Journal of Nondestructive Evaluation (2025) 44:119
https://doi.org/10.1007/s10921-025-01262-1

Superimposing Synthetic Defects into Real XCT Data and Segmentation-Based Comparison for Advanced Probability of Detection Evaluation

Miroslav Yosifov[1,2] · Bernhard Fröhler[2] · Jan Sijbers[1] · Jan De Beenhouwer[1] · Johann Kastner[2] · Christoph Heinzl[3,4]

Received: 15 May 2025 / Accepted: 13 August 2025 / Published online: 1 September 2025

Abstract

This research proposes an approach for integrating realistic defects into computed tomography (XCT) scans by using X-ray simulations. It allows full control over different scenarios and measuring the detection algorithm efficiency in real-world situations. Using real XCT data of a pin-fin cooler made of aluminum alloy with complex internal structures, synthetic spherical and irregular defects ranging from 56 µm to 300 µm in diameter are superimposed to create a comprehensive dataset that mimics a wide range of realistic scenarios. This XCT dataset with superimposed defects is then utilized to apply a probability of detection analysis to detect defects of varying sizes and shapes. This analysis shows that for spherical pores, the detectability limit is up to 2.5 times higher in the superimposed case with a minimum voxel similarity of 95%, while for irregular pores, this limit is 3.3 times higher when a minimum voxel similarity of 80%. The integration of synthetic defects into real XCT images allows for a more rigorous and controlled assessment of detection algorithms, providing valuable insights into their performance under realistic conditions. Our findings demonstrate that this method can significantly improve the accuracy and reliability of measurements of defect detectability, offering a powerful tool for quality assurance in critical manufacturing processes.

Keywords Computed tomography · Superimposing defects · Shape variations · Probability of Detection · XCT Simulation

1 Introduction

XCT is an advanced imaging technique commonly employed in nondestructive testing due to its ability to provide detailed and accurate internal views of industrial specimens without causing mechanical damage [1]. It is especially valuable for analyzing and visualizing internal defects such as pores, fibers, and cracks in industrial components [2]. Segmentation and detection algorithms are required for deep and accurate analysis of XCT data for extraction of defects. Selecting an appropriate segmentation method is crucial for achieving reliable and accurate results. Recently, the probability of detection (POD) [3–5] analysis has become a crucial tool in assessing the performance of detection methods for XCT data. POD provides a quantitative measure of the likelihood that a defect detection algorithm will correctly identify and characterize flaws of a given size, thereby supporting the evaluation of imaging systems, processes, and detection techniques [6, 7]. However, to create POD curves, multiple XCT simulations and the repeated application of segmentation on the reconstructed datasets are necessary. This process is computationally intensive and time-consuming, particularly for large XCT datasets. Moreover, even though realistic simulations can be created, it is not possible to fully integrate all types of artifacts that occur during the scanning process. Improving POD accuracy in XCT requires a

✉ Miroslav Yosifov
miroslav.yosifov@fh-wels.at

1 imec-Vision Lab, Dept. of Physics, University of Antwerp, Universiteitsplein 1, Antwerpen 2610, Belgium

2 Research Group Computed Tomography, University of Applied Sciences Upper Austria, Stelzhamerstrasse 23, Wels 4600, Austria

3 Division Development Center X-ray Technology, Fraunhofer Institute for Integrated Circuits IIS, Flugplatzstraße 75, 90768 Fürth, Germany

4 Faculty of Computer Science and Mathematics, University of Passau, Innstraße 43, 94032 Passau, Germany

multidisciplinary approach considering technical as well as human aspects. Several elements influence POD accuracy, including measurement precision, operator expertise, XCT system characteristics, scanning parameters, reconstruction algorithms, and segmentation methods [5]. Precision in the measurement process is one of the most critical aspects of detecting defects. Kim et al. [8] showed that even slight variations in scanning parameters, calibration, or data processing can significantly affect defect detection and POD accuracy. Changes in XCT acquisition parameters—such as source power, acceleration voltage, exposure time, and the number of radiographs—impact on image noise, contrast, and general image quality [8]. These changes influence the detectability of defects and can shift the POD curve, impacting both sensitivity and confidence in detection results. The level of experience of the inspector also plays a very influential role in the interpretation of XCT results as well as analysis of the XCT data. The effectiveness of defect detection and the optimization of scanning parameters are both strongly influenced by the inspector's background and expertise [9]. XCT hardware also plays a key role, as different systems vary in resolution, signal-to-noise ratio, and scanning capabilities. Furthermore, the choice of reconstruction algorithm for 3D image generation directly affects image quality and thus the accuracy of defect detection [10]. The final step in defect detection involves segmentation and extraction of the identified defects. The accuracy of segmentation is crucial for ensuring that the defect is correctly identified and its size, shape, and location are precisely measured, all of which contribute to a more reliable POD analysis [7]. In current applications of POD generation, XCT simulations are essential. However, these processes currently lack integration of the full range of measurement accuracy of XCT scanning and the influence of the inspector during the scanning process. By collectively addressing these aspects—scanning parameters, inspector expertise, hardware capabilities, and effective reconstruction and segmentation—the reliability and efficiency of POD assessments in XCT can be significantly improved. Real-world XCT data is limited by the availability and variability of actual defects, which renders it challenging to comprehensively test and refine detection algorithms. Conversely, synthetic models or XCT simulations, despite their versatility, may lack the realism needed to accurately represent true material and scanning conditions. To address these challenges, this study proposes a hybrid methodology that merges the realism of actual XCT scans with the controlled variability of synthetic defects. By superimposing artificial defects on real XCT images, we create a comprehensive dataset that encompasses a broad spectrum of potential defect scenarios. This process also enables us to refine our POD pipeline by reducing the number of necessary simulations to just two: one with defects and one without.

2 Related Work

This section reviews related research on synthetic defect superimposition in XCT (Subsection 2.1), advances in XCT simulation (Section 2.2) (Subsection 2.2), and developments in POD analysis (Subsection 2.3).

2.1 Superimposing Synthetic Defect

Few studies have explored the superimposition of synthetic defects into real XCT scans, primarily due to the time-consuming nature of creating accurate and realistic defect representations. This approach requires careful integration of synthetic defects, which introduces computational complexity and increases processing time. Tenscher-Philipp et al. [11] discuss the challenges and time-consuming nature of generating realistic synthetic defects in XCT data.

The process of superimposing synthetic defects is a challenging task that requires careful consideration of several steps to achieve a realistic representation of the defects. There are two different ways to superimpose defects in real XCT data images. The first method involves superimposing the defects in radiographs or projected images. The second method involves superimposing the defects in a 3D reconstructed volume created from radiographs. The first method was presented by Brierley et al. [12], who used simulation and experimental XCT data by subtracting each other in projections with the application of seven defects. In our application, we have utilized both methods and decided to accelerate the process by superimposing defects onto the reconstructed volume, thereby avoiding the need for precise registration of each projection. Another critical point and difference in our approach is the presence of several pores already existing in the XCT data. We need to ensure that these do not overlap with the defects.

2.2 XCT Simulation

XCT simulation is increasingly applied in fields such as biomedical applications [13] and material science [14], where precise and optimized scanning parameters are critical for XCT imaging tasks, such as in-line inspections of industrial specimens and training data for deep learning models. In these cases, XCT simulation allows for the generation of synthetic ground truth data for defect detection and the comparison of different segmentation and defect detection algorithms [6, 7].

XCT simulations play a key role in mimicking XCT system behavior, helping to evaluate imaging performance and optimize scanning parameters [15]. They enable engineers and developers to model and visualize complex physical effects such as attenuation, scattering, and geometric

distortions that occur during X-ray scans. By simulating these factors, users can gain insights into how these effects influence image quality. XCT simulation also plays a vital role in optimizing scan parameters, ensuring that factors like noise, image blur, and artifacts are minimized, thus improving overall image quality. Additionally, it is essential for estimating measurement uncertainty, giving users a clearer understanding of the limitations and accuracy of their imaging systems [16].

There are two main types of XCT simulation tools. The first category includes numerical simulation tools such as XSimulation [17] and the ASTRA toolbox [18], which simulate radiographic images and support system parameter optimization. These tools employ efficient ray-tracing algorithms to create radiographic images and typically include visualization modules for analyzing results. The second type is Monte Carlo XCT simulations, which focus on simulating X-ray photon interactions with high accuracy. Tools like MC-GPU [19] and PENELOPE [20] generate realistic radiography projection images and XCT scans by simulating complex photon behavior. While Monte Carlo-based simulations are highly accurate, they are computationally expensive and time-consuming, making them less suitable for our work, which is why we focus on more efficient numerical simulation methods. In this study, the numerical simulation tool SimCT [21] was used to generate respective simulated data.

2.3 Probability of Detection

POD is fundamentally a statistical measure used to define and quantify the detection limits for task-specific inspection data. POD has been applied in various fields, including XCT, eddy current testing [22], and ultrasonic inspection [23]. As additional example, in structural health monitoring, a model-assisted POD framework was developed for optical fiber sensors, allowing researchers to better quantify and optimize the damage detection capabilities of these sensors under various environmental and operational conditions [24]. In recent years, the role of POD analysis has gained momentum in the field of Non-Destructive Testing (NDT) and XCT, especially for safety based critical applications such as aerospace and additive manufacturing. Advancements in simulation, deep learning, and model-assisted approaches have led to more precise and explainable defect detection pipelines. Yosifov et al. [6] demonstrated how numerical simulations could be employed to model POD curves, providing a practical framework for evaluating detection performance in XCT systems. This has been expanded upon by Lee et al. [25], who proposed a model-assisted POD methodology to further quantify inspection reliability. More recently, Zhou et al. [26] introduced a sinogram-based approach using deep learning to localize defects at the raw data level, potentially improving POD by preserving information typically lost in reconstruction. In parallel, Liu et al. reviewed the integration of object detection algorithms in NDT, emphasizing how such methods contribute to more reliable detection outcomes [27].

However, most of these studies rely on simulated data. In this study, we take a step further by increasing the accuracy of POD through an approach that superimposes defects directly onto experimental XCT scans.

3 Materials and Methods

In this section, we introduce our methodology, which encompasses XCT imaging, simulated data generation, defect superimposition and modeling, the processing framework, as well as the generation, segmentation, and assessment of defects using POD analysis.

