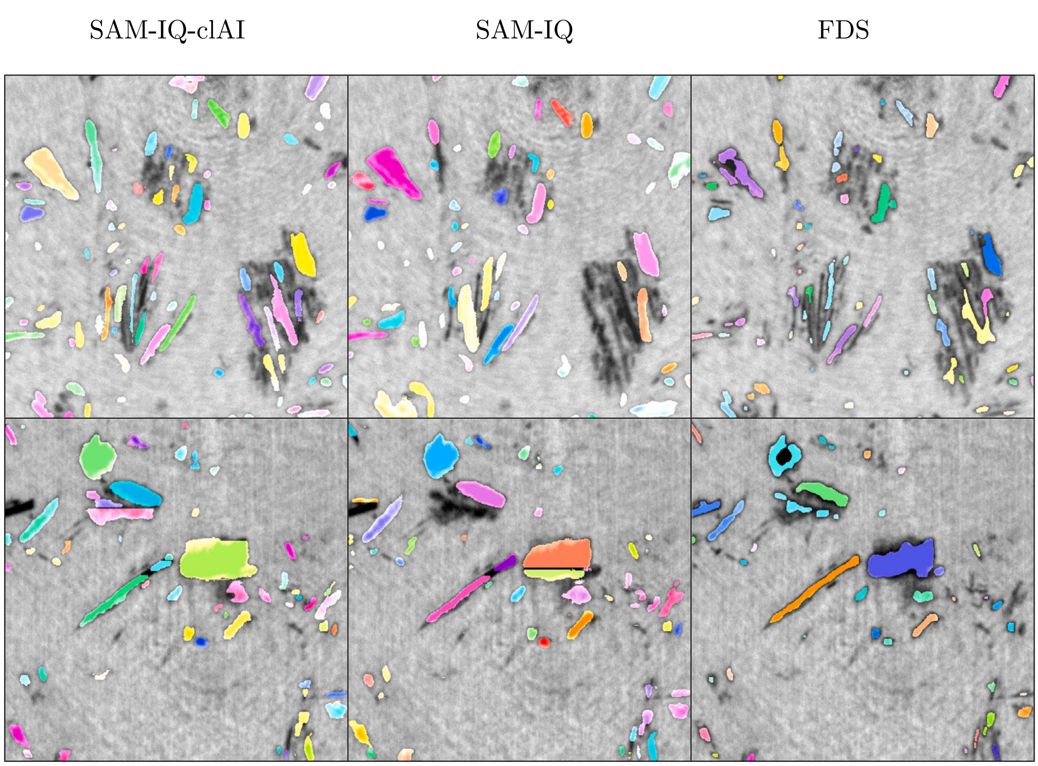
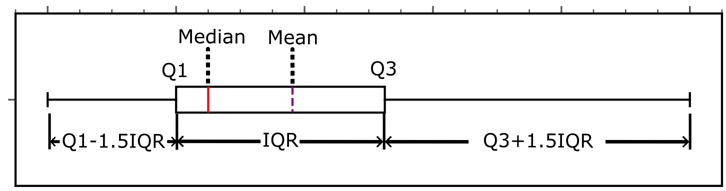
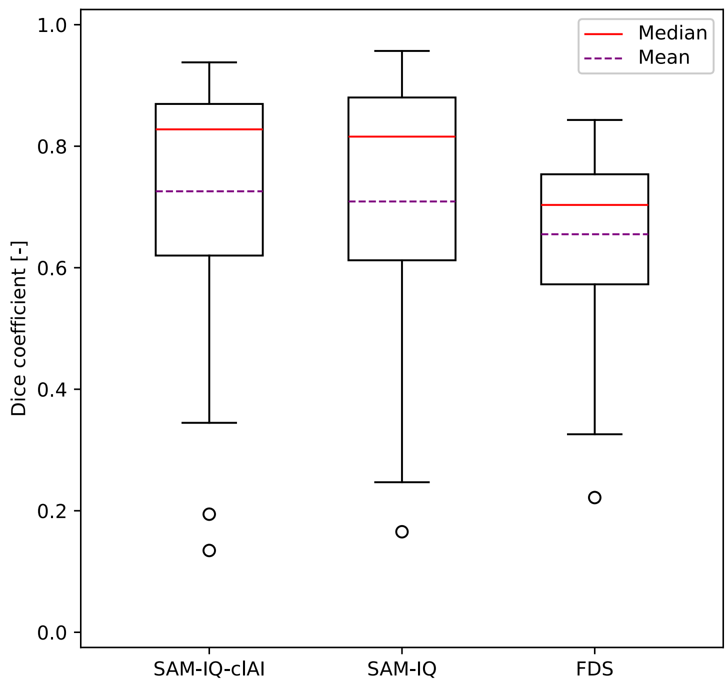


Fig. 8. Selected slices showing the segmentation performance of SAM-IQ-clAI, SAM-IQ, and the FDS. The segmented masks are overlayed onto the respective greyscale slice of a subset of the original holotomography volume of Speswhite Kaolin clay in suspension. The first row shows a slice in the x–y plane while the second row shows a slice in the z–y plane.



(a) Definition of box and whisker plot.



(b) Dice coefficients.

Fig. 9. The Dice coefficients (b) for SAM-IQ-clAI, SAM-IQ, and FDS. The box extends from the first to the third quartile of the data. The whiskers extend to the farthest point within 1.5 times the width of the box added to the box width, as described in (a). Outliers are presented as circles.

followed by SAM-IQ (1027), and lastly, FDS (918). The SA-models produce comparable size distributions, but we find several differences when compared to FDS. All models show a disparity between mean and median particle volume further reinforcing that the distribution is uneven.

The manually labelled reference set (clAI-GT) shows a much higher median volume (4.00 μm^3), reflecting the selection bias introduced by the user, who prefers clearly identifiable and well-separated (therefore larger) particles during manual labelling.

