Automatic segmentation for synchrotron-based imaging of porous bread dough using deep learning approach

Salah Ali,[a]* Sherry Mayo,[b] Amirali K. Gostar,[a] Ruwan Tennakoon,[c] Alireza Bab-Hadiashar,[a] Thu MCann,[d] Helen Tuhumury[d] and Jenny Favaro[d]

[a]School of Engineering, RMIT University, Australia, [b]CSIRO Manufacturing, Clayton, Victoria, Australia, [c]School of Science, RMIT University, Australia, and [d]CSIRO Agriculture and Food, Werribee, Victoria, Australia. *Correspondence e-mail: s3629495@student.rmit.edu.au

In recent years, major capability improvements at synchrotron beamlines have given researchers the ability to capture more complex structures at a higher resolution within a very short time. This opens up the possibility of studying dynamic processes and observing resulting structural changes over time. However, such studies can create a huge quantity of 3D image data, which presents a major challenge for segmentation and analysis. Here tomography experiments at the Australian synchrotron source are examined, which were used to study bread dough formulations during rising and baking, resulting in over 460 individual 3D datasets. The current pipeline for segmentation and analysis involves semi-automated methods using commercial software that require a large amount of user input. This paper focuses on exploring machine learning methods to automate this process. The main challenge to be faced is in generating adequate training datasets to train the machine learning model. Creating training data by manually segmenting real images is very labour-intensive, so instead methods of automatically creating synthetic training datasets which have the same attributes of the original images have been tested. The generated synthetic images are used to train a U-Net model, which is then used to segment the original bread dough images. The trained U-Net outperformed the previously used segmentation techniques while taking less manual effort. This automated model for data segmentation would alleviate the time-consuming aspects of experimental workflow and would open the door to perform 4D characterization experiments with smaller time steps.

1. Introduction

Bread is a spongy, porous material consisting of open and closed pores. The open pores comprise the majority of its texture, with a single interconnected pore-network accounting for about 99% of the total porosity. Breads produced worldwide come in different appearances and features. For example, French baguette is highly aerated crumb, the Middle East flatbread features a dense and crispy crust, and the Chinese steamed bread is usually soft with a moist skin (Gao et al., 2018).

There has been great interest in studying the micro-structure of bread made from Australian wheat. Australia produces around 17.3 million tons of wheat per year, the majority of which is used in noodles, pasta and stemmed and flat breads. There are notable differences compared with other types of wheat worldwide. Australian wheat flour typically has lower protein content than North American flour, so it is not generally considered as suitable for commercial leavened
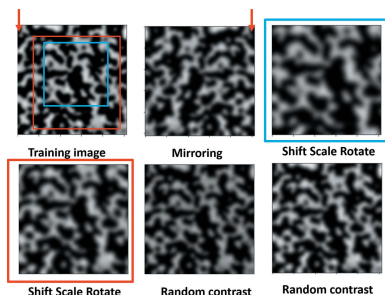

Training image   Mirroring   Shift Scale Rotate

Shift Scale Rotate   Random contrast   Random contrast

Table 1
List of imaging sources and segmentation techniques used in existing studies on bread.

	Wheat type	Imaging technique	Reference	Year	Segmentation method
1	Doughs mixed at different pressures	FEI Quanta 200 environmental scanning electron microscope	(Trinh *et al.*, 2013)	2013	*Avizo*, thresholding
2	Canadian and Australian	SkyScan 1176 high resolution	(Chakrabarti-Bell *et al.*, 2014; Wang, 2015; Wang *et al.*, 2013)	2014	Automatic thresholding
3	Bread (local bakery)	Nikon metrology 160 Xi Gun set X-ray source	(Dyck *et al.*, 2014)	2014	Otsu's method
4	Bread doughs	European Synchrotron Radiation Facility	(Turbin-Orger *et al.*, 2015)	2015	Automatic thresholding
5	Cleaned and sifted samples of Canadian wheat	SkyScan 1172 micro-computed tomography scanner	(Suresh & Neethirajan, 2015)	2015	Multi-thresholding techniques
6	Australian and North American bread dough	Australian Synchrotron	(Mayo *et al.*, 2016)	2016	*Avizo*
7	German commercial	Eclipse Ti-U inverted microscope	(Bernklau *et al.*, 2016)	2017	*AngioTool64*
8	Bread doughs noodles	Canadian Light Source	(Koksel *et al.*, 2016, 2017*a*,*b*; Koksel & Scanlon, 2016)	2016–2018	Thresholding after histogram equalization

bread-making (Park *et al.*, 2017). In recent years researchers have sought to understand the fundamentals of the performance of Australian wheat flour in bread to improve its performance with additives or new processing methods (Chakrabarti-Bell *et al.*, 2014).