3.1 XCT Imaging and Simulated Data Generation

In this study, the specimen used for scanning and analysis is a pin-fin cooler made from an aluminum alloy and manufactured at Materialise NV using metal laser powder bed fusion (PBF-LB/M), with dimensions of 103 mm in length, 18 mm in height, and 5 mm in width. Scanning parameters were optimized for a fast XCT scanning with a total acquisition time of 5 min. Key settings included an acceleration voltage of 200 kV and an anode current of 102 µA, with an integration time of 500 ms per projection. A total of 500 projections were acquired over a complete scan cycle lasting approximately 5 min. The system was configured with a source-object distance of 304.14 mm and a source-detector distance of 1091.9 mm, resulting in a magnification of 3.59 and an effective voxel size of 55.7 µm, allowing the full specimen to be captured on the detector. No binning was applied, and the full detector resolution of 2000×2000 pixels was used. The focal spot size was set to 1 mm. These parameters are summarized in Table 1.

For the simulated data, the same scanning parameters were applied, and the STL (Standard Tessellation Language) model of the specimen was positioned nearly identically to its placement in the experimental XCT scan. Three different simulations were generated with physical effects disabled: one with defects inserted, one without defects to create superimposed data, and one with defects and physical effects enabled, in order to compare the POD results between the superimposed data and the simulations with physical effects enabled.

3.2 Superimposing Defects in XCT Data

This section describes the process of superimposing simulated defects into real XCT scans. To ensure accurate POD

Table 1 Parameters for XCT Imaging

Parameters	Unit	Value	Parameters	Unit	Value
Acceleration voltage	kV	200	Source-object-distance	mm	304.14
Anode current	μA	102	Source-detector-distance	mm	1091.9
Integration time	ms	500 ms	Magnification	–	3.59
Detector pixels	–	2000 × 2000	Voxel size	μm	55.7
Binning	–	1	Focal spot size	mm	1
Number of projections	–	500	Total cycle time	min	5

results, several key factors must be considered. In the first, scanning parameters and specimen locations were optimized to create realistic simulation images. Two different simulations were then created: one defect-free and one containing defects. The defect-free simulation was subtracted from the defect-containing simulation, and the result was added to the real XCT scan. A key requirement of this process is achieving precise alignment between the simulated and real XCT scans to ensure accurate defect placement. To ensure precise alignment and accurate superimposition of synthetic defects, we converted the reconstructed XCT volume into an STL model, preserving all transformation information, including position and orientation. Porosity was removed during surface mesh generation to create a defect-free reference geometry. Registration accuracy was validated using the overlap analysis function in VG Studio Max, achieving a 100% surface match between the converted STL model and the reference surface geometry, confirming robust alignment for reliable defect insertion. However, some errors remain, with a maximum of $\pm$ 0.12 mm (two voxel), particularly on the surface.

Another important aspect to consider is the analysis of intensity profiles. Figure 1 compares intensity profiles along a selected line across the specimen and shows how simulated data with and without physical effects align with the superimposed CT data. We examine profiles from simulated data with and without introduced physical effects such as noise, unsharpness, scattering and focal spot blurriness effect, as well as from the superimposed XCT data. We optimize the background and porosity gray values across all simulated datasets (see Fig. 1). This similarity is essential for achieving realistic defect superimposition, even though there may be minor differences in the gray values within the material regions. It is also crucial to preserve the integrity of structural features at defect boundaries. To address this, no interpolation or resampling is performed during the insertion process. Instead, defects are embedded directly into spatially aligned voxel grids. This approach ensures that the original intensity values are preserved at the defect boundaries, effectively minimizing edge blending and avoiding interpolation artifacts.

Figure 2 shows two XCT slices extracted from a 3D volume of an experimental scan. The image on the right shows the original, unaltered XCT slice, containing only naturally occurring structural features and defects. The image on the

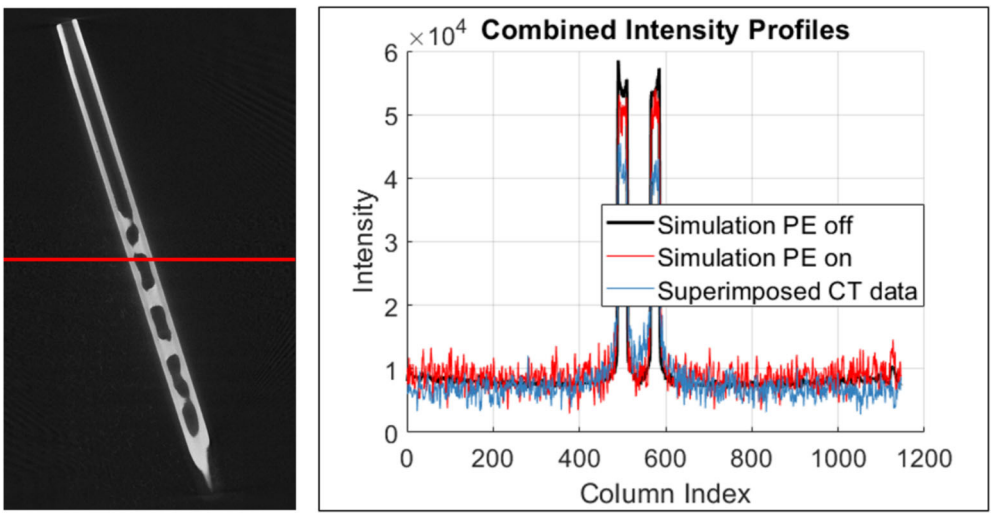


Fig. 1 Intensity profiles across a selected row of the XCT slice, comparing simulations with physical effects enabled (PE on) and disabled (PE off) against superimposed XCT data. The red line on the XCT slice (left) indicates the row from which the intensity profiles (right) were extracted for comparison

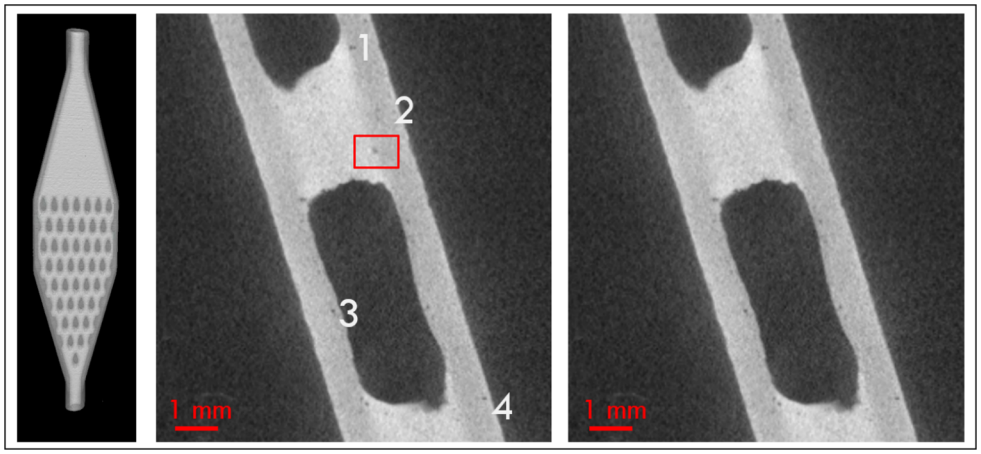


Fig. 2 3D rendering and experimental XCT slices from a 3D volume. **Left:** 3D rendering of the scanned specimen showing the overall structure. **Middle:** XCT slice with a superimposed synthetic defect (labeled as 2) embedded into the real scan. Four visible pores or defects are highlighted and labeled 1 through 4. **Right:** Corresponding original XCT slice without the added defect, used as a reference for comparison

left shows the same XCT slice after the superimposition of a synthetic defect, marked as region 2 in the red box. This synthetic defect was inserted to mimic realistic defect morphology and intensity distribution, closely matching the material characteristics of the surrounding region. Regions labeled 1, 3, and 4 correspond to real structural or defect features in the sample, serving as references for spatial context. This visual similarity ensures that the defect is superimposed realistically into the background, providing a reliable basis for evaluating detection performance more accurately. Additionally, a quantitative analysis of the superimposed defects will be presented in the Section 4.

3.3 Processing Framework

The overall data processing pipeline is illustrated in Fig. 3, showing example XCT slice images for each step. The figure outlines both the defect superimposition workflow and the POD generation process. The processing pipeline followed in our analysis involved several key stages. First, a XCT scan was acquired of a 3D-printed pin-fin cooler component. Two types of synthetic defect series were created: spherical defects of varying sizes and irregularly shaped defects with randomized shape and size. In the next phase, these randomly generated defects were distributed within the boundaries of the specimen. Then, using the XCT scan data, which included position, rotation, and scanning parameters, two sets of simulated data were created without introducing any physical effects: one without defects and another with artificially introduced defects. This stage is crucial before superimposing the defects onto the XCT data. Additionally, further simulations were generated with introducing physical effects such as noise, unsharpness, and scattering. These simulations were generated to compare the results between simulations and the superimposed XCT data. Next, segmentation and defect detection were performed on the XCT data to identify pore locations and determine the maximum and minimum sizes of the pores via VG Studio Max. This information was essential for avoiding overlap with existing defects in the specimen, and the size parameters were used to define the defects in the simulations. Subsequently, we visualized the STL models of the specimen, the defects, and the specimen with defects inserted. This 3D representation ensured accurate positioning of the defects within the specimen's boundaries. The simulated defects were then superimposed onto the experimental XCT data by subtracting the simulation with defects from the one without defects. The resulting data was then added to the experimental XCT

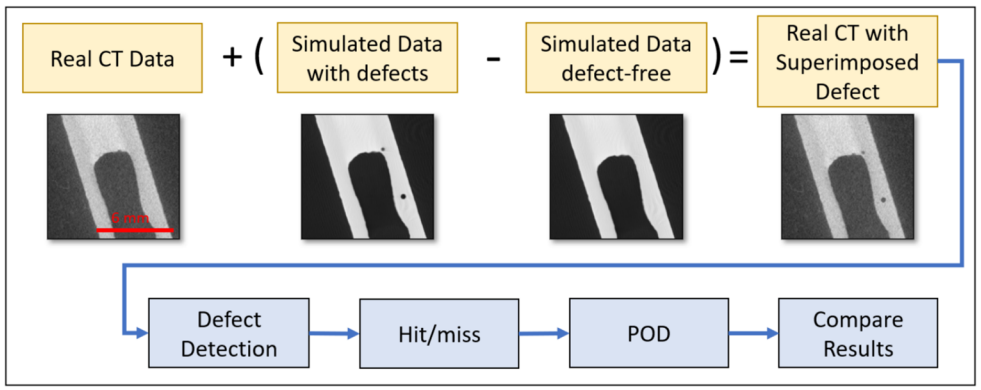


Fig. 3 The processing pipeline including exemplary images (cut from the volume) for XCT scans and simulation use case

data for further analysis. Following this step, a segmentation algorithm was applied to identify and characterize the introduced defects. This step enabled us to measure and compare the accuracy and sensitivity of our detection methods. In line with previous studies [7, 28] and comparisons of different CNN-based supervised and unsupervised segmentation algorithms [29], we applied the 3D U-Net in this study. Finally, POD curves were compared between the simulated data and the XCT data with superimposed defects. This comparison also allowed us to evaluate the effectiveness of our detection methods and identify potential areas for improvement.