Fig. 10 presents a histogram of particle volumes, with the horizontal axis representing particle volume (in μm^3) and the vertical axis showing particle count on a logarithmic scale. The use of a logarithmic y-axis highlights the broad size distribution and makes the occurrence of larger, less frequent particles more apparent. All three segmentation approaches (SAM-IQ, SAM-IQ-clAI, and FDS) reveal a particle size distribution heavily skewed toward small volumes, with the majority of detected particles falling below 1.00 μm^3 (see Fig. 10). This trend is consistent across methods and aligns with findings from other experimental techniques in the literature (e.g., Static Light Scattering (SLS) in [46] and Dynamic Light Scattering (DLS) in [47] for the same clay type).

Table 1
Mean, median, and total volume of all labelled objects as well as number of objects for SAM-IQ-clAI, SAM-IQ, FDS, and clAI-GT. Number of objects and total volume is also presented for all objects below 1 μm^2.

	Mean vol [μm^3]	Median vol [μm^3]	Total vol [μm^3]	Objects [–]	Total vol ≤ 1 [μm^3]	Objects $\leq 1\ \mu m^3$ [–]
SAM-IQ-clAI	1.31	0.50	1660	1270	362	916
SAM-IQ	1.65	0.58	1700	1027	277	689
FDS	0.63	0.19	578	918	198	825
clAI-GT	6.99	4.00	(224)	32	(0)	5

When comparing the statistics in terms of size and particle count for SAM-IQ and SAM-IQ-clAI in Table 1, the fine-tuned model (SAM-IQ-clAI) segments approximately 20% more particles in total, primarily due to improved detection of small particles below 1.00 μm^3. In this size range, SAM-IQ-clAI identifies roughly 25% more particles than SAM-IQ, indicating enhanced sensitivity to fine features. Despite this higher particle count, the total segmented particle volume differs by only about 2%, with SAM-IQ yielding a slightly larger cumulative volume.

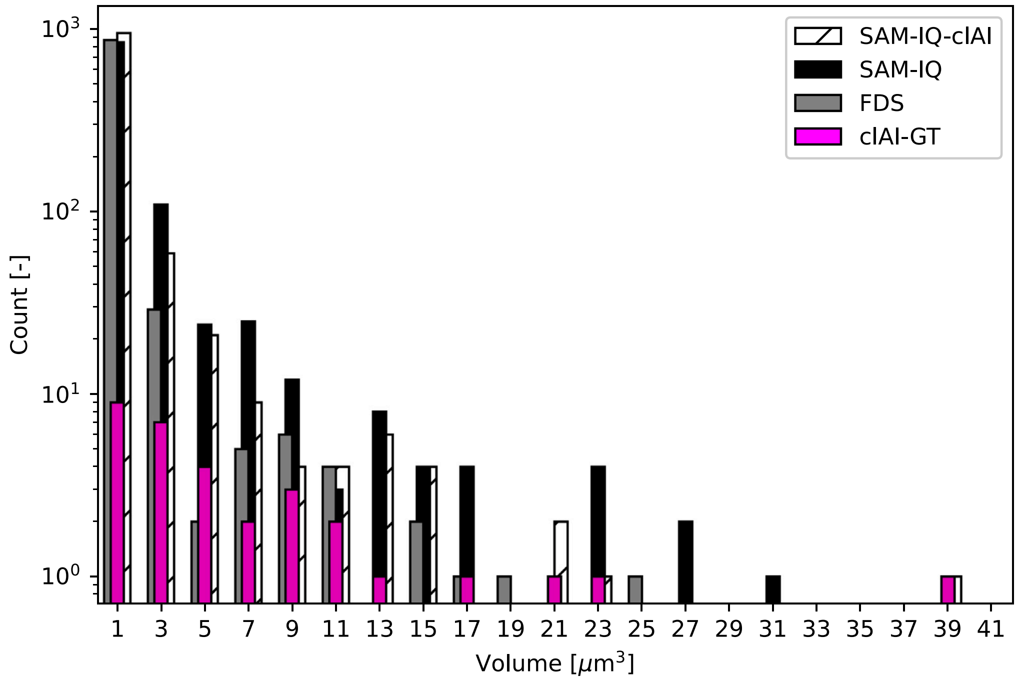


Fig. 10. The distribution of individual particle volumes and equivalent sphere diameters as segmented by SAM-IQ-clAI, SAM-IQ, and FDS.

This implies that SAM-IQ tends to segment generally larger objects, a trend reflected in its mean particle volume, which is about 20% higher than the one of SAM-IQ-clAI. The tendency appears to be limited to the particles larger than $1.00\,\mu m^3$, as the 25% higher particle count in SAM-IQ-clAI below this threshold is accompanied by a comparable 25% increase in total volume. Overall, these results suggest that SAM-IQ may merge small adjacent particles into larger volumes, leading to undersegmentation and/or SAM-IQ-clAI occasionally exhibits oversegmentation of individual particles. Examples of both artefacts can be seen in Fig. 8.

3.3.2. Shape

Particle shape was quantified using flatness and elongation of the ellipsoid fitting each particle. The results for the three methods are given in Fig. 11.

For flatness and elongation, there is a general agreement among the three methods. The platy nature of kaolinite is captured to some degree as median flatness values range from 1.78 to 2.04. However, the values for flatness are consistently lower than those reported in the literature for kaolinite (5–10 in [e.g. 48]). SAM-IQ-clAI produced the highest median flatness value at 2.04, indicating an effectiveness in capturing the platy geometry of the kaolin particles. In contrast, SAM-IQ and FDS exhibited lower median flatness values at (1.86 and 1.78 respectively), indicating a tendency to segment particles as more compact volumes. For the FDS, this is likely a result of the heavy filtering applied during preprocessing within the FDS method, which may suppress fine edges and smooth out thin structures.

Notably, the reference dataset (clAI-GT) exhibits higher flatness values (i.e., thinner particles) than all automated methods. Beyond segmentation effects, this also reflects the annotation bias previously discussed in the size analysis: the reference set contains predominantly larger, visibly platy particles, whereas the SAM-based methods detect many small, rounded particles that naturally exhibit lower flatness.