Investigating the mechanical properties for bread samples requires accurate quantification of the micro-structural parameters such as cell wall thickness, cell shape, void fraction, crumb brightness, and fineness. To achieve this, many techniques have been used (Falcone *et al.*, 2006). Computer-aided X-ray micro-tomography (micro-CT) is a recently popularized method to investigate the internal structure of various materials non-destructively. Many studies have utilized micro-CT to study the bread properties, taking advantage of the high contrast between pore and void phases in the bread (Babin *et al.*, 2005).

A growing trend is the increased use of fast (or ultra-fast) micro-CT at third-generation synchrotron facilities allowing to obtain larger sized and higher resolution data sets. This makes it possible to follow the structural changes of bread over time (Gao *et al.*, 2018; Koksel *et al.*, 2016). Analysing the bread using the micro-CT technique has been employed to classify different types of bread based on quantitatively measured parameters (Cafarelli *et al.*, 2014*a*), and to investigate the correlation between the bread bubble size and its shape using advanced statistical methods (Koksel *et al.*, 2016). Different types of bread have been characterized using micro-CT studies ranging from bread varying in yeast and water content, Canadian wheat with damage caused by insects and sprouting and different types of Italian breads (Cafarelli *et al.*, 2014*b*; Suresh & Neethirajan, 2015).

Tomography experiments using synchrotron radiation can produce a massive amount of 3D data in just a few days. This presents a major challenge for quantitative data analysis using traditional methods. Similar challenges are emerging with increasingly efficient laboratory CT systems. This paper focuses on the development of an automated workflow for accurate image segmentation. The result may lead to more accurate quantitative analysis of sequences of 3D data-sets of samples evolving over time and/or in response to changing experimental conditions, and with less need for user input.

Researchers have used different workflows for analysing the 3D representation of the sample generated by the micro-CT reconstruction software. However, there are common steps shared among the majority of these studies, including pre-processing, segmentation, labelling, post-processing, and quantification of measurements. Researchers typically have used different software modules to perform the different steps through a point-and-click approach (Jensen *et al.*, 2014).

There have been some efforts to automate some of the manual processes that need to be done by an operator. The back-end programming interface in Tool Command Language (TCL) and Java available in *Avizo* and *ImageJ* has been utilized to automate the point-and-click approach performed by the user. However, the automated process still accounts for a very small portion of the overall workflow and its accuracy fails to achieve satisfactory output compared with the manual workflow (Jensen *et al.*, 2014). Table 1 lists a summary of imaging and segmentation techniques used in existing studies on bread.

Machine learning techniques have been used in analysing tomography datasets used in many material science applications. A multi-layer perceptron architecture is used to detect the crack in the tomographic dataset of lithium-ion cells (Petrich *et al.*, 2017). A supervised feature-based technique is used to segment mineral phases of amphibole, plagioclase and sulfide phases (Guntoro *et al.*, 2019). A combination of Trainable Weka Segmentation (TWS) and the level set method is used to segment and reconstruct tomography images of particles of granular geomaterials (Lai & Chen, 2019).

Recently, with the re-emergence of neural networks, researchers have started to use deep neural networks as a powerful tool to analyse CT-images primarily for generic microstructure representation, classification and segmentation. PixelNet architecture has been used to segment ultrahigh carbon steel trained on 24 images (DeCost *et al.*, 2019). Other models are summarized in Table 2. All these models depend

Table 2
List of deep learning models used to segment porous materials.