3.4 Defect Modeling and Generation

Defect modeling and generation are essential steps in our pipeline to produce results that closely resemble real XCT scans. First, various defect types were extracted from real XCT scans, and their sphericity as well as minimum and maximum sizes were analyzed to define the size range for artificial defects. Using the shape variations method—which generates randomized shapes based on a surface model [29]—defects of varying shapes and sizes were synthesized to mimic real-world flaws. Then, a smoothing process was applied to the shape variations of irregular pores using the Laplacian smoothing algorithm [30] with five iterations. In the next phase, these randomly generated defects were distributed within the boundaries of the specimen. Additionally, all detected defect locations were recorded to avoid overlapping with defects already existing in the specimen. In this study, we address spherical types of defects and realistic defects of random shapes and sizes. A ray-casting method was implemented to determine if a defect lies inside a triangular mesh. This function utilizes the Möller–Trumbore algorithm [31] to check for intersections between a ray and a triangle.

3.5 POD and Segmentation

In this study, the POD method was utilized with a hit/miss approach to evaluate the defect detection capabilities. POD is a statistical tool used to quantify the efficiency of inspection processes in identifying specific defects. To construct POD curves, a binary response indicating detection was used, which was derived through the hit/miss method and modeled using the Probit linking function for logistic regression [32, 33]. The defect size at which the POD curve reaches 0.90 is denoted as a_{90}. Additionally, the characteristic value $a_{90/95}$ was calculated, representing the lower bound of the 95% confidence interval for a_{90}. This detection limit corresponds to the smallest defect size that can be reliably detected with 95% confidence.

In our current approach, we have employed the deep learning technique 3D U-Net [34] for segmentation, to compare the results between POD generated from pure XCT simulations and those derived from superimposed defects in real XCT scans. In this study, we aim to improve POD results accuracy and detection limits more quickly and with minimal error by using the superimposition method and real XCT scans.

K-means clustering [35] was applied to high-quality, noise-free simulated data to generate reliable ground truth for voxel-wise 3D comparisons with deep learning-based segmentation. To ensure the generation of high-quality ground truth data with strong contrast between the material and defects, the number of projections was increased from 500 to 1440. The improved image quality in these volumes enables highly accurate segmentation with minimal boundary uncertainty. Additionally, small defects were manually reviewed to prevent mislabeling at the boundaries. This provides a reliable basis for binary hit/miss analysis.

Jaccard Index (JI) [36] is used to quantify the similarity between binary volumes obtained from different segmentation methods. As illustrated in Fig. 4, the workflow begins with segmentation of the volume using both k-means clustering and deep learning method 3D U-Net (1). Next, a voxel-wise comparison between the binary volumes from both approaches is performed (2). The JI is then computed for each case to quantify segmentation agreement (3). A detection is classified as a hit when the JI is at least 75%, and as a miss otherwise (4). In the testing process, different JI thresholds are selected, with a minimum of 75%. This classification yields a binary detection signal, encoded as a hit/miss vector (5). Based on this signal, the POD is estimated, and the minimum detectable limit is determined for each evaluated 3D U-Net segmentation method (6). In the final step, a comparison is made between the simulations and the superimposed data to quantify the differences.

Figure 5 illustrates a voxel-wise comparison between predicted and ground truth defect volumes, shown using a color-coded 3D visualization. Light beige indicates true positives (correct segmentation), red shows false positives (over-segmentation), and dark gray highlights false negatives (missed regions).

3.5.1 Training and Testing

For the simulated dataset, 400 training subvolumes were generated by simulating two pore types: 200 spherical and 200 irregular-shaped pores. These defects were embedded in two separate XCT simulation datasets. The defect-containing regions were then extracted as subvolumes of size $64 \times 64 \times 64$ voxels for use in the training process. Similarly, 400 training subvolumes were created for the superimposed defect dataset using the same procedure. For both pore types, diameters were selected in the range of 56–300 μm, based on real experimental observations. Similarly, an initial diameter

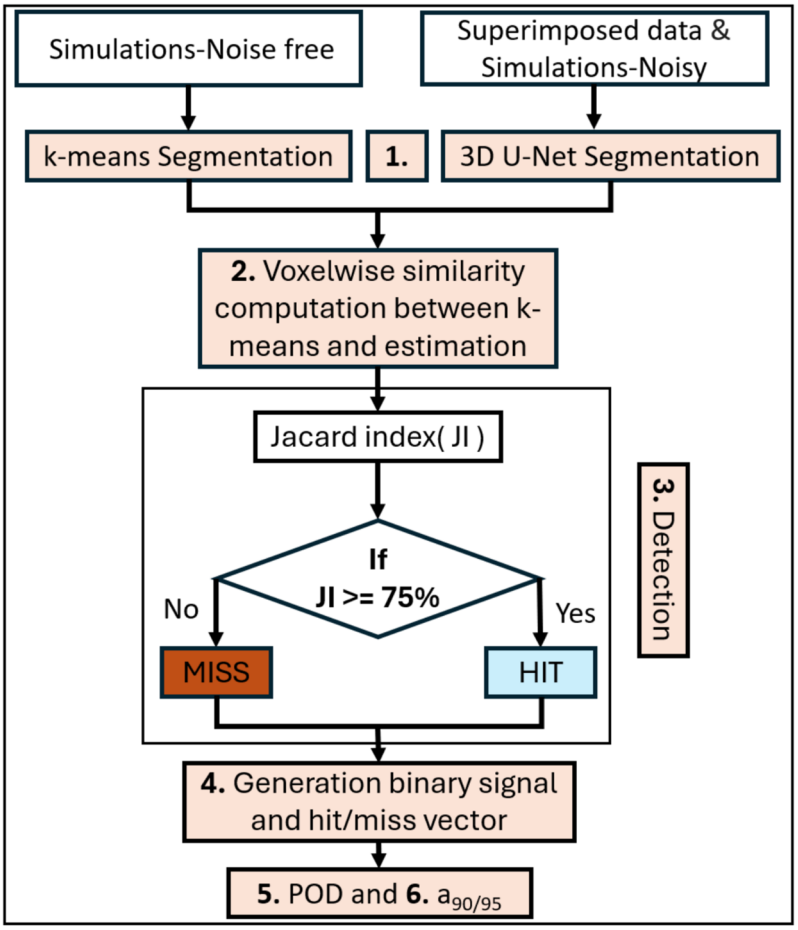


Fig. 4 Workflow for generating ground truth and evaluating the detectability limit through hit/miss analysis and POD

range of 56 to 300 μm was also defined for the irregular-shaped pores. To create the irregular geometry, each point in the corresponding STL model was randomly displaced, and a smoothing function was subsequently applied. Finally, two models were trained and generated for the simulated and superimposed datasets. For ground truth generation, all physical effects such as noise, blur, and scattering were disabled to obtain high-quality simulated images. K-means segmentation was then applied to these clean images to extract accurate ground truth labels. For the testing part, a total of 200 pores were generated for each pore type, resulting in 200 spherical and 200 irregular-shaped pores.

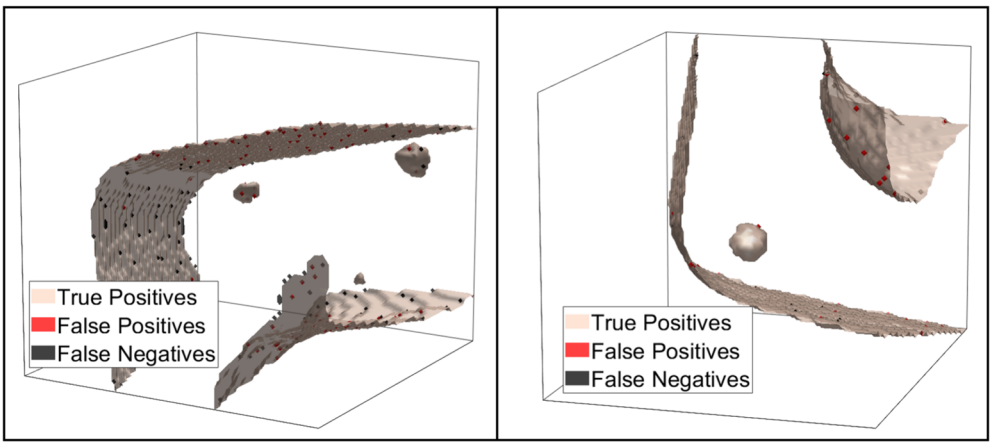


Fig. 5 Three-dimensional visualization of defect segmentation performance. True positives (light beige) indicate correctly segmented defect regions, while false positives (red) and false negatives (dark gray) represent over-segmented and missed areas, respectively

4 Results

This section presents both visual and quantitative analyses of simulated and superimposed XCT data, including gray value profile evaluation, segmentation performance, image similarity analysis to validate the realism of the superimposed defects, and POD analysis.

4.1 Visualization of Defects and Segmentation Performance

This section presents example visualizations of pores with spherical and irregular shapes, along with their corresponding segmentation results generated by the 3D U-Net algorithm.

4.1.1 Analysis of Pores with Spherical Shape

Figure 6 show a comparative analysis of the segmentation results on superimposed and simulated image data for spherical pores. A red bounding box highlights the pore region of interest in each image. To assess segmentation accuracy, the bottom row shows a visual comparison between the ground truth mask and the predicted segmentation (outlined in red). Regions of disagreement between the two are indicated by pink and green overlays, offering a direct visual cue to the types of errors (false positives and false negatives). Crucially, each segmentation result in the bottom row is accompanied by the JI value, a statistical measure that quantifies the similarity between the estimated segmentation and ground truth regions. Higher JI scores indicate greater overlap and thus improved segmentation accuracy. By presenting results across a range of parameter values, these figures aim to assess the robustness and sensitivity of the segmentation algorithm when comparing outputs generated from superimposed and simulated XCT data. Overall, the figures show that, segmentation similarity—reflected by higher JI values—is greater in the simulated images compared to the superimposed ones. However, in some cases (e.g., image 35, 40), the JI values were higher for the superimposed defects, suggesting that the location of pores and local gray value differences may also influence the segmentation accuracy and detection performance in the superimposed case.

4.1.2 Analysis of Pores with Irregular Shape

Similarly, Fig. 7 shows a comparative analysis of segmentation results for superimposed and simulated data, focusing on irregularly shaped pores. As observed in the figures, the JI values are significantly lower for irregularly shaped pores compared to those with spherical geometry, indicating a notable decrease in segmentation performance. From Fig. 7, it can also be observed that the simulated specimen data exhibits slight differences in certain border regions, likely due to discrepancies between the manufactured specimen and and the exported STL model used for simulation.