In contrast to flatness, the elongation values were remarkably consistent across all segmentation methods including the reference GT set. As shown in Fig. 11b, the median elongation values fall within a narrow range across SAM-IQ-clAI, SAM-IQ, FDS, and the manually labelled particles (red lines ranging from 1.54 to 1.60). This value is also consistent with the literature reports [48], as the in-plane geometry of kaolinite platelets is typically associated with a hexagonal-shaped morphology, where the two lateral dimensions are nearly equal, resulting in elongation values close to 1.

The 32 clAI-GT masklets segmented with each of the three different approaches (SAM-IQ-clAI, SAM-IQ, and FDS) are compared in Fig. 12, to quantify the accuracy of the morphological descriptors. This direct comparison isolates segmentation effects on identical reference particles, providing a consistent basis for evaluating descriptor reliability.

Despite the presence of some segmentation artefacts, the median and mean values for SAM-IQ-clAI and SAM-IQ remain highly similar to clAI-GT in terms of flatness, demonstrating that the calculated shape descriptors are captured across segmentation approaches. However, SAM-IQ exhibits notably larger whiskers in elongation compared to the reference data. This may suggest that SAM-IQ occasionally merges neighbouring labels (undersegmentation), producing more rounded, less elongated particles, as visible in the top row of Fig. 8. This behaviour is consistent with the observations from the particle size analysis, where SAM-IQ was found to segment fewer but larger objects.

3.4. SAM-IQ rotational cross correlation

Fig. 13 shows the volume, flatness, and elongation measured using SAM-IQ with two different propagation directions. Fig. 13(a) shows that volume is captured well with a slight bias towards the Z-axis. In contrast, there is a larger disparity for elongation and flatness, where only a weak correlation is found.

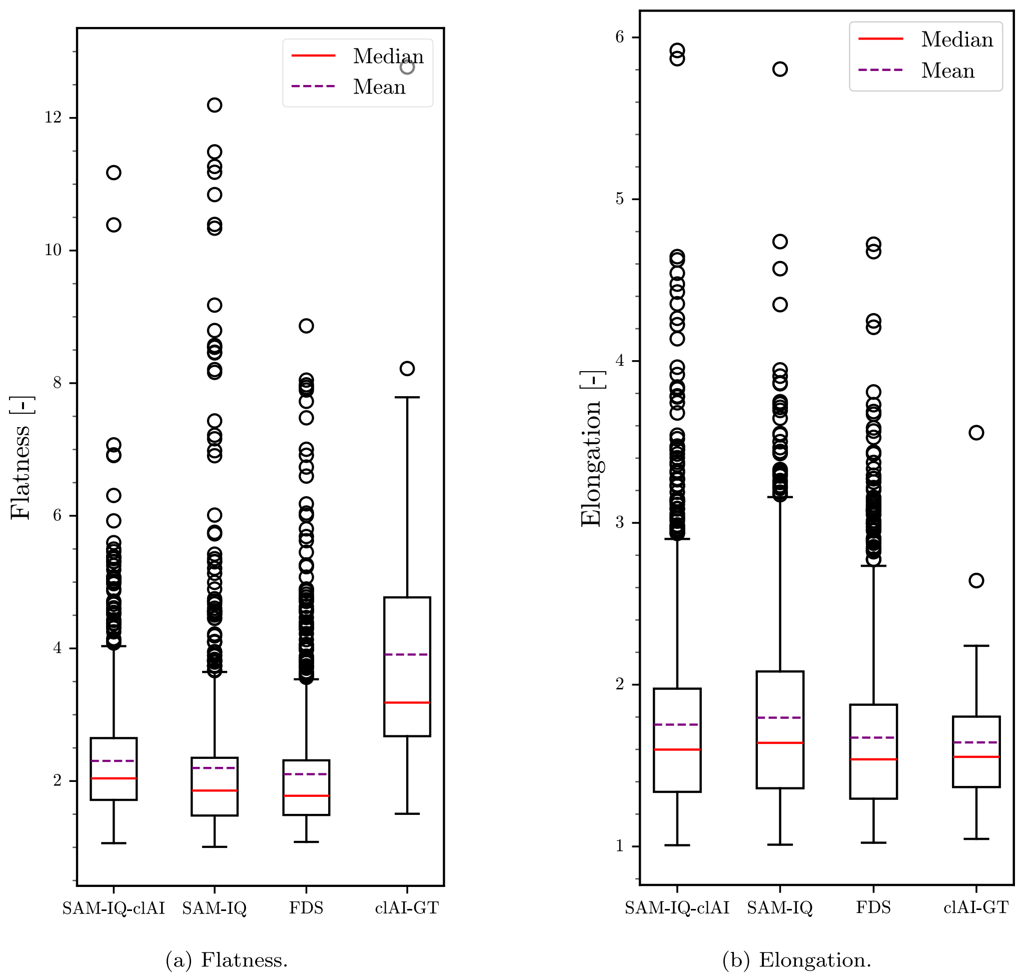


Fig. 11. The flatness (b) and elongation (c) of the output particles of SAM-IQ-clAI, SAM-IQ, and FDS also compared to the reference clAI-D.

