



Degradation Detection in Rice Products via Shape Variations in XCT Simulation-Empowered AI

Miroslav Yosifov^{1,2} · Thomas Lang³ · Virginia Florian³ · Stefan Gerth³ · Jan De Beenhouwer¹ · Jan Sijbers¹ · Johann Kastner² · Christoph Heinzl^{3,4}

Received: 31 January 2024 / Accepted: 3 December 2024 / Published online: 16 December 2024
© The Author(s) 2024

Abstract

This research explores the process of generating artificial training data for the detection and classification of defective areas in X-ray computed tomography (XCT) scans in the agricultural domain using AI techniques. It aims to determine the minimum detectability limit for such defects through analyses regarding the Probability of Detection based on analytic XCT simulations. For this purpose, the presented methodology introduces randomized shape variations in surface models used as descriptors for specimens in XCT simulations for generating virtual XCT data. Specifically, the agricultural sector is targeted in this work in terms of analyzing common degradation or defective areas in rice products. This is of special interest due to the huge biological genotypic and phenotypic variations occurring in nature. The proposed method is demonstrated on the application of analyzing rice grains for common defects (chalky and pore areas).

Keywords X-ray computed tomography · Shape variations · Probability of detection · X-ray simulation · Deep learning · Segment anything

1 Introduction

Rice is one of the main products for feeding more than half of the world's population; therefore, its quality is of significant concern [1]. Considering the exponential increase in the global population, it is estimated that the demand for agricultural products will double over the next 30 years [2]. So, improving the yield of harvest by ensuring quality of seeds and food products will play a crucial role in near future. Additionally, eliminating poor-quality rice grains is equally important for improving the quality of rice grains in terms of superior appearance quality [3]. One of the critical issues is

chalkiness, an attribute signified by an opaque, floury appearance in rice grains which compromises their visual appeal, milling yield, and gastronomic qualities [4, 5]. Chalkiness is substantially influenced by environmental conditions, particularly temperature. So, the increase in temperature due to global warming is a growing concern for rice quality. Elevated temperatures during the grain filling period of the rice crop can enhance the propensity of chalky grain formation, threatening the grain's marketability and milling yields [6, 7]. Therefore, the detection of chalkiness in rice grains has emerged as a necessary need, critically linking our fight against climate change with the pursuit of food security and economic stability.

X-ray computed tomography (XCT) can help in this regard as a powerful imaging technique used for non-destructive testing purposes in industrial quality control. As such, it provides detailed and accurate information about the internal structure of specimens without causing any structural damage. XCT is often used to analyze, visualize and explore defects such as pores, fibers, or cracks [8, 9]. Furthermore, it has proven to be highly valuable for quality assessment and inspection of industrial objects, as well as for examining the internal structure of food products [10–13]. In this area the analysis of chalkiness of rice grains, rotten areas of fruits, or

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

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

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

internal browning, capturing shape variations becomes crucial [14, 15]. For example, rice grains may vary significantly in their main and chalky appearance [16], while fruit may exhibit different degrees of rot in various regions.

For the investigation of the detectability of such unwanted structures, the concept of Probability of Detection (POD) has gained substantial importance in recent years in the field of X-ray-based digital radiography and XCT [17–19]. A POD analysis quantifies the reliability and effectiveness of a defect detection algorithm and is used in the evaluation of imaging systems, processes and techniques. It provides a quantitative measure of the likelihood of correctly identifying and characterizing defects, ultimately aiding the improvement of inspection processes and system performance [20]. For this purpose, incorporating a broad variety of shape variations in the training data is of utmost importance in AI-driven systems in order to obtain robust models [21] for detection. This is especially true in industrial inspection, where objects exhibit complex geometries and defects in diverse forms.

In this work, we focus on the application of generating artificial data of rice grains in terms of shape modeling chalky parts. However, the proposed approach may equally be employed for other use cases, e.g., to detect rotten parts for other food products. By simulating shape variations of rice grains and utilizing advanced deep learning techniques, our methodology thus aims to accurately identify and differentiate rice grains with chalky and corrupted part appearances.

2 Related Work

For the proposed method, algorithms for modeling shape variations and their applications are fundamental components. Furthermore, methods for the detection of rice chalkiness are explored and compared to the proposed method. Finally, unsupervised and supervised deep learning methods for the extraction of relevant features are considered and reviewed.

2.1 Rice Chalkiness Detection

The detection, characterization and finally understanding different levels of chalkiness provides valuable insights into the quality of the crop and its management [22]. Chalkiness can be influenced by various factors such as genetic traits, growing conditions, and post-harvest practices [5]. Thus, the detection and measuring of rice chalkiness is of utmost importance in all stages of rice production and marketing, from breeders developing low-chalkiness varieties, farmers trying to reduce post-harvest losses, to traders and millers aiming to deliver high-quality rice to consumers.

In the past decade, several applications have been proposed for the detection of rice chalkiness. A first approach in

this regard based on high resolution micro-CT imaging was presented by Su and Xiao [23] in 2020. For the segmentation of rice chalkiness, the authors applied threshold and contrast adjustments using the Mimics Innovation Suite software. Another recently published approach demonstrates the use of a convolutional neural network to differentiate between chalky and non-chalky grains, followed by the application of Grad-CAM [7] to highlight the specific areas of a grain that indicate its classification as chalky.

The genetic anatomy and growth conditions significantly contribute to the cultivation of healthy rice. In connection with this, Piao et al. [22] demonstrated an approach for the identification and characterization of the chalkiness endosperm gene CHALK-H in rice. Their findings show insights into the role of CHALK-H in leaf color and grain development. Another important aspect is the analysis of the chalkiness component concerning the percentage of chalkiness, both for individual rice grains and groups, providing a comprehensive insight into rice quality. The evaluation of chalkiness in rice can be categorized into four scales: Scale 0 (no chalkiness), Scale 1 (chalkiness area less than 10%), Scale 5 (chalkiness area ranging from 11–20%), and Scale 9 (chalkiness area exceeding 20%) [24]. The percentage of chalkiness in medium-grain rice is also analyzed by Tam et al. [25]. They presented the average chalkiness result for 33 rice grains in each scale.

In contrast to these approaches, we apply deep learning methods for the segmentation and detection tasks employing virtual X-ray data to create a deep learning model. Using XCT scan data, multiple rice grains are evaluated regarding their chalkiness and the probability of detection for task-specific cases.

2.2 Shape Modeling and Variation

Shape modeling and generating different variations of an object is an essential method to simulate a distribution of data similar to the naturally occurring variance. This technique is especially useful in instances where data is limited or difficult to gather [26, 27]. By manipulating attributes such as size, orientation, texture, and color, among others, researchers can artificially expand their ensemble of datasets, improving the robustness and generalizability of their algorithms, models, or techniques. In our application, we introduce a methodology for generating random shape variations of rice grains, including the chalkiness or defective parts. The objective of this novel approach is to provide a richer, more diverse dataset for training convolutional neural networks (CNNs). The presented method is not limited on the application of rice grains. It is generalizable to other types of shape variations in fruits products.

2.3 Supervised and Unsupervised Deep Learning Methods

Detecting and extracting interesting features in image data are crucial steps in medical, material and agricultural science. Over the last 20 years, numerous techniques have been introduced to develop accurate segmentation algorithms, incorporating both traditional image processing methods as well as the application of deep learning techniques [28, 29]. Especially in the realm of deep learning, particularly with the advancements seen in CNNs, significant progress has been achieved in the field of segmentation tasks [30, 31]. Nevertheless, there is still a need for improvement, especially considering the segmentation of small features such as pores or defects in material science, early-stage tumor detection in clinical imaging, or the identification of rotten parts in agricultural products. Deep learning itself is a broad field within the machine learning domain and may also be coarsely classified into two types: supervised and unsupervised methods [32, 33]. Supervised methods generally infer knowledge from a training data set having labels associated to it. In the context of segmentation, a typical setup considers segmented training images (one label per pixel) and trains a neural network which maps an unannotated image to its segmentation. During inference, an unseen image is propagated through the network and an according segmentation is generated. Contrary, unsupervised methods infer knowledge using the training data samples only without any labels. Within segmentation, this means that a segmentation is achieved mostly by clustering, i.e., grouping similar learned features in feature space [32].

The remarkable results, which were facilitated by U-Net [34] in (bio-)medical applications have also triggered their introduction to material science, particularly in the precise segmentation of pores, a vital process in the analysis of data and quality inspection of specimens. A noteworthy example is the work of Yosifov [35], which introduced structural modifications to the 3D U-Net architecture for the application on XCT datasets of fiber-reinforced composites. However, the generalizability of the created CNN model to other datasets at different resolutions or datasets generated by different XCT devices poses an open problem. To mitigate these effects, it is crucial to train the model on a diverse set of data that covers a wide range of scenarios and variations. An example of a well-generalized unsupervised segmentation algorithm is the Segment Anything Model (SAM) [36], published in 2023, which is capable of multi-class image segmentation in a variety of contexts. SAM was designed and trained with 1.1 billion segmentation masks on 11 million respective images. SAM integrates three stages: an image encoder, a flexible prompt encoder, and a fast mask decoder which is build on vision transformer models [37–39]. Yet, it is applicable only to 2D images. The weaknesses and strengths

of this method were discussed, e.g., in the context of medical data by Mazurowski et al. [40].