4.2 Quantitative Analysis and Comparison of Defects

Figure 8 illustrates a side-by-side comparison between the superimposed and simulated datasets, focusing on selected subvolumes and their respective gray value distributions. For each selected volume index, a central XY-slice is extracted from both datasets. These slices are shown in the first and second rows of the figure, respectively. To analyze local intensity differences, a horizontal line is overlaid on each slice at the center row, from which gray value profiles are extracted. These profiles are plotted in the third row, with blue representing the superimposed data and orange the simulated data. This visualization enables direct comparison of gray value trends at the same spatial locations, revealing variations in contrast, noise, or defect visibility that may impact segmentation.

Gray value profile analysis in the pore region for cases 40 and 50 indicates both visual and quantitative similarity. However, for case 50, slight differences in gray values are observed in the specimen area between the superimposed and simulated XCT data. This is likely due to existing porosity in that region, which affects the overall gray value levels (see Fig. 8). For case 45, the gray values in the specimen region are similar between the datasets; however, a slight difference is observed in the pore region. Additionally, the figure offers both visual and quantitative insight into how accurately synthetic simulations replicate real-world intensity profiles. These observed differences in gray value distributions, whether in the pore or sample regions, can influence the reliability of segmentation results to distinguish between true defect boundaries and background. As a result, even subtle gray value inconsistencies may lead to reduced segmentation accuracy, particularly in regions where the contrast between pore and material is limited.

4.3 Validation of Superimposed Defect Realism

Figure 9 shows a comparison of gray value profiles between superimposed defects and real CT scans for cases 15, 17, and 19, following a similar layout as Fig. 8. The top row shows XY-slices from superimposed data, and the middle row presents corresponding real CT slices. The colored lines mark the profile extraction paths, and the bottom row plots the gray value profiles along these lines. The superimposed data (blue) introduces noticeable gray value decreases at defect locations, closely resembling those in the real CT profiles (orange). These three slices were purposefully selected to

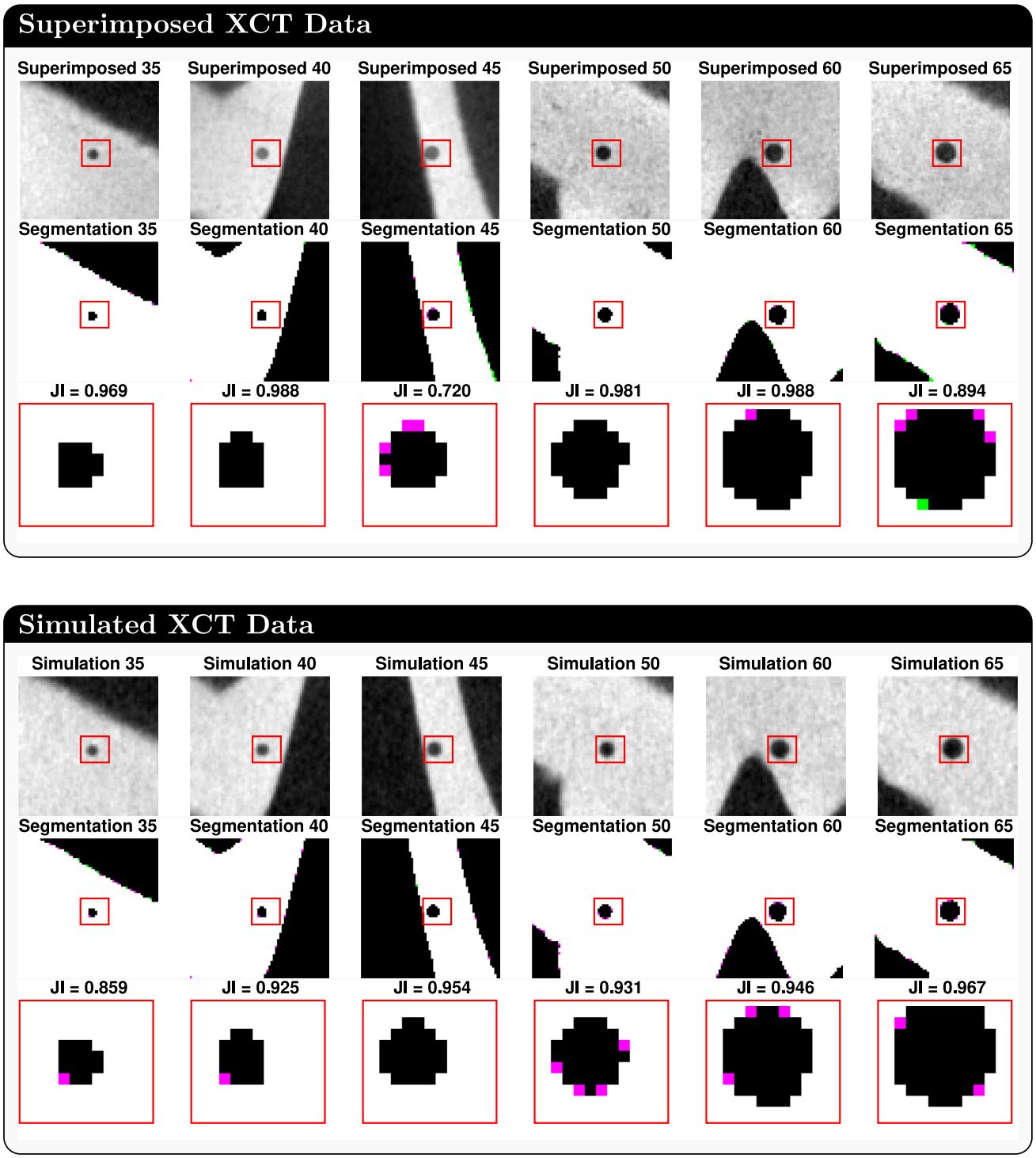


Fig. 6 Example results of **spherical pore** detection using 3D U-Net segmentation. The top block shows results for **superimposed XCT data**, and the bottom block shows results for **simulated XCT data**. Each block includes the original XCT slice, segmentation overlay with ground truth (discrepancies marked in pink and green), and a zoomed-out region. The JI values quantify segmentation accuracy

validate superimposed defect realism and compare different noise and contrast characteristics. The gray value trends align well across datasets, indicating that the superimposed defects produce realistic attenuation patterns, which supports their validity for evaluating detection and segmentation methods. In slice 15, there are two defects—one real and one superimposed. The gray value profiles along the extraction line show similar peak and defect behavior, demonstrating

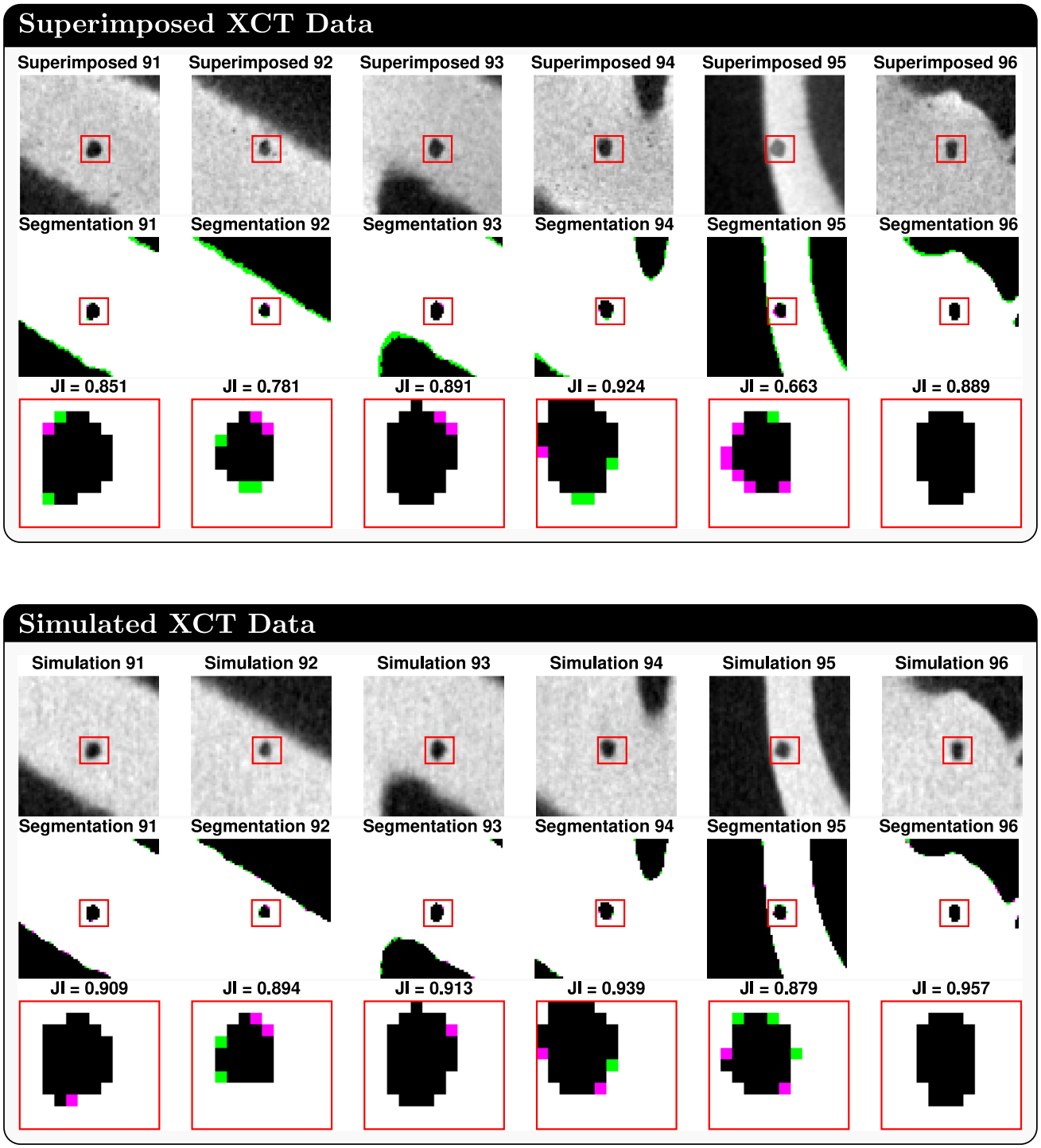


Fig. 7 Example results of **irregular pore** detection using 3D U-Net segmentation. The top block shows results for **superimposed XCT data**, and the bottom block shows results for **simulated XCT data**. Each block includes (top to bottom): the XCT slice with highlighted pores, the segmentation overlay with ground truth (showing discrepancies), and a zoomed-out segmented region. The JI values indicate segmentation performance for each region

that the superimposed defect closely replicates reality both visually and quantitatively.