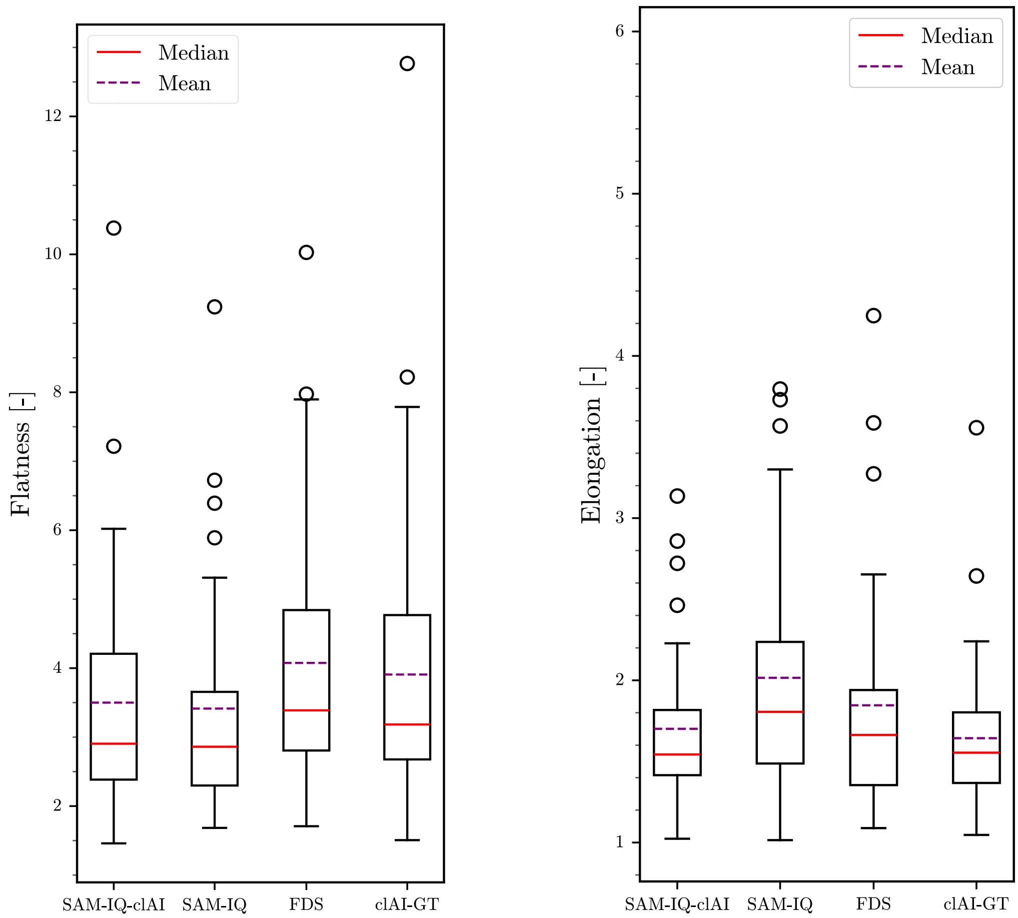
4. Discussion

The three segmentation approaches evaluated in this study (SAM-IQ, SAM-IQ-clAI, and FDS) exhibit distinct behaviours in particle detection, voxel-level accuracy, and morphological characterisation. For image-based materials characterisation, segmentation quality directly governs what can be measured, inferred, or simulated.

Across the dataset, the SAM-based models consistently outperform the conventional Filtering Distance map-based Segmentation (FDS) method in terms of particle count. The key behavioural distinction does not lie in whether particles are found (since all three methods successfully identify ground-truth objects), but rather on how they are partitioned. SAM-IQ-clAI identifies the largest number of particles, especially in the sub-micron range. However, its higher count contrasts with a similar total segmented volume compared to SAM-IQ. This divergence in count–volume balance signals two competing mechanisms: (i) a tendency of SAM-IQ to merge adjacent small particles (undersegmentation), consistent with the high dispersion (large whiskers) in the elongation results observed for clAI-GT particles (Fig. 11); (ii) a tendency of the fine-tuning (SAM-IQ-clAI) to increase fragmentation, introducing mild oversegmentation in some regions. Both tendencies can be observed in Fig. 8. These opposing tendencies are small but scientifically relevant as they emphasise that segmentation choice modulates the size spectrum and thereby influences the extracted morphology.

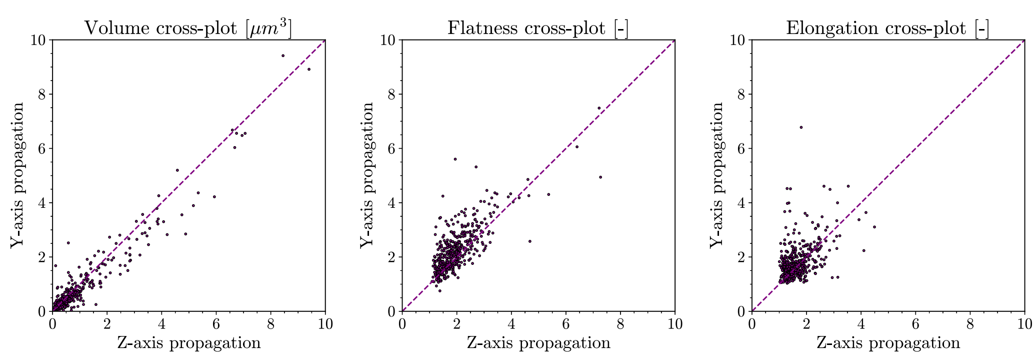
Accuracy metrics reinforce this picture. Both SAM-based implementations outperform FDS in voxel-level fidelity, but the near-equivalence between SAM-IQ and SAM-IQ-clAI suggests that domain-specific fine-tuning does not always translate into substantive volumetric gain. This is a key observation for data-scarce scientific domains. The FDS pipeline, although attractive as a fully 3D method, systematically segments smaller particles on average, reflecting the strong smoothing and edge-filtering effects of its preprocessing pipeline. This behaviour also underscores the need for AI-based segmentation approaches when dealing with noisy, densely populated 3D images, where conventional thresholding and morphological operations struggle to resolve fine boundaries and overlapping structures.

Morphological descriptors further illuminate the consequences of these behaviours. Elongation remains stable across all methods. This is likely a consequence of the intrinsic geometry of clay: particles often exhibit similar in-plane dimensions ($a \simeq b$), resulting in elongation values that remain close to 1 regardless of the size of the particle. Flatness, however, is more sensitive to segmentation: smaller particles generally tend to be more rounded and, therefore, thicker. Moreover, differences in flatness across methods partly reflect the difficulty of reliably estimating the smallest axis (c), which is the dimension most susceptible to segmentation errors, especially in noisy or low-contrast 3D images. SAM-IQ-clAI yields the closest match to expected platelet morphology, suggesting that fine-tuning improves how thin particles



(a) Flatness. (b) Elongation.

Fig. 12. The flatness (b) and elongation (c) of all particles contained in the ground truth set (clAI-GT), for all methods.



(a) (b) (c)

Fig. 13. Cross correlation of volume (a), flatness (b), and elongation (c), for SAM-IQ when analysing the volume from two different directions.

are traced, but all automated methods reproduce shape descriptors well once compared on identical ground-truth objects.