Sample description	Scanning technology	Technique	Ground truth	Number of mages	Reference
Four phases, ultrahigh carbon steel	UHCS dataset	PixelNet	Partially automated	24	(DeCost *et al.*, 2019)
Two phase steel	LOM and SEM	MVFCNN	Manually by group of material experts, and metallographers	21	(Smal *et al.*, 2018)
Fontainebleau sandstone and Grosmont carbonate	Public dataset	SegNet	Manual segmentation	20	(Koksel *et al.*, 2017*b*)

critically on the accuracy of the ground truth images used for training which are time-consuming to prepare, typically relying on manual or only partially automated segmentation.

The U-Net model was initially developed for medical image segmentation at the University of Freiburg, Germany (Ronneberger *et al.*, 2015). Here we have explored the potential of U-Net for automated segmentation of porous structures.

The main contribution of this paper is utilizing U-Net Convolutional Neural Networks (CNNs) (DeCost *et al.*, 2019) to segment micro-CT images of bread doughs using synthetic images to create a training dataset, and thus eliminating any need for manually segmenting the ground truth data. We have investigated the potential for this technique to facilitate quantitative microstructure analyses that conventionally would require a large amount of hands-on image processing. The proposed approach enables the researchers to focus on obtaining insights from their data rather than focusing on manually segmenting and inspecting their data. This in turn would allow researchers to perform further experiments for bread doughs or any other porous material studies with smaller time steps.

The remainder of this paper is organized as follows. First, in Section 2, the dataset of bread dough and the process of generating the synthetic data are explained. It also describes the segmentation models, and how we train the model. The results are discussed in Section 3. Finally, conclusions are given in Section 4.

2. Experimental

2.1. Acquisition of the dataset

Images were acquired at the Australian Synchrotron to study the effect of different salt additives on the performance of dough made from high-protein flours (13.6% protein) and low-protein flour (9.6% protein). For each of the two types of flour (high and low protein) four formulations were made using no salt, NaCl, KCl and NaBr, respectively. Two samples of each formulation were prepared making 16 samples in total as shown in Table 3. The micro-CT images were scanned with the following parameters: monochromatic X-ray beam, energy = 30 keV, 720 views, 0.5° step, 11.13 µm pixel size, sample-to-detector distance of 80 cm, 16 s exposure time, pixels containing the bread doughs $1024 \times 1024 \times 600$. The full details of preparing these samples are described by Mayo

Table 3
Recipes for bread making and sample lists.

Sample number	Protein content	Quantity	Salt type	Water target
1	Low protein (9.6% protein)	4 g	KCl	2.18 ml
2	High protein (13.6% protein)	4 g	KCl	2.18 ml
3	Low protein (9.6% protein)	4 g	NaBr	2.24 ml
4	High protein (13.6% protein)	4 g	NaBr	2.24 ml
5	Low protein (9.6% protein)	4 g	No Salt	2.18 ml
6	High protein (13.6% protein)	4 g	No Salt	2.18 ml
7	Low protein (9.6% protein)	4 g	NaCl	2.24 ml
8	High protein (13.6% protein)	4 g	NaCl	2.24 ml

et al. (2016). The sample-to-detector distance was chosen to be sufficient that the images would benefit from propagation-based phase-contrast (with phase-contrast fringes of the order of the pixel size) in order to improve the visibility of voids, without resulting in excessive blurring due to Fresnel diffraction.

The scans were performed at interval of about 5 min during both proving and baking. This was followed by data reconstruction using the *X-TRACT* software (Gureyev *et al.*, 2011). This includes standard image corrections, phase retrieval using Paganin's algorithm (Paganin *et al.*, 2002) and filtered back-projection algorithm tomographic reconstruction. Phase-retrieval is used because the data were collected in propagation-based phase-contrast mode and phase-retrieval improves signal-to-noise in the resulting images and results in reconstructed images more suitable to image segmentation. The delta-to-beta ratio used for reconstruction was tuned to remove phase-contrast fringes and improve signal-to-noise without introducing additional blurring (Paganin *et al.*, 2002). The final output for all the samples was a total of 464 3D volumes, representing scans at 29 time-points for each of the 16 dough samples.