Figure 10 presents gray value probability histograms from XCT slices at image index 15, evaluating the fidelity of the superimposition. Figure 10a compares superimposed data with purely simulated defect data, revealing noticeable differences, particularly in higher gray value ranges. This indicates that purely simulated data may not fully capture the

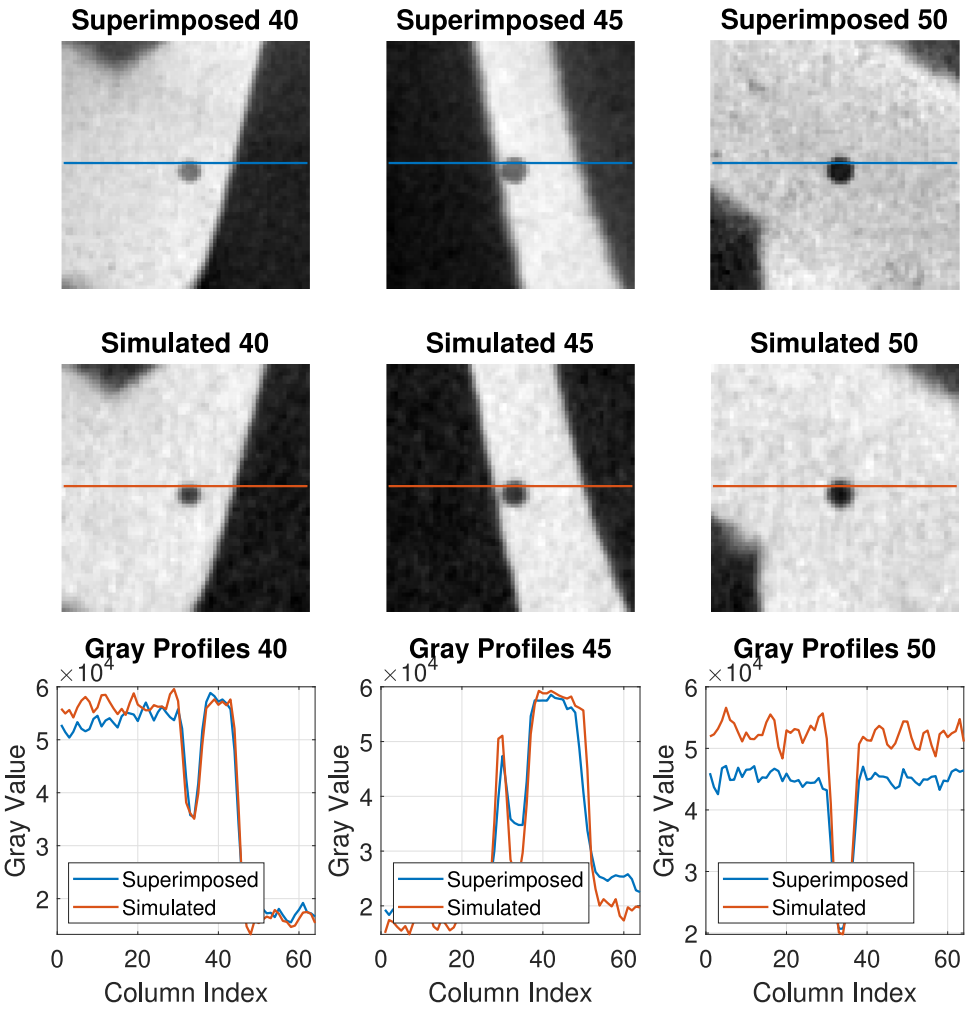


Fig. 8 Comparison of extracted subvolumes and gray value profiles from superimposed and simulated data. The first row presents central XY-slices from the superimposed dataset, and the second row displays the corresponding slices from the simulated dataset. A horizontal line—colored specifically for each dataset—indicates the row along which the gray value profile is extracted. The third row shows the gray value profiles: blue lines represent the superimposed data, and orange lines represent the simulated data

complexity of real scan gray distributions. Conversely, Fig. 10b shows nearly identical distributions between the superimposed XCT slices and the real CT data. This close alignment demonstrates that the superimposition method successfully preserves the inherent characteristics and noise profile of the original CT scan, which is crucial for achieving accurate POD results. Both the gray value line analysis and the histogram comparison demonstrate that superimposed defects exhibit similar characteristics to those in real CT data, confirming improved realism compared to purely simulated data.

4.3.1 Image Quality Metrics and Analysis

In this section, image quality is evaluated through comparisons conducted both on selected slices and across the full volume. Table 2 presents a comparison of image quality between superimposed and real CT slices, as well as simulated and real CT slices, using MS-SSIM [37, 38] and PSNR [39] as quantitative metrics. Across all evaluated slices (15, 17, and 19), the superimposed images exhibit consistently high structural similarity to the real CT data, with MS-SSIM values exceeding 0.99 and PSNR values above 37 dB. In contrast, the simulated images demonstrate lower similarity, with MS-SSIM values ranging from 0.76 to 0.88 and notably lower PSNR values between 18.74 dB and 23.85 dB. The similarity metrics between the superimposed and real volumetric data indicate a near-perfect match, with a MS-SSIM of one and a PSNR of 60.16 dB. In contrast, when comparing the real and simulated data volumes, the MS-SSIM drops to 0.891, and the PSNR to 29.10 dB. These results indicate that the superimposed approach yields image quality closer to real CT, particularly in terms of structural fidelity and noise characteristics.

Table 3 shows the contrast-to-noise ratio (CNR) for each slice and volume across the three data sources. When comparing the real data with the superimposed case, the results are nearly identical across all slices (15, 17, and 19) (see Table 3). However, the simulated images, particularly for slices 17 and 19, exhibit substantially higher CNR values (e.g., 28.51 and 43.67). For slice 15, the CNR values are very close (9.93 vs. 9.48). This suggests that the simulated contrast may not consistently replicate the physical imaging characteristics

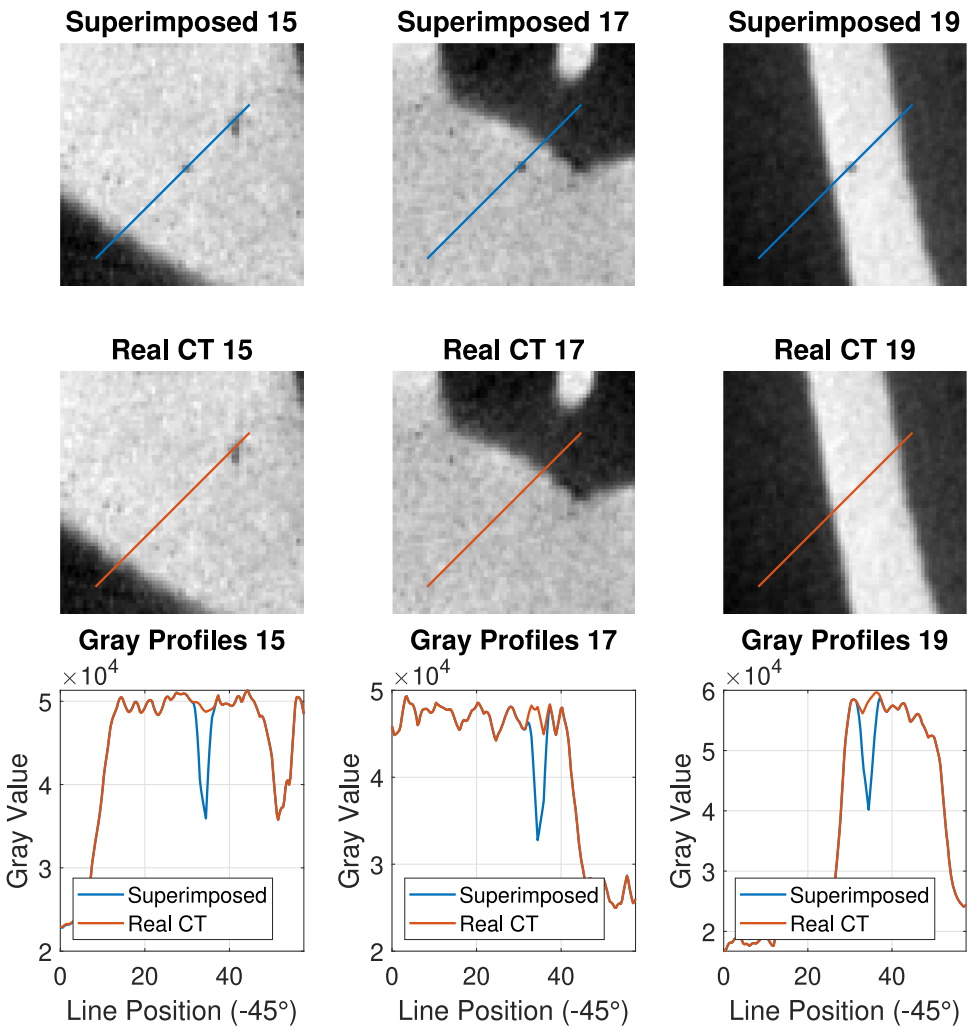
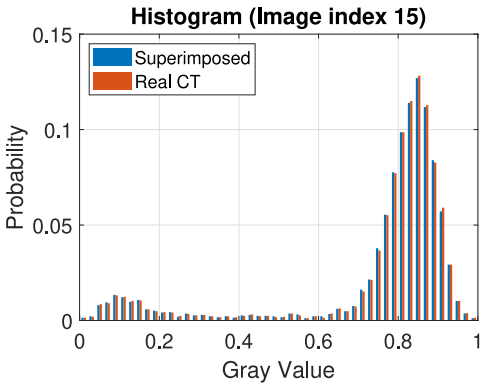
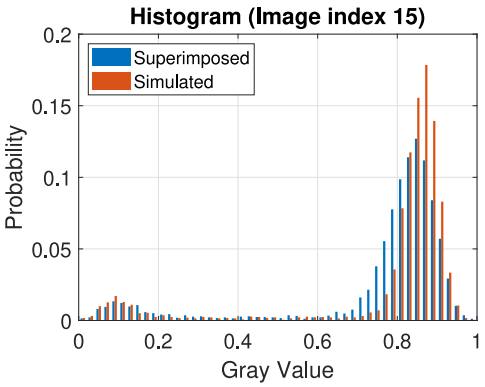


Fig. 9 Comparison of extracted subvolumes and gray value profiles from superimposed and real XCT data. The first row shows central XY-slices from the superimposed dataset, and the second row displays the corresponding slices from the simulated dataset. A diagonal line—colored specifically for each dataset—indicates the row along which the gray value profile is extracted. The third row shows the gray value profiles: blue lines represent the superimposed data, and orange lines represent the real XCT data



(a) Gray value probability histograms for superimposed and simulated XCT slices at image index 15.