Together, these findings point towards the fact that when the objects of interest reside at the edge of resolution, the segmentation algorithm becomes a modelling choice: one that can shift particle statistics, bias morphology, and reshape mechanistic interpretation.

After completion of this work, the successor to SAM 2, SAM 3 was published. SAM 3 includes a number of crucial additions that could improve 3D instance segmentation of particulate systems, i.e. multiplex segmentation, simultaneous segmentation propagation of multiple objects. Moreover, the addition of language-grounded segmentation could enable hybrid prompting that identifies concepts, rather then geometric spatial locations [49].

5. Conclusions

This study evaluated the performance of SAM-IQ (a new spatial 3D adaptation of SAM-2), its fine-tuned extension SAM-IQ-clAI, and a traditional 3D filtering-based pipeline (FDS) for segmenting kaolinite particles from synchrotron X-ray nano-holotomography performed at 50 nm voxel size.

The results demonstrate that SAM-IQ (untuned model) provides a strong quantitative baseline under the limited-data conditions investigated in this work, while requiring substantially less manual parameter tuning than the conventional FDS pipeline. Within the investigated dataset, fine-tuning increases particle detection performance; however, the resulting gains in derived morphological descriptors were comparatively modest relative to the additional annotation effort required. For data-scarce domains such as nano-CT (where acquiring annotated volumes is expensive), this is particularly significant. It suggests that large-scale pretraining may provide transferable structure recognition even in scientific volumes on which the model was never trained. Traditional FDS remains useful for generating additional labelled 3D data, but even with manual corrections it struggles in dense regions. SAM-IQ therefore represents a more suitable solution for quantitative analysis in low-contrast, sub-micron imaging.

While the present validation dataset remains limited, the results demonstrate that the choice of segmentation method has a direct influence not only on particle counts but also on the derived morphological descriptors. These differences arise from the distinct ways in which the methods handle ambiguous particle boundaries, overlapping platelets, low-contrast interfaces, and 3D prompt propagation. This is particularly relevant for applications where these descriptors are used for quantitative and modelling applications. Thus, segmentation must be treated not as a neutral preprocessing step but rather as a critical part of the measurement pipeline, influencing both the accuracy and reliability of scientific conclusions drawn from image-based materials characterisation.

Looking forward, we advocate for the development of domain-specific 3D training datasets, the use of hybrid workflows, and increased emphasis on quantitative validation standards for segmentation in particulate systems. Such developments will accelerate automated analysis of complex particulate systems imaged at the edge of spatial resolution in geomechanics and materials science.

CRediT authorship contribution statement

Theo Örtendahl: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Angela Casarella:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Formal analysis. **Olga Stamati:** Writing – review & editing, Methodology, Investigation, Funding acquisition. **Jelke Dijkstra:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Jelke Dijkstra reports financial support was provided by Swedish Transport Administration. Jelke Dijkstra reports financial support was provided by Swedish Research Council. Jelke Dijkstra reports equipment, drugs, or supplies and travel were provided by ESRF. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The financial support from the Swedish Research Council (Vetenskapsrådet) under Grant No. 2020-03982 and Trafikverket in the Branschsamverkan i Grunden framework is greatly acknowledged (TRV 2023/29710). The computations and data handling were enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS), partially funded by the Swedish Research Council through grant agreement no. 2022-06725, and by Chalmers e-Commons at Chalmers. We acknowledge the European Synchrotron Radiation Facility (ESRF) for provision of synchrotron radiation facilities under proposal number ME1668 and for the support in using beamline ID16B for the experimental campaign.

Appendix A. Filtering distance map based segmentation

The following are the applied steps for segmenting the 3D greyscale reconstructed image using the FDS method. While all steps can be implemented in various softwares, they were implemented here in ImageJ [31]. The prerequisites for ImageJ in addition to the standard functions include: clij2, Morpholibj and 3D suite. Depending on the image size, a VRAM of at least 6 GB is recommended.

Firstly, the greyscale image was loaded and a *Difference of Gaussian filter 3D* from clij2 was applied. The difference of gaussian filter applies two different gaussian blur filters to the image and subtracts one image from the other. This filter accentuates important features in the image and increases its contrast. Two different thresholding algorithms were applied separately, resulting in two binary images. One using an Otsu's threshold and one using a local Bernsen's threshold (clijx). The Otsu threshold is usually able to separate large background areas from particles, while the local Bernsen filter is able to separate more close-lying particles (but it often fails to remove large background areas). By multiplying these binarised images, a high quality binarisation of the greyscale image was acquired.

Using a distance map (clij2) on the binarised image and subsequently applying another Otsu filter, some close-lying particles were separated. The resulting binarised image contained mostly separated particle areas, which were labelled using a connected components algorithm (clij2). By alternating erode labels with *relabel islands* checked (clij2) and dilate labels (clij2) undersegmentation was further improved. To remove very small objects (which are assumed to be erroneously labelled particles) label size filtering (MorpholibJ) was applied, where the minimum size was set to 200 voxels. Furthermore, any object with a smallest axis (as calculated by MorpholibJ's max inscribed sphere) of less than 10 px was removed. To exclude particles that were not fully labelled because they lied partially outside the image crop, *Exclude labels touching image edges* (clij2) was applied. Lastly, the labels were remapped (clij2) to aid in the analysis of the labelled image (e.g. object counting).

Fig. A.14 displays sample slices for FDS along with SAM-IQ and three experimental segmentation techniques.

Appendix B. SAM 2 image post-processing

The following steps were applied to post-process the initial segmentation masks as produced by SAM 2. Firstly, to remove erroneously labelled particles (background object masks) a custom Python script was used. See the following (under construction) github repository [50]. To remove z-axis artefacts, morphological opening and closing operations were applied, consisting of dilation followed by erosion and vice versa, respectively. Subsequently, an oversegmentation correction script [38] was applied. Lastly, for quantitative analysis, labels touching any of the image edges were excluded from further analysis.

Appendix C. SAM 2 fine-tuning

See the following github repo for the fine-tuning code [50].

Data availability

Data will be made available on request.