In our current approach, we investigated supervised (U-Net [34], V-Net [41] as well as unsupervised (Vector-Quantized Variational model [42]) deep learning methods 3D chalkiness and pore detection via simulation-empowered AI methods. Although these models share a common foundational design, they each have unique features that have an influence on segmentation performance and accuracy. The techniques were thus compared regarding their segmentation results in order to determine which one is more suitable for the detection of bad parts in rice products.

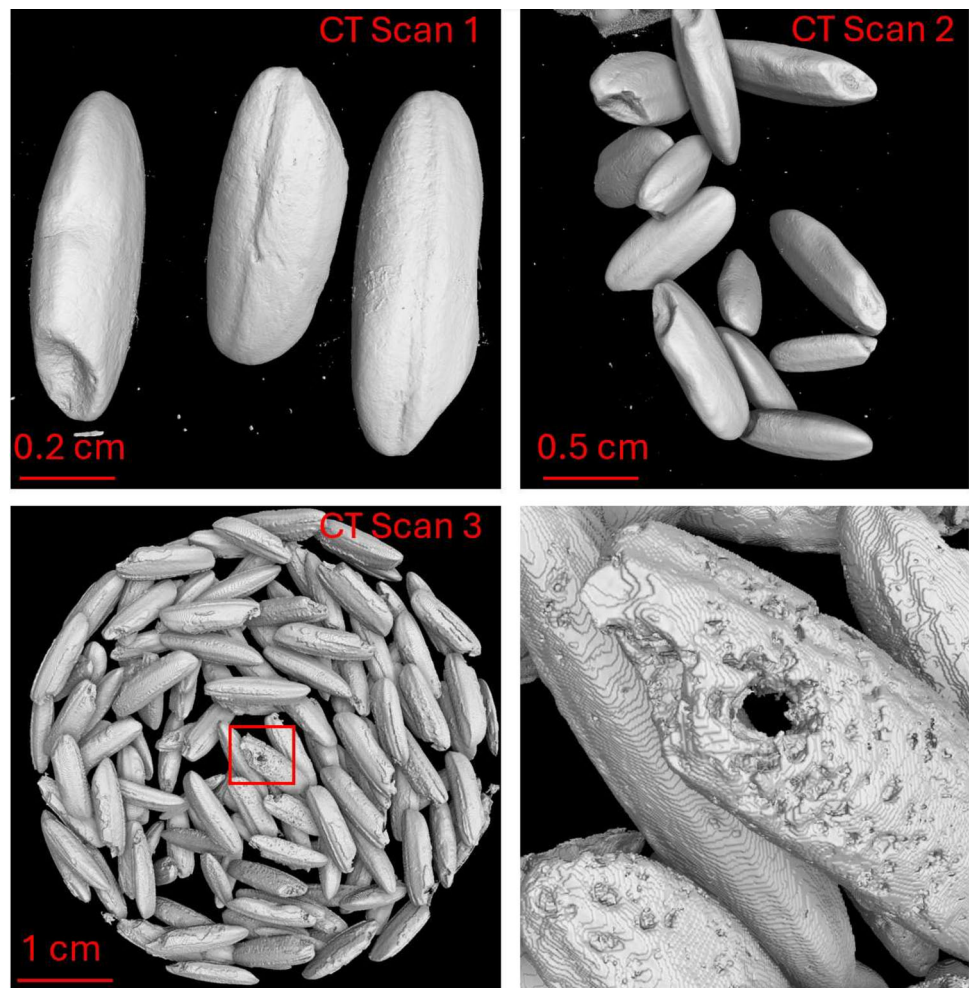
3 Methods and Investigations

In this section, we outline our methodology concerning shape variations of a given surface model, the segmentation of bad parts, and the assessment of Probability of Detection, in relation to chalky parts within simulated XCT data of rice grains.

3.1 Real XCT Data Generation

The generation of real XCT data is an important step to integrate realistic information, as well as to test and validate the model trained on virtual XCT data. Therefore, XCT data was acquired using a RX Solutions Easytom 160 imaging device. A total of 180 rice grains were scanned in three scanning campaigns using different resolutions of 12, 17 and 17.48 μm (see Fig. 1). The CT scans at 12, 17, and 17.48 μm resolutions were generated on different rice grains. We individually analyzed 180 rice grains using XCT data to evaluate porosity and chalkiness. In scans with resolutions of 12 and 17 μm , neither chalkiness nor porosity was observed. However, in the 17.48 μm resolution scans, while chalkiness remained undetected, porosity was successfully segmented and identified with different set of rice. The largest defect (air inclusion) was then exported as an STL model using VG Studio Max (see Fig. 1). Variations of this STL model were then integrated into the rice grain to generate XCT simulations and testing data (see Fig. 2). To simulate the training data, a spherical internal pore was exported, and its variations were created with different sizes and shapes. From Scan 3, we quantified that five rice grains exhibited damage (see Fig. 1), and most of the remaining grains contained small pores (less than 40 μm). The X-ray source operated at an acceleration voltage of 80 kV and anode current of 68 μA . A consistent integration time of 143 ms was maintained throughout the scanning procedure. The detector, configured at a resolution of 1920 \times 1536 pixels, facilitated the precise reconstruction of acquired images. A total of 2112 projections were created during the scanning process. The distances between the X-ray source and the scanned object (21.26 mm) and the source to the detector

Fig. 1 Visualizations of rice grains obtained through three different XCT scans, and highlighting the zoomed-in areas of damage from CT scan 3



(224.99 mm) were strategically chosen to achieve an optimal imaging geometry yielding a magnification factor of 1.2766, and a voxel size of $12\mu\text{m}$. The entire scanning cycle was completed within 35 min. The detailed scanning parameters used for real XCT data ($12\mu\text{m}$ resolution) generation can be found in Table 1. Scan 2 shares similar parameters with Scan 1, except for a different voxel size of $17\mu\text{m}$ and current is $64\mu\text{A}$. For Scan 3, the magnification was 2.83, resulting in a voxel size of $17.48\mu\text{m}$, with 2200 projections generated for the reconstruction.

3.2 Processing and Simulation-Empowered Training Data Generation

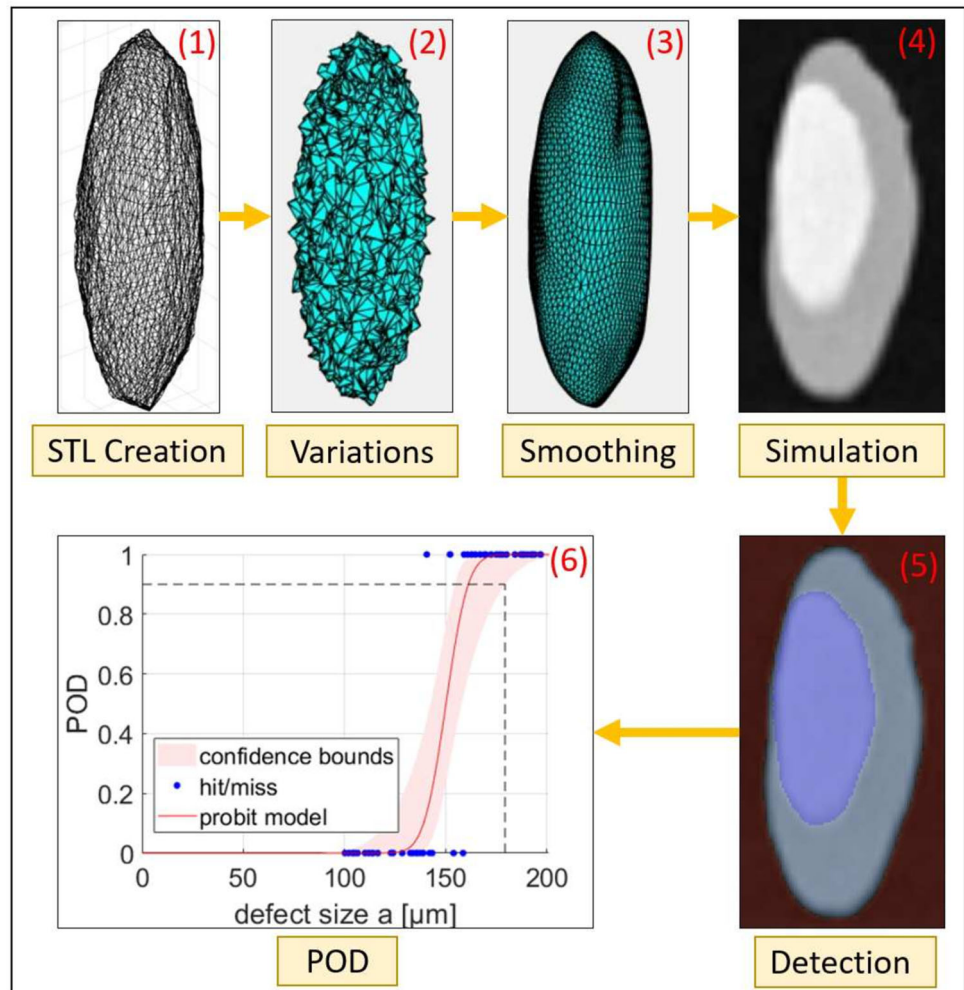
For the generation of simulation-empowered training data, the XCT simulation tool *SimCT* [43] was used, incorporating models of the source, the specimen, the detector and the XCT imaging process. The RX Solutions Easytom 160 imaging device used for the generation of the real data was modeled and parameterized in terms of scanning geometry, including source-object distance, source-detector distance,

focal spot size, etc. in *SimCT*. Additionally, the XCT components, encompassing source, detector, rotary table, etc., were incorporated along with their respective specifications. Finally, physical phenomena such as noise, beam-hardening and scattered radiation were introduced to the radiographs.