(b) Gray value probability histograms for superimposed and real XCT slices at image index 15.

Fig. 10 Gray value histograms (image index 15) comparing superimposed XCT slices with simulated data (a) and with real XCT data (b). The superimposed and real CT distributions show stronger similarity, especially in higher gray values, whereas differences are more pronounced between the superimposed and simulated data

Table 2 Image quality metrics (normalized slices and full volume) – comparison between superimposed vs. real CT and simulated vs. real CT

Region	Superimposed vs. Real CT		Simulated vs. Real CT	
	MS-SSIM	PSNR (dB)	MS-SSIM	PSNR (dB)
Slice 15	0.9908	39.42	0.7611	23.85
Slice 17	0.9936	37.65	0.8078	18.74
Slice 19	0.9974	38.43	0.8845	22.69
Volume	1.0000	60.16	0.9310	29.10

observed in real CT data across different regions, highlighting potential limitations in the simulation model's ability to fully capture the complexity of real-world imaging.

To further validate the accuracy of the alignment, segmentation performance was evaluated on both the real XCT data and the purely simulated data. The segmentation achieved a 3D JI of 0.98, with a precision of 0.99 and a recall of 0.98. These results confirm the robustness of the registration and defect superimposition process. It is important to note that the real XCT data also includes segmentations of existing pores.

4.4 POD Results for Spherical Pores

Figure 11 shows the POD results for spherical pore diameters using a probit regression model, evaluated separately for superimposed and simulated XCT datasets with a segmentation threshold of 0.9%. The red curves represent the fitted probit models, and the shaded areas indicate 95% confidence bounds. Each blue dot corresponds to the outcome of a hit/miss experiment, where a value of one indicates successful pore detection and zero indicates a miss. A notable difference was observed in the minimum detectable pore diameter at 90% probability level (a_{90}): for the superimposed data, $a_{90/95}$ was approximately 219 µm, whereas for the simulated data, it was significantly lower at approximately 193 µm. This indicates that the segmentation algorithm achieves better detectability performance on simulated pores, likely due to reduced complexity or less porosity in the synthetic dataset. The difference highlights how additional artifacts and gray value variations in the superimposed data can degrade detection and segmentation sensitivity. Figure 11 illustrates the POD results in pore diameters in µm along the x-axis, while Figs. 12 and 13 show the number of voxels in the porosity region along the x-axis for POD evaluation.

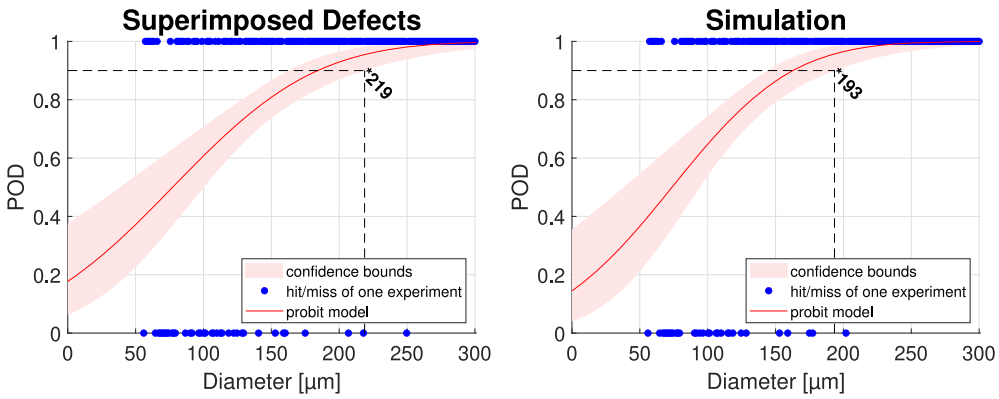
Table 3 CNR comparison across image slices for different image sources

Slice	Superimposed	Real CT	Simulated
15	9.8939	9.9311	9.4897
17	15.9694	15.9409	28.5137
19	35.8017	35.9627	43.6661

The results presented in Fig. 12 demonstrate the influence of voxel similarity (JI values) thresholds on pore detectability using both superimposed and simulated datasets. For superimposed pores, the detectability limit values, $a_{90/95}$ (in voxels) increase with higher thresholds: 86 voxels at threshold 0.85, 255 voxels at threshold 0.90, and 424 voxels at threshold 0.95. In contrast, simulated pores exhibit significantly lower detection limits: 115, 146, and 170 voxels for thresholds 0.85, 0.90, and 0.95, respectively. These values suggest that defects in the simulated dataset are more readily detected, possibly due to their idealized structure, higher contrast, and absence of fully representative background noise. Overall, lower segmentation thresholds lead to earlier detection, but also increase the risk of false positives or false negatives, especially in realistic conditions. The gap in $a_{90/95}$ between the two datasets underlines the importance of the superimposing method as a more realistic approach when estimating detection performance on representative data.

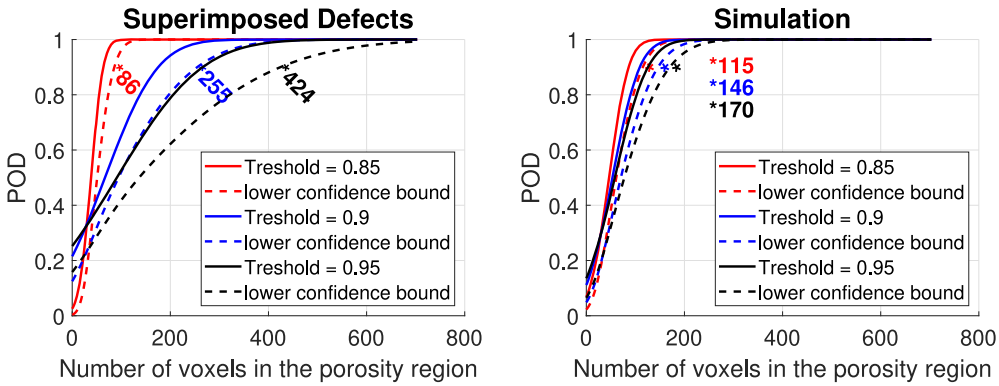
4.5 POD Results for Irregular Pores

Figure 13 presents the POD curves for pore identification using a hit/miss approach in two datasets: (a) superimposed defects and (b) simulated defects. The analysis is performed for three segmentation thresholds: 0.8 (red), 0.9 (blue), and 0.95 (black). For each case, the solid lines denote the fitted POD curves, while the dashed lines represent the corresponding lower 95% confidence bounds. The annotated values along the curves show the $a_{90/95}$ values—the defect size at which 90% detection probability is achieved with 95% confidence. The numeric labels along each curve (e.g., 59, 274, 426 in (a) and 47, 223, 353 in (b)) indicate the detectability limit values, $a_{90/95}$ (in voxels) at which there is 90% probability of detection with 95% confidence. Detection sensitivity is consistently higher across all cases for the simulated data; however, the threshold level has a substantial impact. At the 80% detection threshold, the detection limit for superimposed defects is approximately three times higher than that for simulated ones. When the threshold is increased to 90%, the detection limits between the two cases become more comparable. At the 95% threshold, the detection limit for superimposed defects remains roughly 1.5 times higher than for the simulated data.



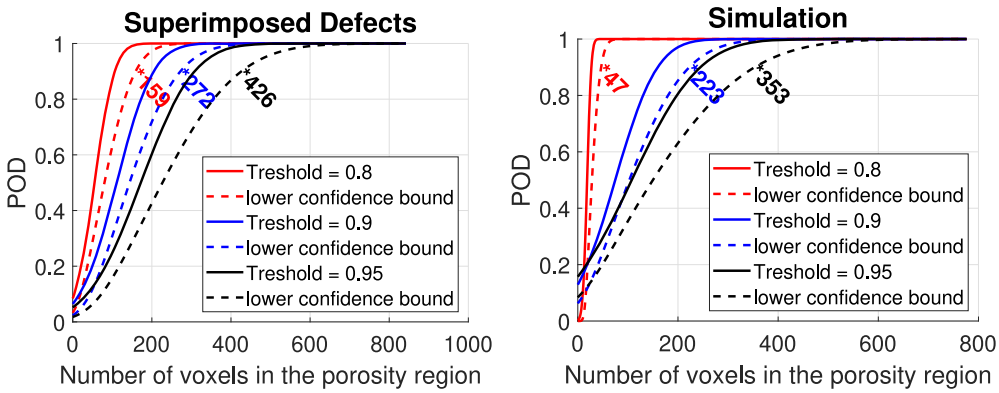
(a) Superimposed pores, threshold 90% (b) Simulated pores, threshold 90%

Fig. 11 POD model (red) for pore detection with confidence bounds based on a hit/miss experiment and $a_{90/95}$ for diameter. Blue dots indicate detected defects (probability one) and undetected defects (probability zero)



(a) Superimposed pores (b) Simulated pores

Fig. 12 POD models for **spherical** pore segmentation using superimposed (a) and simulated (b) defects. POD curves are fitted for three segmentation thresholds (0.85, 0.90, and 0.95), each with associated lower 95% confidence bounds. The x-axis shows the number of voxels within the porosity region, while the y-axis represents the detection probability. Estimated $a_{90/95}$ values are annotated next to each curve, indicating the voxel size above which defects are reliably detected with 90% probability at 95% confidence



(a) Superimposed defects (b) Simulated defects

Fig. 13 Extended POD analysis using adjusted **irregular** pore segmentation thresholds (0.80, 0.90, and 0.95) for both superimposed (a) and simulated (b) defect datasets. As before, each curve includes the lower confidence bound, and $a_{90/95}$ thresholds are displayed. The results illustrate how detection capability varies depending on both the segmentation threshold and the source of the data (superimposed vs. simulated)

As known, lower thresholds lead to higher sensitivity (lower $a_{90/95}$), whereas higher thresholds shift the curves rightward, requiring larger defects to ensure confident detection. Notably, the POD curves derived from superimposed defects exhibit higher $a_{90/95}$ values than those based on purely simulated defects, suggesting improved accuracy of detectability. When comparing the segmentation accuracy of spherical and irregularly shaped defects using the 90% threshold, it can be seen that, as expected, the segmentation accuracy for irregular defects is significantly lower in both the superimposed and simulated cases.