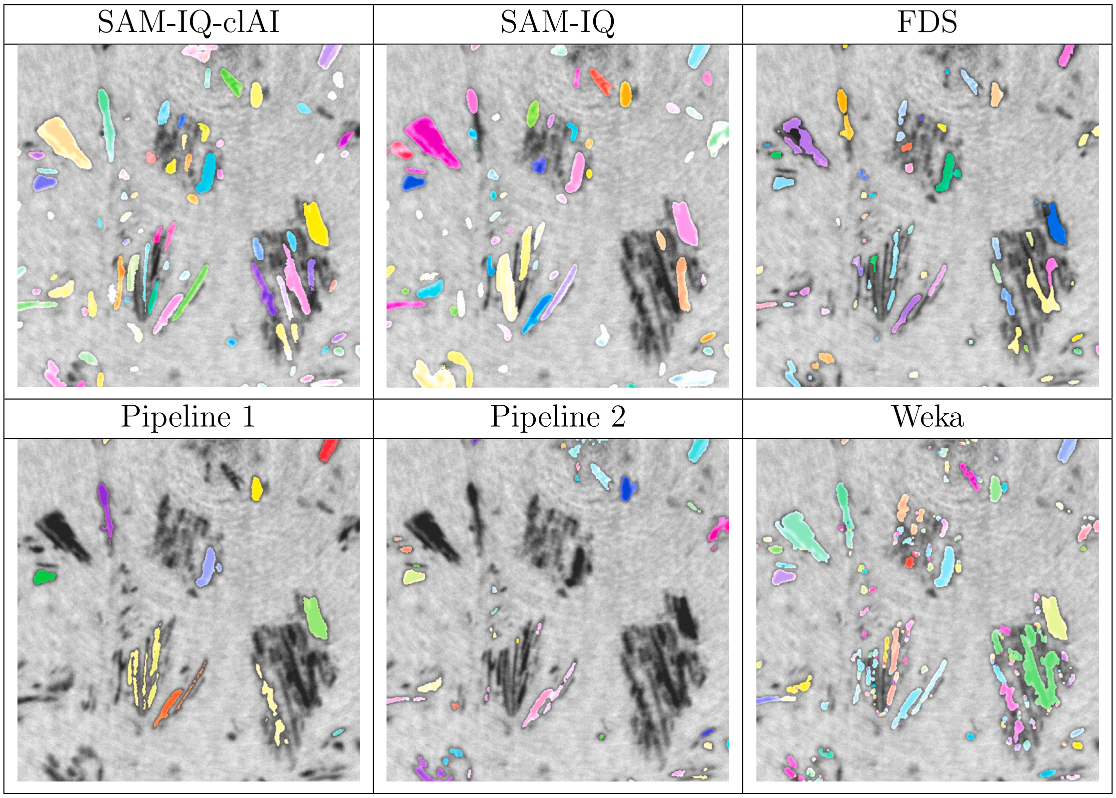


Fig. A.14. Segmentation result overlays for six different segmentation pipelines. Pipeline 1, Pipeline 2, and weka are experimental techniques that do not provide good segmentation results.

References

[1] E. Maire, P.J. Withers, Quantitative X-ray tomography, Int. Mater. Rev. 59 (1) (2014) 1–43, http://dx.doi.org/10.1179/1743280413Y.0000000023.

[2] E. Andò, S. Hall, G. Viggiani, J. Desrues, P. Bésuelle, Grain-scale experimental investigation of localised deformation in sand: A discrete particle tracking approach, Acta Geotech. 7 (2012) http://dx.doi.org/10.1007/s11440-011-0151-6.

[3] J. Fonseca, C. O'sullivan, M. coop, P. Lee, Quantifying the evolution of soil fabric during shearing using directional parameters, Géotechnique 63 (6) (2013) 487–499, http://dx.doi.org/10.1680/geot.12.P.003.

[4] P. Wang, Z. Karatza, C. Arson, DEM modelling of sequential fragmentation of zeolite granules under oedometric compression based on XCT observations, Powder Technol. 347 (2019) 66–75, http://dx.doi.org/10.1016/J.POWTEC.2019.02.050.

[5] M. Fang, A. Sibellas, J. Drummond, Y. Cao, A. Phillion, M. Martinez, V.K. Pediredla, B. Gopaluni, Deep learning-based hybrid reconstruction algorithm for fibre instance segmentation from <scp>3D</scp> x-ray tomographic images, Can. J. Chem. Eng. 101 (12) (2023) 6817–6826, http://dx.doi.org/10.1002/cjce.24939.

[6] D. Franco-Barranco, Z. Lin, W.-D. Jang, X. Wang, Q. Shen, W. Yin, Y. Fan, M. Li, C. Chen, Z. Xiong, R. Xin, H. Liu, H. Chen, Z. Li, J. Zhao, X. Chen, C. Pape, R. Conrad, L. Nightingale, J. de Folter, M.L. Jones, Y. Liu, D. Ziaei, S. Huschauer, I. Arganda-Carreras, H. Pfister, D. Wei, Current Progress and Challenges in Large-Scale 3D Mitochondria Instance Segmentation, IEEE Trans. Med. Imaging 42 (12) (2023) 3956–3971, http://dx.doi.org/10.1109/TMI.2023.3320497.

[7] T. Barisin, C. Jung, A. Nowacka, K. Schladitz, C. Redenbach, Crack segmentation in 3d concrete images: perspectives and challenges, E-Journal Nondestruct. Test. 27 (9) (2022) http://dx.doi.org/10.58286/27207.

[8] M. He, K. He, Q. Huang, H. Xiao, H. Zhang, G. Li, A. Chen, Lightweight mask R-CNN for instance segmentation and particle physical property analysis in multiphase flow, Powder Technol. 449 (2025) http://dx.doi.org/10.1016/J.POWTEC.2024.120366.

[9] P.C. Baveye, M. Laba, W. Otten, L. Bouckaert, P. Dello Sterpaio, R.R. Goswami, D. Grinev, A. Houston, Y. Hu, J. Liu, S. Mooney, R. Pajor, S. Sleutel, A. Tarquis, W. Wang, Q. Wei, M. Sezgin, Observer-dependent variability of the thresholding step in the quantitative analysis of soil images and X-ray microtomography data, Geoderma 157 (1–2) (2010) 51–63, http://dx.doi.org/10.1016/j.geoderma.2010.03.015.