As shown in Fig. 2, in our application scenario, we first extracted surface models of the rice grains as STL models from real XCT scans (1) using VG Studio MAX, Version 2023.4. The chalky part was artificially defined and modeled based on the shape of rice grains in our initial version, while the bad parts (air inclusions) were extracted from the real scans instead. Then, the shapes and locations of the chalky parts were adjusted via shape variations based on images from a previous study [23]. As a starting point, the density of the rice grains was set at 1.45 g/ml [44], while the density of the chalkiness region was experimentally determined to be approximately 0.6 g/ml . To determine this value, we optimized grey values on the reconstructed images, visually comparing the color levels between the rice grains and chalkiness, based on an image from a previous study [23]. Next, the shape variation method proposed in Sect. 3.4 was

Table 1 Parameters for XCT Imaging for CT Scan 1 with resolution $12\mu\text{m}$

Parameters	Unit	Value	Parameters	Unit	Value
Acceleration voltage	kV	80	Source-object-distance	mm	21.26
Anode current	μA	68	Source-detector-distance	mm	224.99
Integration time	ms	143 ms	Magnification	–	10.58
Detector pixels	–	1920×1536	Voxel size	μm	12
Binning	–	1	Focal spot size	mm	1
Number of projections	–	2112	Total cycle time	min	35

Fig. 2 The schematic proposed processing pipeline including an example POD curve for the chalky rice grain use case

employed to generate 240 randomly modified variants of the original mesh for both the shape of the rice grain itself and also the chalky part (2). To ensure the successful creation of POD curves and avoid perfect separation between detected and undetected chalkiness, a diverse range of chalky part sizes were created for the XCT simulations. The defect size, specifically its height, was varied randomly between 10 and $400\mu\text{m}$. Additionally, we carefully defined a specific position for the defect at the edge, and, in order to simulate a critical source of uncertainty to the specimen's placement, we randomly shifted this position by ± 0.5 voxels. This was fol-

lowed by Laplacian smoothing for obtaining a more natural shape than simple randomly offset vertices (3). Using XCT simulation, we accordingly generated XCT datasets for each shape (4). Given the virtual XCT scans, all aforementioned segmentation methods were employed for the detection of the chalky and bad parts in the rice grain (5). In the final stage, we determine the minimum detectability by using POD (6). Then, the efficiency of different deep learning architectures is compared using the minimum detectability values obtained from the POD analysis. A detailed explanation of the ground truth generation segmentation and detection pro-

cess is provided in Sect. 3.5. The ground truth images were generated with an XCT simulation where all implemented physical effects were disabled to precisely control chalkiness size and locations. Then, XCT simulations were segmented with k-means [45] and a binary threshold was applied to create ground truth images. Figure 2 illustrates the overall data processing pipeline employed in this study. Additionally, we created physical informed simulations with and without noise to measure the noise effect on chalkiness detection.

3.3 Deep Learning Models and Training Process

In order to detect the aforementioned defects, 3D segmentation methods are necessary. This process ensures a comprehensive and visual analysis, identifying and characterizing defects in different spatial dimensions regarding their size, shape and orientation. As SAM demonstrated proficiency in segmenting objects within medical datasets, we utilized SAM in our current approach for instance segmentation and defect detection for bad parts of rice on simulated projections and slices of reconstructed volumes. A pretrained vision transformer [46] model which is called ViT-B¹ was applied for the segmentation task with default settings. As a supervised method, 3D U-Net and V-Net were applied individually for the 3D detection of chalky and porous parts. Considering unsupervised methods, most commonly an autoencoder architecture is used, in particular variational autoencoders (VAEs). In principle, it is possible to perform the detection task with VAEs simply by training it on clean images using a L1 or L2 reconstruction loss. However, highly representational VAEs can even reconstruct unseen anomalies. Thus, as unsupervised method we adopt a vector-quantized variational autoencoder (VQ-VAE) [42] paired with an autoregressive model (AR). In particular, an anomaly is encoded in latent space differently from a defect-free sample, implying that a model (here an AR model), trained to learn the defect-free distribution from the VQ-VAEs latent representation, can detect anomaly information as out-of-distribution. For the encoding of sample images, we adopt a VQ-VAE in order to quantize the defect-free space. A PixelCNN architecture is adopted as AR prior [47]. The architecture introduced in this paper is referred to as VQ-VAE+AR.

For the supervised training processes of the used CNNs, we generated three distinct sets of training data: one set with 240 XCT scans, comprising 480 artificially created rice grains; another one with 240 XCT scans, comprising 480 rice grains extracted from real scans for chalkiness; and a final set featuring 240 XCT scans, comprising again 480 artificial rice grains that include pores inside. Porosity was

generated using an exported spherical pore and its variations, with five pores assigned to each rice grain at different locations. To ensure an accurate training process, the training data was augmented by XCT simulations carefully designed to include a diverse range of both in rice grain and defective part sizes and variations. The training input size was $128 \times 128 \times 128$. The simulated images had a voxel size of $24 \mu\text{m}$, with 2×2 binning applied to speed up the simulation process and reduce the size of the resulting 3D volumes. Each training input, reconstructed from the simulated projection images, consisted of two rice grains: one with defects and one without. For unsupervised methods, the same training sets were utilized, and additionally, defect-free simulations were generated for the training process. The models for chalkiness detection were trained using a combination of two sets from chalkiness data with using 240 total samples. For pore detection, the models were trained with 240 samples that include pore inside. Additionally, two distinct set of 240 testing samples was created for testing stage of deep learning methods, covering both the chalkiness and porosity tasks.

3.4 Shape Variation of Surface Models

The most crucial part in our approach is the generation of randomized shape variations based on a surface model, i.e., a surface mesh (see Fig. 3). Consider such a surface mesh and its set of vertices $M \in \mathbb{R}^{3 \times N}$ represented as a matrix of N three-dimensional vectors. Furthermore, let $V \in \mathbb{R}^{3 \times N}$ be a matrix whose elements are random numbers samples from a standard normal distribution, i.e., $V_{ij} \sim \mathcal{N}(0, 1)$ for all $V_{ij} \in V$. To allow for a modification of the magnitude of change, let $l \in \mathbb{R}_{>0}$ be some scaling factor. Then, the randomized shape variation of the mesh vertices M is given by the transformed vertex coordinates $W = (M + lV)s$, where $s \in \mathbb{R}_{>0}$ is an additional scaling factor to control the overall size of the resulting shape.

As this simple procedure randomly shifts vertices in or out of the initial boundary position, the resulting modified shapes tend to have a lot of “spikes” on their surface. In order to keep these models more closely to natural rice grains, a subsequent Laplacian smoothing [48] step was applied with five iterations. As reasonably accurate results with smooth surfaces were achieved after five iterations, higher values were not applied.

3.5 Probability of Detection

In our approach, we applied the probability of detection (POD) [17–19] method using a hit/miss approach [17] to quantify the detection capability of the chalky parts in rice grains. Here, POD serves as a statistical metric for assessing the effectiveness of an inspection process in terms of detecting specific features of interest. To generate POD curves, we

¹ ViT-B is available on github <https://github.com/facebookresearch/segment-anything>.

Fig. 3 Schematic representations of shape variations in five different modeled rice variants regarding surface displacement of the rice grain which is extracted from CT scan

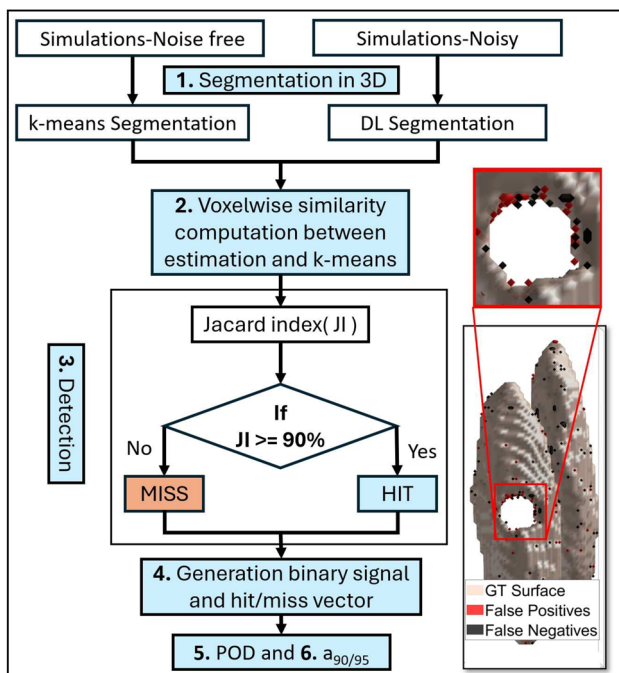
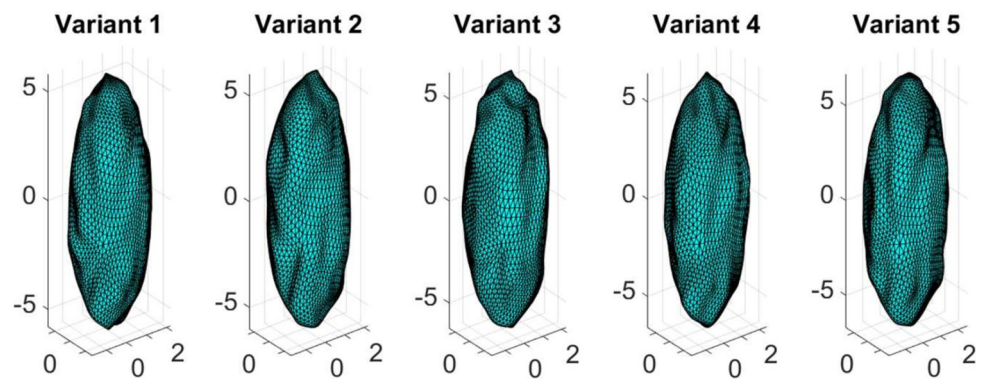