Figures 12 and 13 provide a comparative view of POD performance for spherical and irregular pore shapes, respectively, using both superimposed and simulated datasets across multiple segmentation thresholds. In both figures, superimposed defects consistently yield higher $a_{90/95}$ values compared to simulated ones, indicating a higher voxel count requirement for reliable detection. For spherical pores (Fig. 12), the $a_{90/95}$ estimates in the superimposed case are 86, 255, and 424 voxels for thresholds 0.85, 0.90, and 0.95, respectively, whereas the simulated pores show lower values of 115, 146, and 170 voxels. Similarly, in the case of irregular pores (Fig. 13), the detection limits for superimposed data reach 159, 272, and 426 voxels, while simulated counterparts show reduced values of 47, 223, and 353 voxels. These results clearly illustrate that both shape complexity and segmentation threshold play critical roles in detection performance. Moreover, the consistently higher detection limits in the superimposed datasets underscore the increased realism introduced by this method, thereby providing a more representative benchmark for evaluating defect detectability. Overall, the comparison confirms that relying solely on simulated data may underestimate the practical challenges associated with defect detection in real imaging scenarios.

5 Discussion

In this section, we discuss several aspects of the proposed approach, including noise modeling in superimposed defect simulation, the impact of dataset characteristics on model performance, the measurement uncertainty inherent to XCT data, and computational efficiency and generalizability. We also address some limitations of the current methodology and suggest directions for future improvements.

5.1 Physical Validity and Artifacts in Defect Superimposition

In this study, the simulations used in the superimposition method include beam hardening and scattering artifacts, but do not incorporate a noise model. Beam hardening and scattering effects can be considered consistent and repeatable when simulating the same specimen under identical conditions [40]. When the same specimen is simulated with and without defects, the only differences between the two datasets are localized in the defect regions and neighboring regions of the defects. Therefore, by subtracting the defect-free simulation from the defect-containing one and adding the result to the real XCT data, these nonlinear effects are inherently transferred into the final superimposed output. The current noise implementation introduces random variations that differ across simulations, which can significantly alter the gray value distribution across the volume [41]. This becomes problematic for the subtraction-based method, as the noise changes stochastically on the entire XCT dataset rather than just the added defect information. Instead, noise and other system-related effects are already inherently present in the real XCT data used as the background. As a result, consistent and realistic noise integration into the superimposed data remains a limitation of the current approach. Nevertheless, our findings indicate that the resemblance between real and superimposed defects is greater than that of purely simulated defects, supported by both visual assessment and quantitative similarity metrics. Moreover, by leveraging the existing noise in the real CT scans, we avoid introducing artificial distortions or altering the size and shape of the superimposed defects. While this approach has clear advantages, improving the realism of noise specifically within the defect regions of simulated data remains a valuable direction for future development.

5.2 Impact of Dataset Characteristics on Model Performance

The quality and structure of the input dataset significantly influence the performance of the deep learning model used for estimation [42]. To ensure a fair comparison, we used the same CNN architecture and training parameters across all experiments, with no signs of overfitting or underfitting. Yet, the model's performance varied depending on the dataset, suggesting that data-driven features—particularly those related to shape complexity—affect the CNN's ability to generalize.

For example, when comparing the detection limits between spherical and irregularly shaped pores, the detectability values (in voxels) were higher for irregular shapes in both simulation and superimposed cases. Specifically, in the simulated dataset, there was a noticeable increase—up to two times—for high POD thresholds (e.g., 0.90 and 0.95). While the simulation includes certain physical effects, it tends to produce images with more linear gray value distributions and lower overall complexity compared to real CT scans. As a result, the shape of the porosity has a more pronounced effect on detectability in simulated data, suggesting that the dataset itself is a major contributor to the observed differences.

Interestingly, for the superimposed dataset, the differences in POD results between spherical and irregular pores are relatively small. For a threshold of 90%, the detection limits for spherical and irregular pores are 255 and 272 voxels, respectively, while for 95%, the values are nearly identical (424 vs. 426 voxels). The increased complexity and porosity in the superimposed dataset appear to influence the CNN's learning process, as the model can learn the richer structural information present in the data.

In summary, while both factors play a role, the dataset appears to have a stronger influence on the observed quantitative differences—especially in simulated settings, where the simplicity of the data amplifies shape-based effects. However, the CNN and its training process also contribute, particularly in cases where pore complexity is high and realistic features are present.

5.3 Uncertainty

Measurement uncertainty is a critical aspect in industrial XCT, especially when using XCT data for dimensional metrology or accurate defect detection [43]. In this context, uncertainty encompasses both systematic errors (bias) and random variations (standard deviation or repeatability) in the measurement process. When defects are superimposed onto real CT scans, the resulting data inherently includes the measurement uncertainty already present in the real scan. In contrast, simulated data lacks the inherent and fully representative uncertainty present in real CT scans, and thus, it is important to explicitly introduce physical effects to approximate real-world conditions. In this study, we incorporated several physical models into the simulation pipeline to emulate measurement uncertainty: noise, focal spot blur, modulation transfer function of the detector, scattering, and gain correction. These additions of physical effects into the simulations are essential to mimic the imaging characteristics and simulate the measurement uncertainty observed in experimental CT data. While we did not directly quantify measurement uncertainty in terms of standard uncertainty or bias, we validated the similarity between real and simulated data (volume) both at the image and segmentation levels. Specifically, the MS-SSIM between real and simulated volumes was 93%, indicating strong perceptual similarity. Additionally, segmentation accuracy was evaluated using the 3D JI, which reached 0.98. These high similarity scores support the effectiveness of our registration process and the reliability of the defect superimposition method.

5.4 Computational Efficiency and Generalizability

The superimposition method introduces computational overhead primarily in its post-processing phase, which is not present in purely simulated pipelines. Specifically, the method required 9 minutes 20 seconds to generate a defect-free simulation (with physical effects disabled) and another 9 minutes 20 seconds for a defect-containing version (also with physical effects disabled). An additional 102 seconds were needed to subtract the defect-containing data from the defect-free volume and superimpose the result onto a real CT scan. In total, the superimposition pipeline took approximately 19 minutes. In comparison, generating a fully simulated dataset with physical effects enabled required approximately 14 minutes. Each dataset had a volume size of $823 \times 1146 \times 1968$ voxels, corresponding to approximately 3.6 GB of data. All processing times were recorded on a system equipped with an Intel Core i9-11900K CPU (3.50 GHz, 8 cores, 16 threads) and an NVIDIA GeForce RTX 3090 GPU (24 GB).

While this study focuses on a single geometry, the methodology is generalizable to other component types, materials, and defect types, provided equivalent registration accuracy and contrast conditions are achieved. It can be readily applied to different materials, including various metals, fiber-reinforced composites, and polymer-based manufactured parts.

6 Conclusion

In this study, we proposed an approach for superimposing realistic defects into XCT scans using X-ray simulations. The superimposed and purely simulated datasets were compared for both spherical and irregular-shaped pores through visual inspection and quantitative analysis. The evaluation was performed using gray value profile analysis, JI across different segmentation thresholds, and POD assessment. The results demonstrate that while segmentation performance—quantified by the JI—is generally higher for simulated XCT data, the superimposed datasets provide a more challenging and realistic test environment that better reflects actual XCT conditions. The largest difference in detectability limit was observed for the irregular pore type, where the superimposed case showed a detection limit of 159 voxels using the 80% threshold criterion, compared to only 47 voxels in the purely simulated case. Specifically, for spherical pores, the detectability limit in the superimposed dataset was up to 2.5 times higher (at a minimum voxel similarity of 95%). For irregular pores, this factor increased to a maximum of 3.3 times (at 80% similarity), highlighting the increased complexity of detecting non-spherical defects under realistic conditions.

Both visual and quantitative results demonstrate that superimposing defects can increase the accuracy of POD determination. Additionally, quantitative analysis using the JI revealed that spherical pores are generally segmented with higher accuracy, with an average JI of 0.93 ± 0.06, compared to 0.84 ± 0.01 for irregularly shaped pores, highlighting the

influence of pore morphology on segmentation performance. These observations suggest that the superimposing method, despite slightly lower segmentation metrics, improves the realism and reliability of defect representations, ultimately supporting more accurate and trustworthy determination of the POD. In conclusion, the superimposing approach facilitates a more controlled and realistic assessment of defect detectability, ultimately advancing POD accuracy and contributing to the development of more reliable non-destructive testing methodologies.

For future work, we plan to develop a method that can directly estimate POD curve and detection limits for selected pores within a 2D images or 3D volumes. To achieve this, we intend to explore two different strategies that incorporate information from scanning parameters (e.g., number of projections, voltage), image similarity metrics, noise estimation, and pore characteristics (shape, size). By leveraging these metrics, our goal is to derive a formula capable of directly calculating the detection limit for specific pores. To achieve this and validate the results, superimposed data will play a crucial role. In the second strategy, we plan to use the superimposed data along with the corresponding POD results to train a deep learning model capable of directly estimating the POD from selected pores. This approach aims to significantly accelerate the POD estimation process and eliminate the need to re-simulate data each time a new type of pore or inclusion is encountered.

Author Contributions M.Y. authored the first draft of the manuscript. Additionally, he undertook all aspects of the project, including conception, design, methodology, software development, data generation and analysis, and interpretation of data. B. F. contributed in conception, design and review and editing. J. K. contributed to review and editing. J. S., J. D. B. and C. H. provided support in conception, design, methodology and contributed to review and editing.

Funding This research was co-financed by the European Union H2020-MSCA-ITN-2020 under grant agreement no. 956172 (xCTing). JS acknowledges the Flemish Government under the "Onderzoeksprogramma Artificiele Intelligentie (AI) Vlaanderen programme. Additionally, we acknowledge financial support for the project sustaiNDT (grant nr. 909801) funded within the framework "Production & Materials" of Austrian Research Promotion Agency (FFG). Open access funding provided by FH Oberösterreich Studienbetriebs GmbH.

Data Availability Data can be provided upon request.

Declarations

Declaration of Generative AI and AI-assisted technologies in the writing process Artificial intelligence tools were employed exclusively for linguistic refinement, with no contribution to or influence on the scientific findings or interpretations.

Ethics Approval Not applicable.

Consent for Publication Not applicable.

Competing Interests The authors declare no competing interests.