[10] A. Casarella, G. Pinzón, V. Sadasivan, O. Stamati, B. Lukic, J. Villanova, J. Dijkstra, Non-invasive quantification of fabric in consolidated clay systems via synchrotron X-ray nano-tomography, Appl. Clay Sci. (2025).

[11] 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. URL: https://segment-anything.com.

[12] N. Ravi, V. Gabeur, Y.-T. Hu, R. Hu, C. Ryali, T. Ma, H. Khedr, R. Rädle, C. Rolland, L. Gustafson, E. Mintun, J. Pan, K. Vasudev Alwala, N. Carion, C.-Y. Wu, R. Girshick, P. Dollár, C. Feichtenhofer, M. Fair, SAM 2: Segment Anything in Images and Videos, 2024, URL: https://arxiv.org/abs/2408.00714v1.

[13] C. Chen, J. Miao, D. Wu, Z. Yan, S. Kim, J. Hu, A. Zhong, Z. Liu, L. Sun, X. Li, T. Liu, P.-A. Heng, Q. Li, MA-SAM: Modality-agnostic SAM adaptation for 3D medical image segmentation, Med. Image Anal. 98 (2024) 103310, http://dx.doi.org/10.1016/J.MEDIA.2024.103310, URL: http://arxiv.org/abs/2309.08842.

[14] L.B. Ferreira, V.S. Martins, U.R. Aires, N. Wijewardane, X. Zhang, S. Samiappan, FieldSeg: A scalable agricultural field extraction framework based on the Segment Anything Model and 10-m Sentinel-2 imagery, Comput. Electron. Agric. 232 (2025) http://dx.doi.org/10.1016/j.compag.2025.110086.

[15] Z. Ye, L. Lovell, A. Faramarzi, J. Ninić, Sam-based instance segmentation models for the automation of structural damage detection, Adv. Eng. Informatics 62 (2024) 102826, http://dx.doi.org/10.1016/J.AEI.2024.102826, URL: https://www.sciencedirect.com/science/article/abs/pii/S1474034624004749.

[16] R. Girshick, J. Donahue, T. Darrell, J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation, 2014.

[17] K. He, G. Gkioxari, P. Dollár, R. Girshick, Mask R-CNN, IEEE Trans. Pattern Anal. Mach. Intell. 42 (2) (2017) 386–397, http://dx.doi.org/10.1109/TPAMI.2018.2844175, URL: https://arxiv.org/abs/1703.06870v3.

[18] Z. Zhou, W. Hu, G. Xu, Y. Dong, Self-evolving prompting segment anything model for crack segmentation through data-driven cyclic conversations, Adv. Eng. Inform. 68 (2025) 103626, http://dx.doi.org/10.1016/j.aei.2025.103626.

[19] G. Taherian, E. Rezazadeh Azar, I-FCSAM: An integrated framework of few-shot learning and segment anything model for vision-based indoor built environment management, Adv. Eng. Inform. 70 (2026) 104126, http://dx.doi.org/10.1016/j.aei.2025.104126.

[20] M. Ali, T. Wu, H. Hu, Q. Luo, D. Xu, W. Zheng, N. Jin, C. Yang, J. Yao, A review of the Segment Anything Model (SAM) for medical image analysis: Accomplishments and perspectives, Comput. Med. Imaging Graph. 119 (2025) 102473, http://dx.doi.org/10.1016/J.COMPMEDIMAG.2024.102473.

[21] Y. Yang, L. Xu, L. Tian, SAM2-3dMed: Empowering SAM2 for 3D Medical Image Segmentation, 2025.

[22] L.R. Dice, Measures of the Amount of Ecologic Association Between Species, Ecology 26 (3) (1945) 297–302, http://dx.doi.org/10.2307/1932409, /doi/pdf/10.2307/1932409 https://onlinelibrary.wiley.com/doi/abs/10.2307/1932409 https://esajournals.onlinelibrary.wiley.com/doi/10.2307/1932409.

[23] G. Martínez-Criado, J. Villanova, R. Tucoulou, D. Salomon, J.-P. Suuronen, S. Labouré, C. Guilloud, V. Valls, R. Barrett, E. Gagliardini, Y. Dabin, R. Baker, S. Bohic, C. Cohen, J. Morse, ID16B: a hard X-ray nanoprobe beamline at the ESRF for nano-analysis, J. Synchrotron Radiat. 23 (1) (2016) 344–352, http://dx.doi.org/10.1107/S1600577515019839.

[24] P. Cloetens, W. Ludwig, J. Baruchel, D. Van Landuyt, J.P. Guigay, M. Schlenker, Holotomography: Quantitative phase tomography with micrometer resolution using hard synchrotron radiation X–rays, Appl. Phys. Lett. 75 (1999) 2912–2914, http://dx.doi.org/10.1063/1.125225.

[25] A. Mirone, E. Brun, E. Gouillart, P. Tafforeau, J. Kieffer, The PyHST2 hybrid distributed code for high speed tomographic reconstruction with iterative reconstruction and a priori knowledge capabilities, Nucl Instrum Methods Phys Res Sect B 324 (2014) 41–48.

[26] Y. Wang, C. Lin, J. Miller, 3D image segmentation for analysis of multisize particles in a packed particle bed, Powder Technol. 301 (2016) 160–168, http://dx.doi.org/10.1016/j.powtec.2016.05.012.

[27] G. Kerckhofs, J. Schrooten, T. Van Cleynenbreugel, S.V. Lomov, M. Wevers, Validation of x-ray microfocus computed tomography as an imaging tool for porous structures, Rev. Sci. Instrum. 79 (1) (2008) http://dx.doi.org/10.1063/1.2838584.

[28] C. Shannon, Communication in the Presence of Noise, Proc. the IRE 37 (1) (1949) 10–21, http://dx.doi.org/10.1109/JRPROC.1949.232969.