Fig. 4 Workflow for ground truth generation and measurement of the detectability limit using hit/miss analysis and POD. Additionally, an example image is presented, illustrating segmentations from the estimation overlaid on the GT surface, highlighting false-positive and false-negative voxel representations (Color figure online)

collected a binary response in terms of detection by each model for each defect size. Here, these responses correspond to the hit/miss approach, i.e., indications, if the model was able to detect the result. Given the collection of these responses, a Probit function was fitted to the responses using logistic regression [20, 49]. The resulting continuous curves are the POD analysis result curves. We define a_{90} as the defect size at which the POD reaches a value 0.90. Additionally, we calculate further the characteristic value $a_{90/95}$, representing the lower 95% confidence level of a_{90} . This so called detectability limit provides a measure of the smallest defect size that can be confidently detected at a 95% confidence level.

In the first step of the workflow (see Fig. 4), segmentation was performed using k-means and deep learning (DL) algorithms. Next, each voxel was compared between the k-means and deep learning algorithms in the binary volume, and the Jaccard similarity index (JI) was calculated to create a binary detection signal. A detection was defined as a hit if the JI was greater than or equal to 90%; otherwise, it was considered a miss. Finally, we generated the binary detection signal (zeros or ones) and created a hit/miss vector. We then determined the probability of detection (POD) and the minimum detectable limit for different types of DL segmentation methods.

The following steps (see also Fig. 4) were defined to measure detectability limit:

1. Segmentation of samples.
2. Voxelwise similarity computation between estimation and ground truth.
3. Chalkiness and porosity detection.
4. Generation of a binary signal of detected and undetected defects (hit/miss vector).
5. Creation of POD curves.
6. Determination of $a_{90/95}$ for chalky part and bad parts.

In our previous works [20, 31], we defined defect detection as a binary signal that checks for the presence of any defect voxel. To improve the efficiency of this detection method, the defect detection method was updated by incorporating voxelwise size similarity metrics such as the Jaccard index (Eq. 1).

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \quad (1)$$

Chalkiness or bad parts are defined to be detected if the similarity index is 90 percent or higher, i.e., if the detected defect and the ground truth defect overlap by at least 90%. The k-means segmentation method was applied to noise-free simulations to generate accurate ground truth data for 3D voxel-wise comparison with deep learning methods.

4 Results and Discussion

Our methodology demonstrates its applicability in conducting various types of detection of defects and degradation in rice samples. In this work, we focused mainly on artificially generated data as well as supervised and unsupervised algorithms for defect detection. To detect features (chalkiness or pore) on 2D projection images, we applied the SAM method. To detect features (chalkiness or pore) on 3D reconstructed images, we utilized the 3D U-Net and V-Net models as a supervised method and a vector-quantized variational autoencoder paired with an autoregressive model (VQ-VAE+AR) as unsupervised method. Figure 5 shows real projection and simulated projections with their segmentation's from SAM. Obviously, the modeled artificial rice grains are quite realistic in shape, also considering the chalky part in the top-most grain. Furthermore, the segmentation model is able to accurately distinguish the individual rice grains from the chalky parts and from each other.

4.1 Chalkiness Detection via SAM

In the first application case, we focus on chalkiness detection in rice grains for which we applied and evaluated SAM, resulting in 240 different segmentations of rice and chalkiness parts. We generated XCT simulations with and without noise to measure detectability limits and compare the results between these conditions. As expected, we observed that the SAM method demonstrates a significantly better performance on noise-free images than noisy data. The POD results indicate minimum detection limits of 162 μm for noise-free images and 368 μm for noisy images, respectively. The segmentation outcomes for noise-free and noisy scenarios are illustrated in Figs. 6 and 7, showcasing the efficacy of the SAM method. Moreover, the POD results from SAM demonstrate that the detection capability is approximately two times higher in noise-free images than in the noisy case (see Fig. 8).

4.2 Chalkiness Detection via CNNs

We evaluated three different detection algorithms for chalkiness detection: two supervised methods (3D U-Net, V-Net) and Vector-Quantized Variational method as unsupervised method, coupled with an autoregressive model. Two representative rice type dataset samples including chalkiness and pores parts generated using XCT simulation series are presented with the corresponding example of 3D visualization of rice grains in Fig. 9. In the classification evaluation, true positives (TP) signify instances where the model correctly identified the positive class, aligning with the actual positive ground truth. False positives (FP) represent cases where the model inaccurately predicted the positive class for negative ground truth, and false negatives (FN) represent instances

where the model missed positive instances. To measure similarity, the Jaccard index was calculated voxelwise between all scan instances, considering true positives (TP), false positives (FP), and false negatives (FN). Then, POD curves were computed with respect to number of voxels in the chalky region and height of the chalkiness parts. Moreover, number of voxels in the chalky region was determined for both noisy and noise-free volume cases (see Fig. 10). The threshold for the similarity index is again set at 90% for general cases. The minimum detectability achieved for considering the number of voxels of volume at 319 μm and for height 109 μm , respectively, see Fig. 11, in realistic images (on noise inclusion). These results show that a 3D volume analysis of rice grains using CNN methods is able to outperform 2D analyses, e.g., using the SAM algorithm. The minimum detectability for chalkiness height using the unsupervised deep learning method VQ-VAE+AR was measured at 164 μm . However, during our observations, we identified the presence of extra segmented voxels, classified as artifacts, particularly in the corners of rice grains. In the detection process, only the chalkiness similarity has been checked for the unsupervised methods. Both supervised and unsupervised methods were trained with an equal number of training sets (240 samples). Under these conditions, it was observed that supervised methods outperformed the unsupervised methods in terms of accuracy and efficiency for chalkiness detection. To enhance the accuracy and efficiency of rice chalkiness detection in the unsupervised method, it is necessary to increase the amount of training data.

4.3 Pore Detection via CNNs

In the second application case, we focus on the detection of pores, and determined the POD curves considering again the pore's volume in voxels. We employed CNN methods, namely 3D U-Net and V-Net for pore detection in random pore occurrences. Both methods generated similar results, with a detection limit determined to be approximately 200 μm (refer to Fig. 12). Additionally, the POD is determined at various values for the threshold of the Jaccard index, as shown in Fig. 13. The detectability limits for chalkiness (in voxels) determined to be 25, 157, and 319 voxels corresponding to thresholds of 53, 75, and 90%, respectively, using the 3D U-Net model. The POD for the total voxel volume were determined as 84, 151, and 220 voxels corresponding to thresholds of 50, 75, and 90%, respectively, using the 3D U-Net model for pore detection (see Fig. 13b). When comparing the results of chalkiness and pore detection, CNNs tend to be more efficient and accurate, particularly in the context of pore detection. Figure 15 shows the example simulations and estimations of chalkiness results from U-Net. Additionally, Fig. 14 presents example 2D slice images of pore detection obtained from 3D reconstructed data using U-Net and V-Net.

Fig. 5 The figure show, in respective order, real projection alongside its segmentation from SAM, and simulated projection alongside its segmentation from SAM

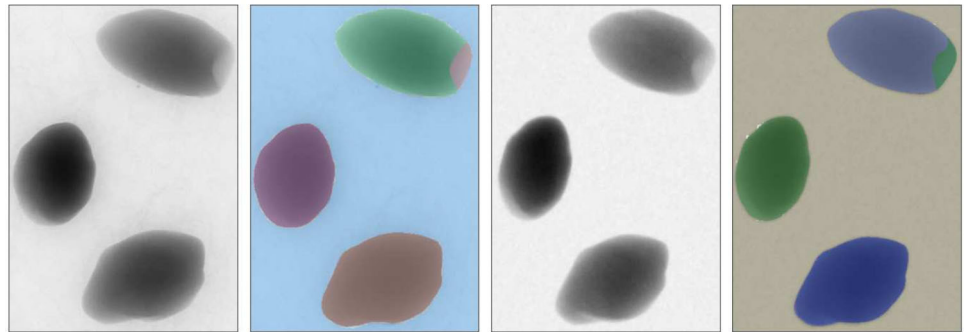


Fig. 6 2D projections and the corresponding segmentation obtained from SAM in noise-free simulations

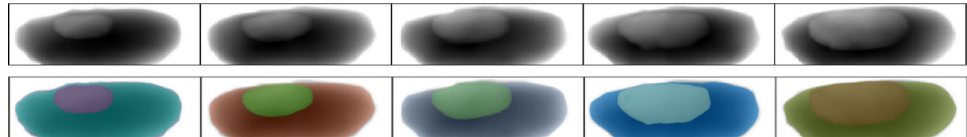
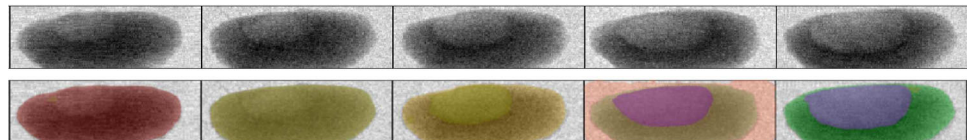


Fig. 7 2D projections and the corresponding segmentation obtained from SAM in noisy simulations



4.4 Porosity Estimation in Real CT Data

In the final stage, the 3D U-Net model was applied to segment pores in real CT scan data (see Fig. 16). Each rice grain was additionally analyzed morphologically for equivalent diameter, length, and volume (see Fig. 17). This analysis provided detailed insights into the pore structure and distribution within the rice grains. A detailed analysis of the internal porosity distribution presented in Fig. 18. Moreover, this information was also utilized to generate simulations of rice grains, such as modeling the size and length of the rice grains. The results demonstrate the model's capability to accurately identify and characterize porosity features in real-world data.