2.2. Preparation of testing datasets

In order to validate and compare our proposed segmentation methods with other segmentation methods, we need a testing dataset based on our real bread CT data images (as input) and corresponding accurately segmented 'ground truth' images to compare with the output of the different segmentation methods. Although the original image size is 1024×1024 pixels, as we use the U-Net model (and initialize its weights using ImageNet dataset, which has a smaller input size), input images are down-sampled to fit the U-Net network
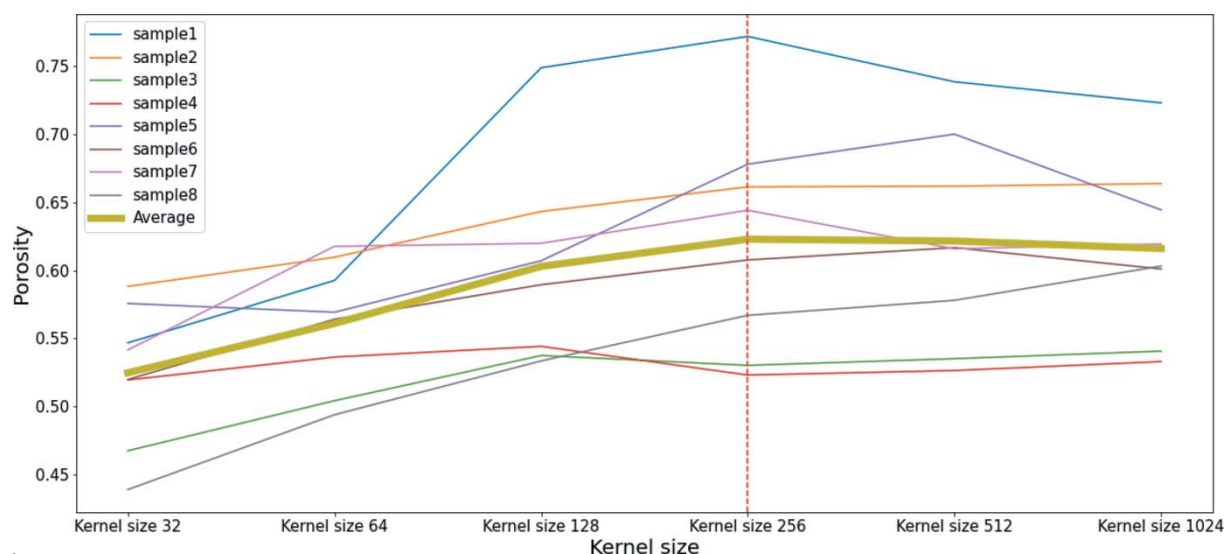
Figure 1
Graph showing the porosity of the samples. Sample 1 is low protein with no salt, sample 2 is high protein with no salt, sample 3 is low protein with NaCl salt, sample 4 is high protein with NaCl salt, sample 5 is low protein with NaBr, sample 6 is high protein with NaBr, sample 7 is low protein with KCl salt, sample 8 is high protein with KCl salt.

(without introducing interpolation errors). To show that the downsampling process does not affect the overall porosity, 32 images of different sizes (32×32, 64×64, 256×256, 512×512, and full size 1024×1024) were selected for porosity comparison. This analysis is conducted slice-wise, where each image is a single x–y slice. As shown in Figure 1, the porosity for images at or above 256×256 does not change significantly and the average porosity for those sizes remains relatively constant. As such, we chose to use images at 256×256 pixels in this work. Given that the inference time for segmentation is less than 30 ms for images of this size, the method could be used for larger datasets as well.

Figure 2 shows a grayscale sample image of a bread dough slice and histogram of grayscale pixel intensities. The histogram of grayscale images in some cases has two peaks, where one represents the pore and the other represents the bread. However, in many cases theses peaks are not distinguishable. Therefore, the current solutions for automated segmentation does not return accurate results.

From each bread sample, three 2D images at different time steps were selected and annotated to create the ground-truth dataset for testing step. As shown in Figure 3, the testing

images have different pore size and intensity brightness, which provides a baseline evaluation for the trained network.