[29] M. Wiebicke, E. Andò, I. Herle, G. Viggiani, On the metrology of interparticle contacts in sand from x-ray tomography images, Meas. Sci. Technol. 28 (12) (2017) 124007, http://dx.doi.org/10.1088/1361-6501/aa8dbf.

[30] C. Garcia-Lopez-de Haro, C. Fuster-Barcelo, C.T. Rueden, J. Heras, V. Ulman, D. Franco-Barranco, A. Ines, K.W. Eliceiri, J.-C. Olivo-Marin, J.-Y. Tinevez, D. Sage, A. Munoz-Barrutia, SAMJ: Fast Image Annotation on ImageJ/Fiji via Segment Anything Model, 2025, URL: https://www.arxiv.org/pdf/2506.02783.

[31] C.T. Rueden, J. Schindelin, M.C. Hiner, B.E. DeZonia, A.E. Walter, E.T. Arena, K.W. Eliceiri, ImageJ2: ImageJ for the next generation of scientific image data, BMC Bioinformatics 18 (1) (2017) 1–26, http://dx.doi.org/10.1186/s12859-017-1934-z, URL: https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-017-1934-z.

[32] P. Bankhead, M.B. Loughrey, J.A. Fernández, Y. Dombrowski, D.G. McArt, P.D. Dunne, S. McQuaid, R.T. Gray, L.J. Murray, H.G. Coleman, J.A. James, M. Salto-Tellez, P.W. Hamilton, QuPath: Open source software for digital pathology image analysis, Sci. Rep. 2017 7:1 7 (1) (2017) 1–7, http://dx.doi.org/10.1038/s41598-017-17204-5, URL: https://www.nature.com/articles/s41598-017-17204-5.

[33] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinform.) 9351 (2015) 234–241, http://dx.doi.org/10.1007/978-3-319-24574-4{_}28/SAVE-RESEARCH, URL: https://link.springer.com/chapter/10.1007/978-3-319-24574-4_28.

[34] L.C. Chen, Y. Zhu, G. Papandreou, F. Schroff, H. Adam, Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 11211 LNCS (2018) 833–851, http://dx.doi.org/10.1007/978-3-030-01234-2{_}49, URL: https://arxiv.org/abs/1802.02611v3.

[35] R. Szeliski, Computer Vision, second ed., Springer Nature, Cham, Switserland, 2021.

[36] N. Otsu, Threshold selection from gray-level histograms, IEEE Trans. Syst. Man Cybern. SMC-9 (1) (1979) 62–66, http://dx.doi.org/10.1109/TSMC.1979.4310076.

[37] J. Bernsen, Dynamic thresholding of grey-level images , in: Eighth International Conference on Pattern Recognition, Paris, 1986, pp. 1251–1255.

[38] O. Stamati, E. Andò, E. Roubin, R. Cailletaud, M. Wiebicke, G. Pinzón, C. Couture, R. Hurley, R. Caulk, D. Caillerie, T. Matsushima, P. Bésuelle, F. Bertoni, T. Arnaud, A. Ortega Laborin, R. Rorato, S. Yue, A. Tengattini, O. Okubadejo, G. Birmpilis, spam: Software for Practical Analysis of Materials, J. Open Source Softw. 5 (2020) 2286, http://dx.doi.org/10.21105/joss.02286.

[39] I. Loshchilov, F. Hutter, Decoupled Weight Decay Regularization, in: 7th International Conference on Learning Representations, ICLR 2019, International Conference on Learning Representations, ICLR, 2017, URL: https://arxiv.org/abs/1711.05101v3.

[40] S. Narang, G. Diamos, E. Elsen, P. Micikevicius, J. Alben, D. Garcia, B. Ginsburg, M. Houston, O. Kuchaiev, G. Venkatesh, H. Wu, Mixed Precision Training, in: 6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings, International Conference on Learning Representations, ICLR, 2017, URL: https://arxiv.org/abs/1710.03740v3.

[41] X. Jia, R.A. Williams, 3D-measurement of particles and particulate assemblies - A review of the paradigm shift in describing anisotropic particles, Powder Technol. 447 (2024) 120109, http://dx.doi.org/10.1016/J.POWTEC.2024.120109.

[42] J. Ollion, J. Cochennec, F. Loll, C. Escudé, T. Boudier, TANGO: a generic tool for high-throughput 3D image analysis for studying nuclear organization, Bioinformatics 29 (14) (2013) 1840–1841, http://dx.doi.org/10.1093/BIOINFORMATICS/BTT276.

[43] D. Legland, I. Arganda-Carreras, P. Andrey, MorphoLibJ: integrated library and plugins for mathematical morphology with ImageJ, Bioinformatics 32 (22) (2016) 3532–3534, http://dx.doi.org/10.1093/bioinformatics/btw413.

[44] V. Angelidakis, S. Nadimi, S. Utili, Elongation, flatness and compactness indices to characterise particle form, Powder Technol. 396 (2022) 689–695, http://dx.doi.org/10.1016/j.powtec.2021.11.027.

[45] E. Pirard, G. Dislaire, Sensitivity of particle size and shape parameters with respect to digitizationURL: www.occhio-usa.com.

[46] A. Casarella, Multi-scale investigation of the thermo-mechanical behaviour of non-active clay (Ph.D. thesis), Université Grenoble Alpes, 2022.

[47] E. Vitale, D. Deneele, G. Russo, G. Ouvrard, Short-term effects on physical properties of lime treated kaolin, Appl. Clay Sci. 132–133 (2016) 223–231, http://dx.doi.org/10.1016/j.clay.2016.04.025.

[48] J.K. Mitchell, K. Soga, Fundamentals of Soil Behavior, third ed., John Wiley & Sons, New Jersey, 2005.

[49] R. Sapkota, K.I. Roumeliotis, M. Karkee, The SAM2-to-SAM3 Gap in the Segment Anything Model Family: Why Prompt-Based Expertise Fails in Concept-Driven Image Segmentation, 2025, URL: https://arxiv.org/abs/2512.06032v1.

[50] T. Örtendahl, Deep learning-based methods for segmentation and labelling of clay, 2025, URL: https://github.com/theoortendahl/SAM-clAI.