4.5 Comparative Discussion with Related Research

In this section, we compare and discuss the results in the context of related studies on chalkiness detection, segmentation, and the validation of these outcomes.

4.5.1 Chalkiness Detection Applications

When contrasting the approach presented here with the methodology of Wang et al. [7], which employs CNNs for detecting chalkiness in 2D photographic images, several key differences are apparent. In our study, we trained the models using 3D reconstructed CT data, enabling us to develop a deep learning model capable of detecting very

small instances of chalkiness. Specifically, the model identifies features as small as $109\ \mu\text{m}$ in height and 69 voxels within a $12\ \mu\text{m}$ voxel size, showcasing the advantage of 3D analysis over 2D photographic methods in detecting chalkiness or pores. However, it is important to note that this approach is currently limited to simulated images and has yet to be validated with real CT data for chalkiness detection.

Su et al. [23] presented both 2D and 3D visualizations of rice grains using Micro CT systems and quantified chalkiness at various scanning resolutions ranging from 2.5 to $15\ \mu\text{m}$. In comparison, this study utilized three resolution levels: 12 , 17 , and $17.48\ \mu\text{m}$. A significant difference lies in the segmentation algorithms; our study applied both supervised and unsupervised techniques, while Su et al. relied on traditional thresholding methods. Furthermore, Su et al. [23] did not include simulations or Probability of Detection (POD) analysis for different segmentation algorithms. Additionally, Su et al. examined rice grains with four types of chalkiness (white-belly, white-core, white-whole, and white-back), detecting chalkiness ranging from 0.21 to 0.499% , with an average of 0.39% . In contrast, this study assessed rice grain heights ranging from 0.6 to $0.8\ \text{mm}$ and defined chalkiness heights between 10 and $400\ \mu\text{m}$. The POD results indicate that the U-Net model achieved a minimum detection rate of 90% for chalkiness at $116\ \mu\text{m}$, with an approximate chalkiness rate of 13% . It is essential to acknowledge that, at this stage, the method has been validated only with simulated images due to the unavailability of real chalkiness data.

Fig. 8 POD model (red) for chalkiness height with confidence bounds of one hit/miss experiment and $a_{90/95}$ from the SAM model on **a** noise-free and **b** noisy simulations, respectively. Blue dots shows detected defects at probability one and undetected defects at probability zero (Color figure online)

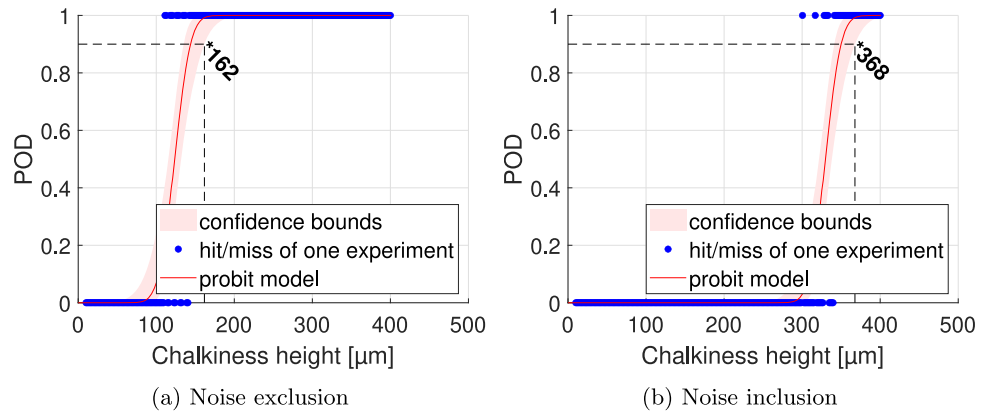


Fig. 9 Simulated slices images and the superimposed predictions on the ground truth (GT) including false positives (green), and false negatives (purple) with a 4× zoom images of the edge region highlighted in red (Color figure online)

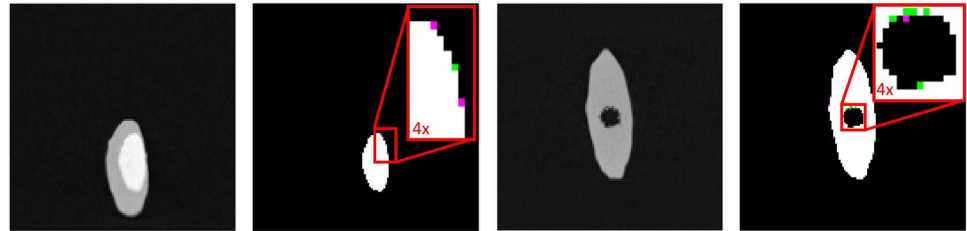


Fig. 10 POD model (red) for the number of voxels in the chalkiness part with confidence bounds of one hit/miss experiment and $a_{90/95}$. Blue dots shows detected defects at probability one and undetected defects at probability zero (Color figure online)

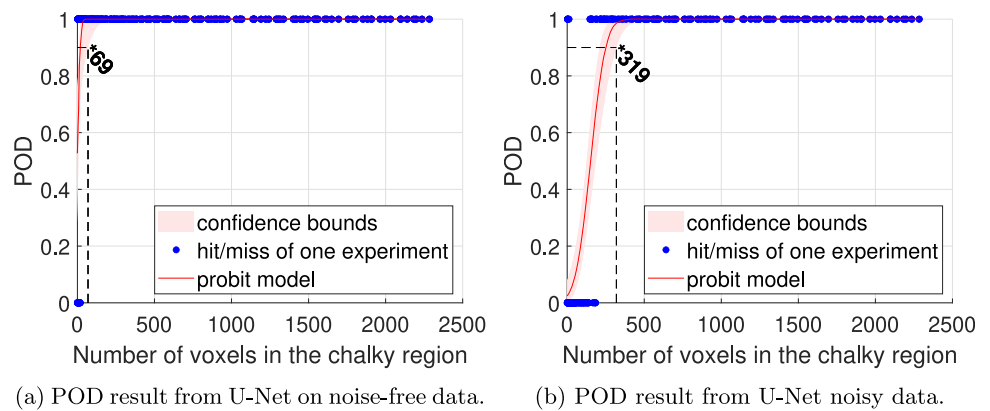


Fig. 11 POD model (red) for chalkiness height with confidence bounds of one hit/miss experiment and $a_{90/95}$. Blue dots shows detected defects at probability one and undetected defects at probability zero (Color figure online)

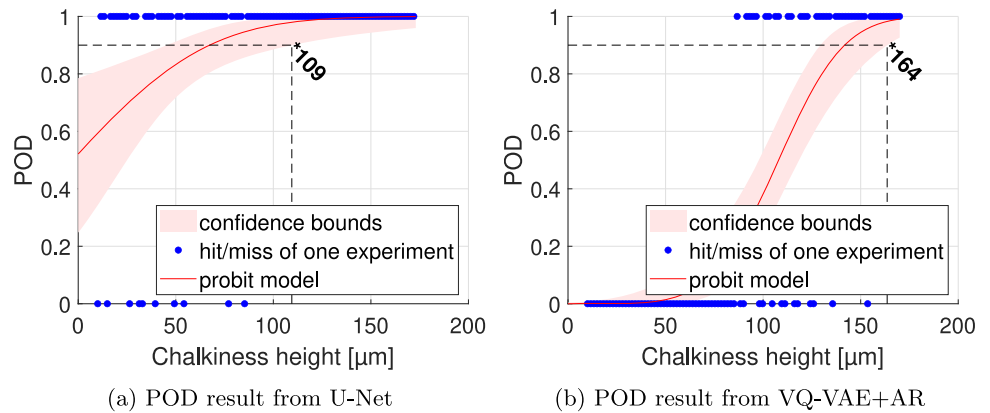
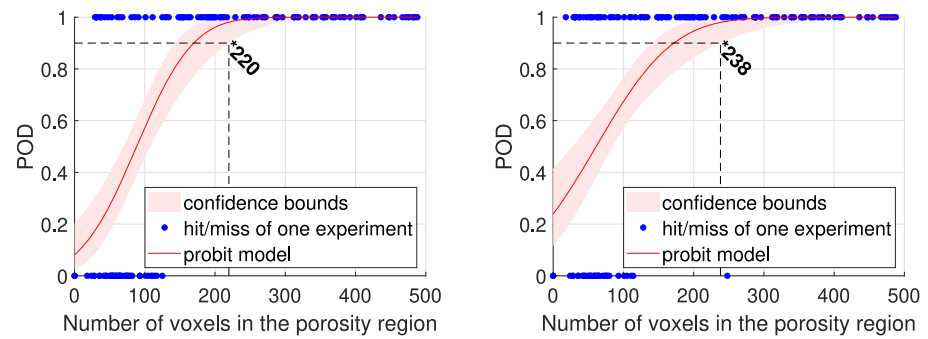
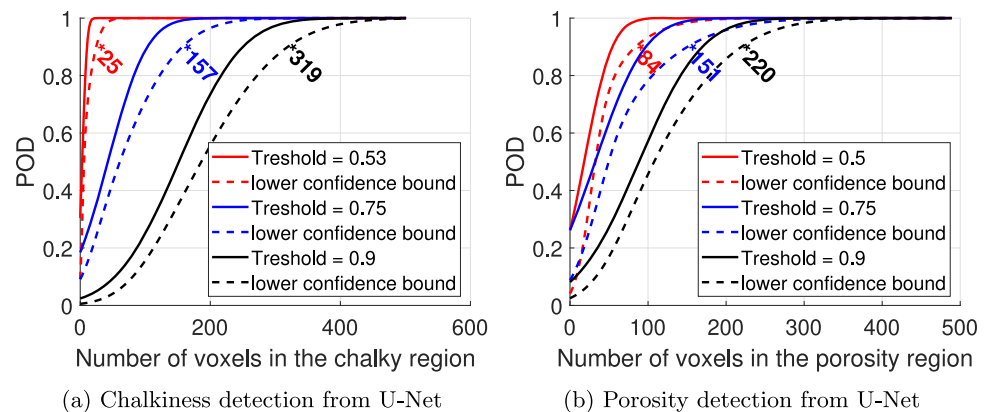