Creating the ground truth images involves three steps. Creating the accurately segmented image from the original image involved using a 2D Otsu thresholding technique followed by the use of a manual annotation tool to fix the mislabeled pixels as an expert supervisor. An open source software tool (*VIA*) is used to select the over-segmented and
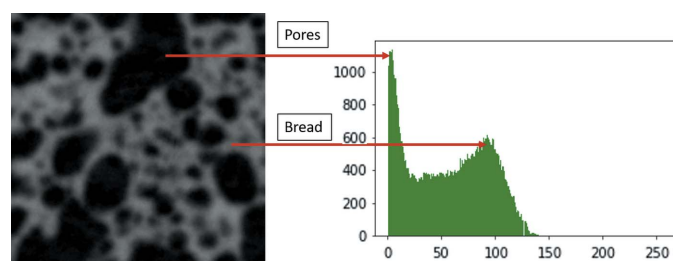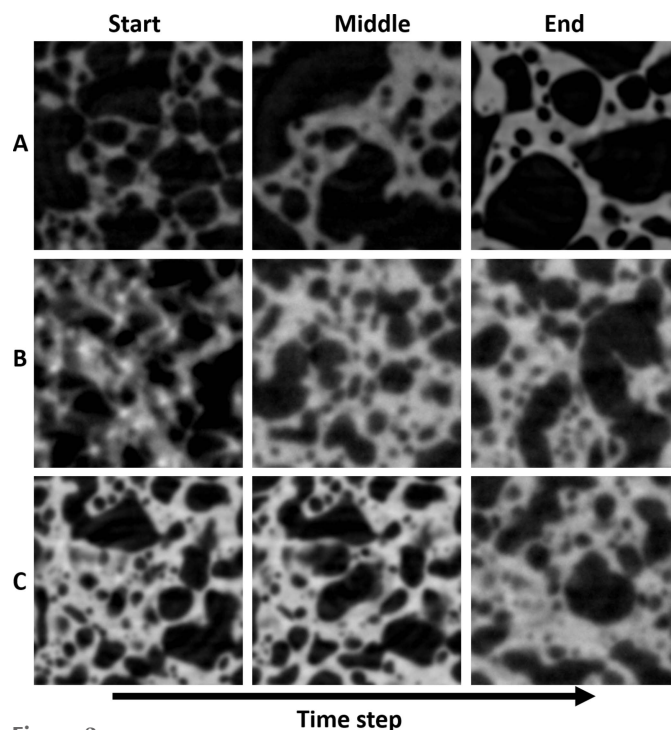

Figure 3
Testing images. (A) Image for salt NaCl–high protein sample, (B) image for salt KCl–low protein sample, and (C) image for salt NaBr–low protein sample.

Figure 2
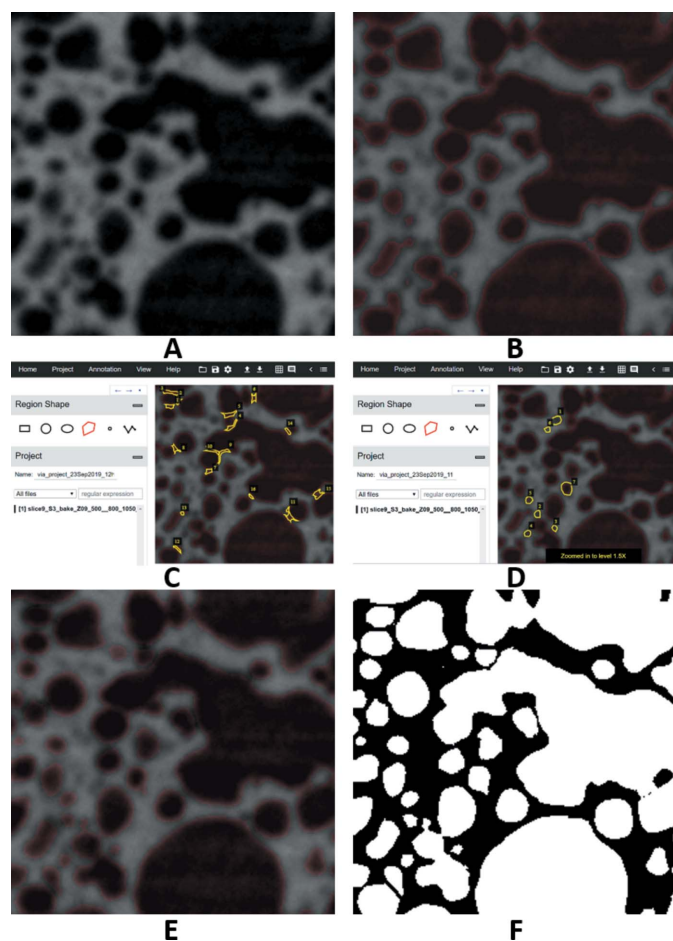Left: grayscale image of bread dough. Right: histogram of grayscale pixel intensities.