References

1. Carmignato, S., Dewulf, W., Leach, R. (eds.): Industrial X-Ray Computed Tomography, 1st edn. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-59573-3
2. Heinzl, C., Stappen, S.: STAR: Visual computing in materials science. Comput. Graph. Forum **36**(3), 647–666 (2017). https://doi.org/10.1111/cgf.13214
3. Berens, A.P.: NDE reliability data analysis. In: ASM Handbook. Volume 17 - Nondestructive Evaluation and Quality Control, pp. 689–701. ASM International, Materials Park, Ohio, US (1989)
4. Amrhein, S., Rauer, M., Kaloudis, M.: Characterization of computer tomography scanners using the probability of detection method. J. Nondestruct. Eval. **33**, 643–650 (2014). https://doi.org/10.1007/s10921-014-0258-4
5. U.S. Department of Defense: Handbook Non-destructive Evaluation System Reliability Assessment vol. MIL-HDBK-, p. 1823. U.S. Department of Defense, Washington, D.C. (1999)
6. Yosifov, M., Reiter, M., Heupl, S., Gusenbauer, C., Fröhler, B., Fernández-Gutiérrez, R., De Beenhouwer, J., Sijbers, J., Kastner, J., Heinzl, C.: Probability of detection applied to X-ray inspection using numerical simulations. Nondestruct. Test. Eval. **37**(5), 536–551 (2022). https://doi.org/10.1080/10589759.2022.2071892
7. Yosifov, M., Weinberger, P., Reiter, M., Fröhler, B., De Beenhouwer, J., Sijbers, J., Kastner, J., Heinzl, C.: Defect Detectability Analysis via Probability of Defect detection Between Traditional and Deep Learning Methods in Numerical Simulations. e-Journal of Nondestruct. Test. **28**(3) (2023). https://doi.org/10.58286/27716
8. Kim, F., Pintar, A., Moylan, S., Garboczi, E.: The influence of X-ray computed tomography acquisition parameters on image quality and probability of detection of additive manufacturing defects. J. Manuf. Sci. Eng. **141**(11), 111012 (2019). https://doi.org/10.1115/1.4044515
9. Carver, J.: The impact of background and experience on software inspections. Empir. Softw. Eng. **9**, 259–262 (2004). https://doi.org/10.1023/b:emse.0000027786.04555.97
10. Boas, F., Fleischmann, D.: CT artifacts: Causes and reduction techniques. Imaging Med. **4**, 229–240 (2012). https://doi.org/10.2217/iim.12.13
11. Tenscher-Philipp, R., Schanz, T., Harlacher, F., et al.: AI-driven synthetization pipeline of realistic 3D-CT data for industrial defect segmentation. J. Nondestruct. Eval. **43**(67) (2024). https://doi.org/10.1007/s10921-024-01080-x
12. Brierley, N., Nye, B., McGuinness, J.: Mapping the spatial performance variability of an X-ray computed tomography inspection. NDT E Int. **107**, 102127 (2019). https://doi.org/10.1016/j.ndteint.2019.102127

13. Thompson, A., McNally, D., Maskery, I., Leach, R.K.: X-ray computed tomography and additive manufacturing in medicine: a review. Int. J. Metrol. Qual. Eng. (IJMQE) **8**, 17 (2017). https://doi.org/10.1051/ijmqe/2017015
14. Brock, J., Sutton, M.: Materials science and X-ray techniques. Materials Today **11**, 52–55 (2008). https://doi.org/10.1016/S1369-7021(08)70239-6
15. Reiter, M., Harrer, B., Heinzl, C., Salaberger, D., Gusenbauer, C., Kuhn, C., Kastner, J.: Simulation aided study for optimising industrial X-ray CT scan parameters for non-destructive testing and materials characterisation. In: International Symposium on Digital Industrial Radiology and Computed Tomography, vol. 6, pp. 20–22 (2011). https://www.ndt.net/article/dir2011/papers/p5.pdf
16. Dewulf, W., Bosse, H., Carmignato, S., Leach, R.: Advances in the metrological traceability and performance of X-ray computed tomography. CIRP Annals **71**(2), 693–716 (2022). https://doi.org/10.1016/j.cirp.2022.05.001
17. Makarov, J., Basting, M., Schielein, R., Schmitt, M., Kabitzky, D., Sukowski, F., Voland-Salamon, V.: Computed tomography with or without radiation. e-J. Nondestruct. Test. **28**, 8 (2023). https://doi.org/10.58286/27745
18. Van Aarle, W., Palenstijn, W.J., Cant, J., Janssens, E., Bleichrodt, F., Dabravolski, A., De Beenhouwer, J., Batenburg, K.J., Sijbers, J.: Fast and flexible X-ray tomography using the ASTRA toolbox. Opt. Express **24**(22), 25129–25147 (2016). https://doi.org/10.1364/OE.24.025129
19. Badal, A., Badano, A.: Accelerating Monte Carlo simulations of photon transport in a voxelized geometry using a massively parallel graphics processing unit. Med. Phys. **36**, 4878–4880 (2009). https://doi.org/10.1118/1.3231824
20. Nuclear Energy Agency: PENELOPE 2018: A code system for Monte Carlo simulation of electron and photon transport. OECD, Workshop Proceedings Barcelona, Spain (2019). https://doi.org/10.1787/32da5043-en
21. Reiter, M., Erler, M., Kuhn, C., Gusenbauer, C., Kastner, J.: SimCT: a simulation tool for X-ray imaging. In: 6th Conference on Industrial Computed Tomography (iCT) (2016). https://www.ndt.net/article/ctc2016/papers/ICT2016_paper_id20.pdf
22. Aldrin, J.C., Sabbagh, H.A., Murphy, R.K., Sabbagh, E.H., Knopp, J.S., Lindgren, E.A., Cherry, M.R.: Demonstration of model-assisted probability of detection evaluation methodology for eddy current nondestructive evaluation. AIP Conf. Proc. **1430**(1), 1733–1740 (2012). https://doi.org/10.1063/1.4716421
23. Dominguez, N., Feuillard, V., Jenson, F., Willaume, P.: Simulation assisted pod of a phased array ultrasonic inspection in manufacturing. AIP Conf. Proc. **1430**(1), 1765–1772 (2012). https://doi.org/10.1063/1.4716425
24. Falcetelli, F., Yue, N., Rossi, L., Bolognini, G., Bastianini, F., Zarouchas, D., Sante, R.D.: A model-assisted probability of detection framework for optical fiber sensors. Sensors (Basel, Switzerland) **23** (2023). https://doi.org/10.3390/s23104813
25. Lee, D., Yoon, S., Park, J., Eum, S., Cho, H.: Demonstration of model-assisted probability of detection framework for ultrasonic inspection of cracks in compressor blades. NDT E Int. **128**, 102618 (2022). https://doi.org/10.1016/j.ndteint.2022.102618
26. Zhou, Y., Schneider, L.-S., Fan, F., Maier, A.: A 2D sinogram-based approach to defect localization in computed tomography. arXiv preprint arXiv:2401.16104 (2024)
27. Liu, Y., Wang, X., Zhang, Y., et al.: Application of object detection algorithms in non-destructive testing of pressure equipment: A review. Sensors **24**(18), 5944 (2024). https://doi.org/10.3390/s24185944
28. Iuso, D., Chatterjee, S., Cornelissen, S., Verhees, D., De Beenhouwer, J., Sijbers, J.: Voxel-wise segmentation for porosity investigation of additive manufactured parts with 3D unsupervised and (deeply) supervised neural networks. Appl. Intell. **54**(24), 13160–13177 (2024). https://doi.org/10.1007/s10489-024-05647-z
29. Yosifov, M., Lang, T., Florian, V., Gerth, S., De Beenhouwer, J., Sijbers, J., Kastner, J., Heinzl, C.: Degradation detection in rice products via shape variations in XCT simulation-empowered AI. J. Nondestruct. Eval. **44**(1), 10 (2024). https://doi.org/10.1007/s10921-024-01147-9
30. Sorkine, O., Cohen-Or, D., Lipman, Y., Alexa, M., Rössl, C., Seidel, H.-P.: Laplacian Surface Editing. SGP '04, pp. 175–184. Association for Computing Machinery, New York, NY, USA (2004). https://doi.org/10.1145/1057432.1057456
31. Möller, T., Trumbore, B.: Fast, minimum storage ray-triangle intersection. J. Graph. Tools **2**(1), 21–28 (1997). https://doi.org/10.1080/10867651.1997.10487468
32. Spiess, M., Nagl, W., Hamerle, A.: Probit models: Regression parameter estimation using the ML principle despite misspecification of the correlation structure (1997). https://doi.org/10.5282/ubm/epub.1461
33. Demidenko, E.: Computational aspects of probit model. Math. Commun. **6**, 233–247 (1997)
34. Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3D U-Net: Learning dense volumetric segmentation from sparse annotation. Medical Image Computing and Computer-Assisted Intervention – MICCAI. Lecture Notes in Computer Science **9901**, 424–432 (2016). https://doi.org/10.1007/978-3-319-46723-8_49
35. Hartigan, J.A., Wong, M.A.: Algorithm AS 136: A K-Means clustering algorithm. Appl. Stat. **28**(1), 100–108 (1979). https://doi.org/10.2307/2346830
36. Jaccard, P.: Étude comparative de la distribution florale dans une portion des alpes et des jura. Bull. de la Société Vaudoise des Sci. Nat. **37**, 547–579 (1901)
37. Wang, Z., Simoncelli, E., Bovik, A.: Multiscale structural similarity for image quality assessment **2**, 1398–14022 (2003). https://doi.org/10.1109/ACSSC.2003.1292216
38. Bakurov, I., Buzzelli, M., Schettini, R., Castelli, M., Vanneschi, L.: Structural similarity index (ssim) revisited: A data-driven approach. Expert Syst. Appl. **189**, 116087 (2022). https://doi.org/10.1016/j.eswa.2021.116087
39. Sheikh, H.R., Sabir, M.F., Bovik, A.C.: A statistical evaluation of recent full reference image quality assessment algorithms. IEEE Trans. Image Process. **15**(11), 3440–3451 (2006). https://doi.org/10.1109/TIP.2006.881959
40. Hsieh, J.: Computed Tomography: Principles, Design, Artifacts, and Recent Advances, Third Edition. SPIE Press, SPIE digital library (2015). Chap. 7, Image Artifacts: Appearances, Causes, and Corrections. https://doi.org/10.1117/3.2197756.ch7
41. Lee, S., Lee, M.-S., Kang, M.: Poisson–gaussian noise analysis and estimation for low-dose x-ray images in the nsct domain. Sensors (Basel, Switzerland) **18** (2018). https://doi.org/10.3390/s18041019
42. Zhang, C., Bengio, S., Hardt, M., Recht, B., Vinyals, O.: Understanding deep learning requires rethinking generalization. CoRR arXiv:1611.03530 (2016)
43. Rodríguez-Sánchez, Thompson, A., Körner, L., Brierley, N., Leach, R.: Review of the influence of noise in x-ray computed tomography measurement uncertainty. Precis. Eng. J. Int. Soc. Precis. Eng. Nanotechnol. **66**, 382–391 (2020). https://doi.org/10.1016/j.precisioneng.2020.08.004