Fig. 12 POD model (red) for one hit/miss experiment considering the detection of pores in rice grains, the $a_{90/95}$ value is indicated. Blue dots shows detected defects at probability one and undetected defects at probability zero (Color figure online)



(a) POD results for pore detection from U-Net (b) POD results for pore detection from V-Net

Fig. 13 POD model comparison on different thresholds from Jacard index



(a) Chalkiness detection from U-Net

(b) Porosity detection from U-Net

4.5.2 Image Segmentation

In chalkiness detection applications, there are limited studies that have applied both supervised and unsupervised segmentation methods. However, in agricultural studies, the supervised method ResNet50 has been used for quality checking and defect detection in 'Braeburn' apple fruit, both in 2D and 3D, achieving a model accuracy of 96% [50]. Similarly, in this study, radiographies were generated using the ASTRA Toolbox [51] to train neural networks. In our application, U-Net achieved a accuracy (similarity index) of 0.9755 ± 0.08 for 3D segmentation. In comparison, V-Net and VQ-VAE+AR demonstrated lower performance than U-Net (see Table 2).

4.5.3 Validation

The neural networks model from U-Net, V-Net and VQ-VAE+AR were tested on 3D simulated chalkiness; however, due to the lack of available real rice grains with chalkiness, the model has not yet been evaluated on scans of actual chalky rice grains. When comparing our simulated chalkiness to the visualizations of rice chalkiness-both in terms of gray values and shapes-presented by Su et al. [23], the high visual similarity suggests that this method should be applicable to real data

once such samples become available. Additionally, the neural networks performed successfully on both for simulated porosity and scans of real rice grains that had air inclusion.

5 Conclusion

In this paper, we proposed a shape variation method for the generation of surface model variants in simulation-empowered training data generation. Target application area of this method is the analysis of defects in agricultural products, specifically chalkiness and pores in rice grains. We investigated detection efficiency on both supervised and unsupervised deep learning methods. Initially, we generated XCT scans of rice grains at three different resolutions. The resulting XCT datasets then were processed and converted into surface models, from which a bigger set of realistic simulated XCT datasets were generated. Finally, rigorous POD analyses were conducted with the intent of determining the limits of detectability of chalky parts or pores in XCT datasets of the aforementioned type. Our methodology demonstrates its applicability in conducting various types of detection of defects and degradation in rice samples. The SAM reached accurate results with POD determination in the 2D imaging case, especially on noise-free simulated data. However,

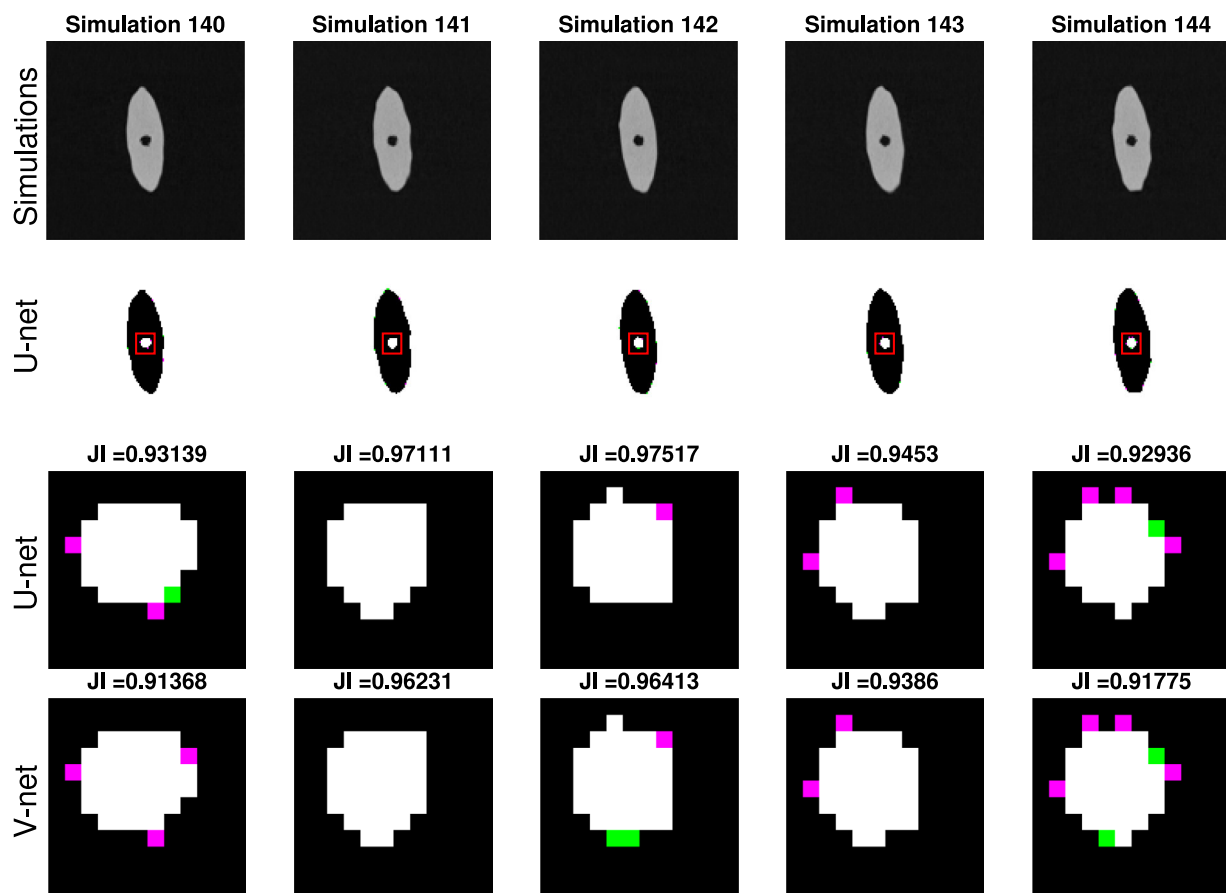


Fig. 14 Example results for air inclusion from 3D U-Net and V-Net segmentation. The figure shows, from the first row respectively: XCT simulations from SimCT, U-Net segmentations, U-Net segmentations (zoomed out), and V-Net segmentations (zoomed out in the region of interest)

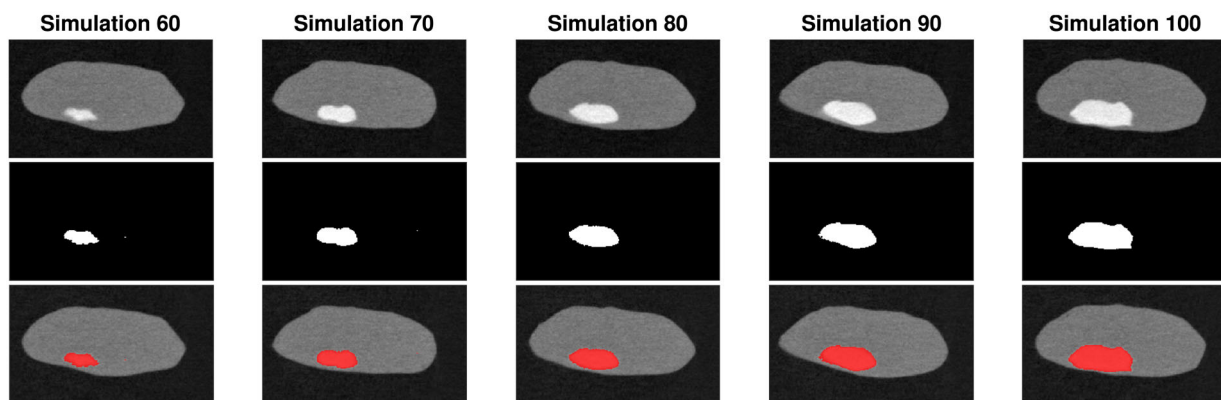


Fig. 15 Example XCT simulations and estimated segmentation slices from the 3D volumes using U-Net. The first row presents the simulated XCT data, followed by the segmentation masks in the second row, and finally, the segmentation overlays applied to the simulated XCT in the third row

Table 2 Segmentation quality metrics averages of the neural network from the testing data in chalkiness detection

Method	Jaccard index	Precision	Recall
U-Net	0.9755 ± 0.08	0.9930 ± 0.004	0.9721 ± 0.087
V-Net	0.9483 ± 0.01	0.9852 ± 0.011	0.9621 ± 0.015
VQ-VAE+AR	0.8695 ± 0.23	0.8992 ± 0.247	0.9620 ± 0.036

Fig. 16 Visualisation of individual rice grains from 3D rendering (a), the original slice extracted from the 3D reconstructed volume (b), and the estimations from 3D U-Net (c)

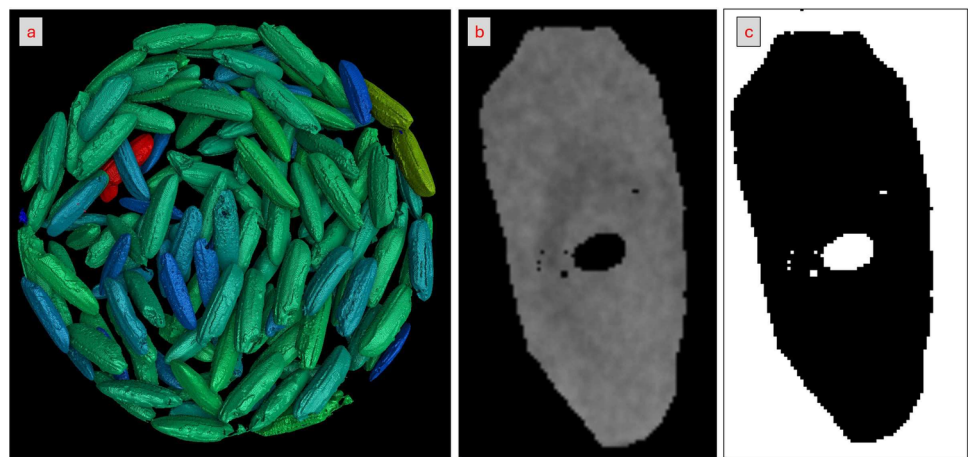


Fig. 17 Statistical analysis of individual rice grains from CT scan 3

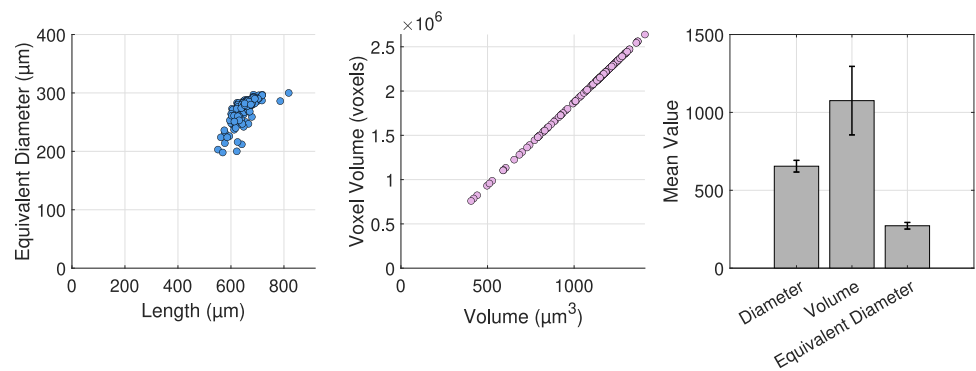
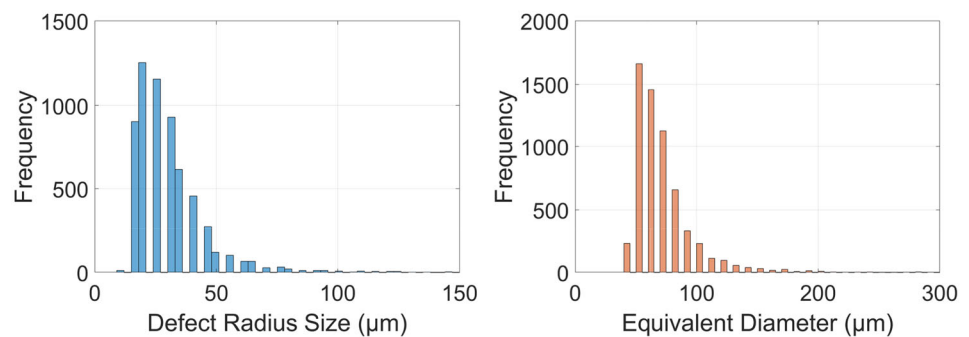


Fig. 18 Porosity distribution for defects radius and equivalent diameter



when noise and artifacts were introduced into the simulations, its performance decreased drastically. In a comparison of the detection performance between the used models on projected radiographs and XCT volumes, both for chalkiness and porosity, 3D CNNs methods outperformed the 2D SAM approach significantly. VQ-VAE+AR also showed promising results; however, it exhibited artifacts at the corners. The POD results indicated that both SAM and VQ-VAE+AR achieved comparable detection limits for chalkiness height.

In terms of future work, our approach shows the potential to be applicable on various other types of food products such

as seeds, fruits or plants. We therefore plan to test the presented method on the detection of rotten parts in both apples and zucchinis in our future work. Additionally, our method can be upgraded for multi-class segmentation and classification regarding the detection for degraded parts in 3D.

Acknowledgements We express our gratitude to Sarah Heupl for providing the CT scans utilized in our paper.

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. T. L. and C. H. provided support in draft preparation, conception, design, methodology and contributed to review and editing. V. F. and S. G. were involved in the generation of data and editing. J. D. B., J. S., and J. K. contributed to conception, design, and provided review and editing.

Funding Open access funding provided by University of Applied Sciences Upper Austria. 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.

Data Availability Data can be provided upon request.

Declarations

Conflict of interest The authors have no Conflict of interest to declare that are relevant to the content of this article.

Ethical approval Not applicable.

Consent for publication Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- UR, F.: Exporting U.S. Rice. <https://www.usarice.com/discover-us-rice/find-a-supplier/exporting-u.s.-rice>
- The Future of Food and Agriculture. Trends and Challenges. <http://www.fao.org/3/a-i6583e.pdf>
- Zhao, D., Zhang, C., Li, Q., Liu, Q.: Genetic control of grain appearance quality in rice. *Biotechnol. Adv.* **60**, 108014 (2022). <https://doi.org/10.1016/j.biotechadv.2022.108014>
- Shi, W., Yin, X., Struik, P.C., Solis, C., Xie, F., Schmidt, R.C., Huang, M., Zou, Y., Ye, C., Jagadish, S.V.K.: High day- and night-time temperatures affect grain growth dynamics in contrasting rice genotypes. *J. Exp. Bot.* **68**(18), 5233–5245 (2017). <https://doi.org/10.1093/jxb/erx344>
- Ashida, K., Iida, S., Yasui, T.: Morphological, physical, and chemical properties of grain and flour from chalky rice mutants. *Cereal Chem.* **86**(2), 225–231 (2009). <https://doi.org/10.1094/CCHEM-86-2-0225>
- Bahuguna, R.N., Solis, C.A., Shi, W., Jagadish, K.S.: Post-flowering night respiration and altered sink activity account for high night temperature-induced grain yield and quality loss in rice (**Oryza sativa** L.). *Physiol. Plantarum* **159**(1), 59–73 (2017). <https://doi.org/10.1111/ppl.12485>
- Wang, C., Caragea, D., Narayana, N.K., Hein, N.T., Bheemahalli, R., Somayanda, I.M., Jagadish, S.V.K.: Deep learning based high-throughput phenotyping of chalkiness in rice exposed to high night temperature. *Plant Methods* **18**(1), 9 (2022). <https://doi.org/10.1186/s13007-022-00839-5>
- 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>
- Fröhler, B., Weissenböck, J., Schiwarth, M., Kastner, J., Heinzl, C.: open_iA: A tool for processing and visual analysis of industrial computed tomography datasets. *J. Open Source Softw.* **4**(35), 1185 (2019). <https://doi.org/10.21105/joss.01185>
- Janssens, E., Alves Pereira, L.F., De Beenhouwer, J., Tsang, I.R., Van Dael, M., Verboven, P., Nicolai, B., Sijbers, J.: Fast inline inspection by neural network based filtered backprojection: application to apple inspection. *Case Stud. Nondestruct. Test. Eval.* **6**, 14–20 (2016). <https://doi.org/10.1016/j.csndt.2016.03.003>. Special Issue: Industrial computed tomography
- Van Dael, M., Verboven, P., Zanella, A., Sijbers, J., Nicolai, B.: Combination of shape and X-ray inspection for apple internal quality control. In silico analysis of the methodology based on X-ray computed tomography. *Postharvest Biol. Technol.* **148**, 218–227 (2019). <https://doi.org/10.1016/j.postharvbio.2018.05.020>
- Metzner, R., Eggert, A., Dusschoten, D., Pflugfelder, D., Gerth, S., Schurr, U., Uhlmann, N., Jahnke, S.: Direct comparison of MRI and X-ray CT technologies for 3D imaging of root systems in soil: potential and challenges for root trait quantification. *Plant Methods* **11**, 17 (2015). <https://doi.org/10.1186/s13007-015-0060-z>
- Gerth, S., Claußen, J., Eggert, A., Woerlein, N., Waininger, M., Wittenberg, T., Uhlmann, N.: Semiautomated 3D root segmentation and evaluation based on X-ray CT imagery. *Plant Phenomics* **2021**, 1–13 (2021). <https://doi.org/10.34133/2021/8747930>
- Li, Q., Jia, W., Sun, M., Hou, S., Zheng, Y.: A novel green apple segmentation algorithm based on ensemble U-net under complex orchard environment. *Comput. Electron. Agric.* **180**, 105900 (2021). <https://doi.org/10.1016/j.compag.2020.105900>
- Wood, R., Schut, D., Trull, A., Marcellis, L., Schouten, R.: Detecting internal browning in apple tissue as determined by a single CT slice in intact fruit. *Postharvest Biol. Technol.* **211**, 112802–1128029 (2024). <https://doi.org/10.1016/j.postharvbio.2024.112802>
- Su, Y., Xiao, L.-T.: 3D Visualization and volume-based quantification of rice chalkiness in vivo by using high resolution micro-CT. *Rice* **13**, 69 (2020). <https://doi.org/10.1186/s12284-020-00429-w>
- Berens, A.P.: NDE reliability data analysis. In: *ASM Handbook. Nondestructive Evaluation and Quality Control*, vol. 17, pp. 689–701 (1989)
- 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>
- U.S. Department Of Defense: *Handbook Non-Destructive Evaluation System Reliability Assessment*, vol. MIL-HDBK-, p. 1823 (1999)
- Yosifov, M., Reiter, M., Heupl, S., Gusenbauer, C., Fröhler, B., Gutiérrez, R.F., Beenhouwer, J.D., 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>
- Andriiashen, V., Liere, R., Leeuwen, T., Batenburg, K.J.: Ct-based data generation for foreign object detection on a single X-ray projection. *Sci. Rep.* **13**(1), 1881 (2023). <https://doi.org/10.1038/s41598-023-29079-w>
- Piao, R.-H., Chen, M.-J., Meng, F.-M., Qi, C.-Y., Koh, H.-J., Gao, M.-M., Song, A.-Q., Jin, Y.-M., Yan, Y.-F.: Identification and characterization of the chalkiness endosperm gene CHALK-H in rice (*Oryza sativa* L.). *J. Integr. Agric.* **22**(10), 2921–2933 (2023). <https://doi.org/10.1016/j.jia.2023.04.020>
- Su, Y., Xiao, L.-T.: 3D Visualization and volume-based quantification of rice chalkiness in vivo by using high resolution

- micro-CT. *Rice* **13**(1), 69 (2020). <https://doi.org/10.1186/s12284-020-00429-w>
24. Rice, I.N., Institute, I.R.R.: Standard Evaluation System for Rice (1996). <https://books.google.com.pk/books?id=xyBIAAAAYAAJ>
 25. Tam, B.P., Tu, P.T.B., Pha, N.T.: Identification of candidate SNPs and genes controlling chalkiness in the medium-grain rice (*Oryza sativa* L.). *Asian J. Plant Sci.* **22**, 618–627 (2023)
 26. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models—their training and application. *Comput. Vis. Image Underst.* **61**(1), 38–59 (1995). <https://doi.org/10.1006/cviu.1995.1004>
 27. Ambellan, F., Lamecker, H., Tycowicz, C., Zachow, S.: Statistical shape models: understanding and mastering variation in anatomy. *Adv. Exp. Med. Biol.* **1156**, 67–84 (2019). https://doi.org/10.1007/978-3-030-19385-0_5
 28. Zaitoun, N.M., Aqel, M.J.: Survey on image segmentation techniques. *Proc. Comput. Sci.* **65**, 797–806 (2015). <https://doi.org/10.1016/j.procs.2015.09.027>. International Conference on Communications, management, and Information technology (ICCMIT'2015)
 29. Sarma, R., Gupta, Y.K.: A comparative study of new and existing segmentation techniques. *IOP Conf. Ser. Mater. Sci. Eng.* **1022**(1), 12–27 (2021). <https://doi.org/10.1088/1757-899x/1022/1/012027>
 30. Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., Terzopoulos, D.: Image segmentation using deep learning: a survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **44**(7), 3523–3542 (2022). <https://doi.org/10.1109/TPAMI.2021.3059968>
 31. 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 Nondestruct. Test.* (2023). <https://doi.org/10.58286/27716>
 32. Chen, Y., Mancini, M., Zhu, X., Akata, Z.: Semi-supervised and unsupervised deep visual learning: a survey. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–23 (2022). <https://doi.org/10.1109/TPAMI.2022.3201576>
 33. Sarker, I.H.: Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN Comput. Sci.* **2**(6), 420 (2021). <https://doi.org/10.1007/s42979-021-00815-1>
 34. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention—MICCAI. Lecture Notes in Computer Science*, vol. 1, pp. 234–241 (2015). https://doi.org/10.1007/978-3-319-24574-4_28
 35. Yosifov, M.: Extraction and quantification of features in XCT datasets of fibre reinforced polymers using machine learning techniques. Master's thesis, Umeå University, Department of Computing Science (2020). <https://umu.diva-portal.org/smash/record.jsf?pid=diva2%3A1463165&dsid=4301>
 36. Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.-Y., Dollár, P., Girshick, R.: Segment anything. *arXiv:2304.02643* (2023)
 37. Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., Zagoruyko, S.: End-to-end object detection with transformers. *CoRR arXiv: 2005.12872* (2020)
 38. Cheng, B., Schwing, A., Kirillov, A.: Per-pixel classification is not all you need for semantic segmentation. In: Ranzato, M., Beygelzimer, A., Dauphin, Y., Liang, P.S., Vaughan, J.W. (eds.) *Advances in Neural Information Processing Systems*, vol. 34, pp. 17864–17875 (2021). https://proceedings.neurips.cc/paper_files/paper/2021/file/950a4152c2b4aa3ad78bdd6b366cc179-Paper.pdf
 39. Li, Y., Mao, H., Girshick, R., He, K.: Exploring plain vision transformer backbones for object detection (2022)
 40. Mazurowski, M.A., Dong, H., Gu, H., Yang, J., Konz, N., Zhang, Y.: Segment anything model for medical image analysis: an experimental study. *Med. Image Anal.* **89**, 102918 (2023). <https://doi.org/10.1016/j.media.2023.102918>
 41. Milletari, F., Navab, N., Ahmadi, S.-A.: V-Net: Fully convolutional neural networks for volumetric medical image segmentation. In: *International Conference on 3D Vision (3DV)*, pp. 565–571 (2016). <https://doi.org/10.1109/3DV.2016.79>
 42. Van Den Oord, A., Vinyals, O., Kavukcuoglu, K.: Neural discrete representation learning. *arXiv preprint arXiv:1711.00937* (2017)
 43. Reiter, M., Malik, M.M., Heinzl, C., Salaberger, D., Lettenbauer, E.G.H., Kastner, J.: Improvement of X-ray image acquisition using a GPU based 3DCT simulation tool. In: *International Conference on Quality Control by Artificial Vision* (2009)
 44. Bhattacharya, K.R., Sowbhagya, C.M., Swamy, Y.M.I.: Some physical properties of paddy and rice and their interrelations. *J. Sci. Food Agric.* **23**(2), 171–186 (1972). <https://doi.org/10.1002/jsfa.2740230204>
 45. Ray, S., Turi, R.H.: Determination of number of clusters in k-means clustering and application in colour image segmentation. In: *International Conference on Advances in Pattern Recognition and Digital Techniques*, pp. 137–143 (1999)
 46. He, K., Chen, X., Xie, S., Li, Y., Dollár, P., Girshick, R.: Masked autoencoders are scalable vision learners. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 16000–16009 (2022)
 47. Oord, A., Kalchbrenner, N., Espeholt, L., Kavukcuoglu, K., Vinyals, O., Graves, A.: Conditional image generation with Pixel-CNN decoders. In: Lee, D., Sugiyama, M., Luxburg, U., Guyon, I., Garnett, R. (eds.) *Advances in Neural Information Processing Systems*, vol. 29 (2016). https://proceedings.neurips.cc/paper_files/paper/2016/file/b1301141feffabac455e1f90a7de2054-Paper.pdf
 48. Sorkine, O., Cohen-Or, D., Lipman, Y., Alexa, M., Rössl, C., Seidel, H.-P.: Laplacian surface editing. In: *SGP'04*, pp. 175–184. Association for Computing Machinery, New York (2004). <https://doi.org/10.1145/1057432.1057456>
 49. Spiess, M., Nagl, W., Hamerle, A.: Probit models: regression parameter estimation using the ML principle despite misspecification of the correlation structure (1997)
 50. Tempelaere, A., Van Doorselaer, L., He, J., Verboven, P., Nicolai, B.M.: Braenet: internal disorder detection in 'braeburn' apple using X-ray imaging data. *Food Control* **155**, 110092 (2024). <https://doi.org/10.1016/j.foodcont.2023.110092>
 51. van Aarle, W., Palenstijn, W.J., De Beenhouwer, J., Altantzis, T., Bals, S., Batenburg, K.J., Sijbers, J.: The astra toolbox: a platform for advanced algorithm development in electron tomography. *Ultramicroscopy* **157**, 35–47 (2015). <https://doi.org/10.1016/j.ultramicro.2015.05.002>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.