Figure 4
Steps for manual annotating of grayscale image. (A) Original image, (B) automatic segmentation output, (C) removing over-segmented areas, (D) adding under-segmented areas, (E) overlay of original image with the fixed binary image, and (F) the final ground truth image.

under-segmented regions (Dutta *et al.*, 2016). Lastly, the output of the annotator tool is a JSON file for polygon points of the boundaries, which are converted into 2D masks and then merged with the initial segmented images. Figure 4 shows the steps followed to create the ground truth images.

By plotting the histogram of the pores and bread pixels for the ground truth images, it is clear that the pixel intensities of the pores and breads overlap. Figure 5 shows that for all testing images there is an overlapping area in the pixel intensities. The automatic threshold would fail to separate the two classes accurately.

2.3. Generating synthetic training data

X-ray CT images should in principle have pixel intensities proportional to the X-ray absorption coefficient in each pixel. In reality, such images differ from the corresponding ideal 'ground truth' image for a number of reasons. These include pixels with intermediate intensities because they are on the boundary of a pore due to the partial volume effect. Other reasons are the image noise (primarily photon shot noise) and blurring due to the point-spread function of the imaging
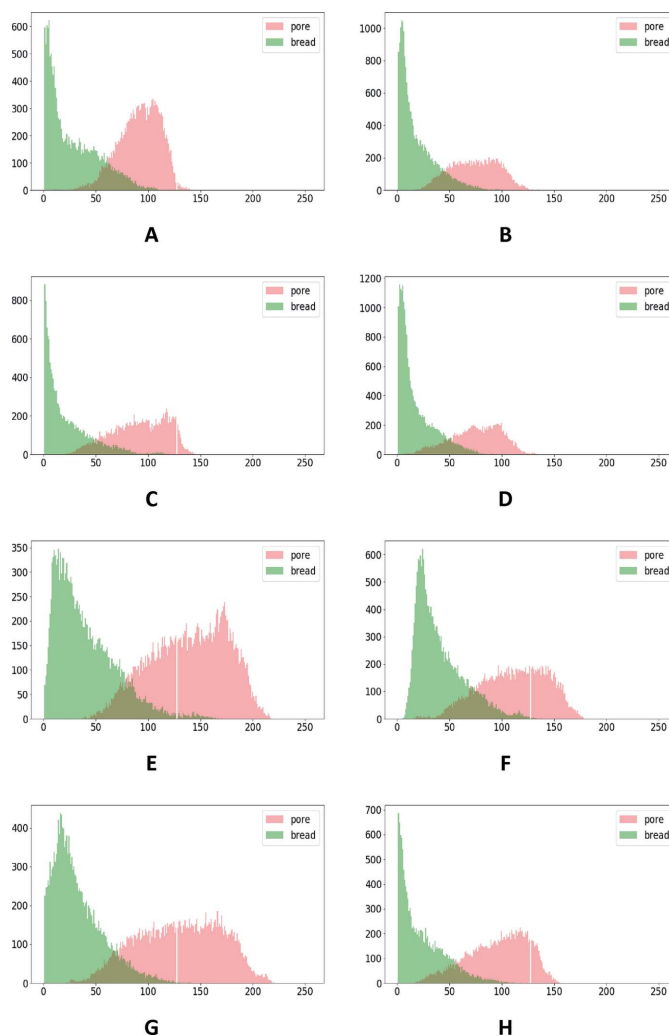

Figure 5
Histogram of annotated images. (A) Low protein with no salt, (B) high protein with no salt, (C) low protein with NaCl salt, (D) high protein with NaCl salt, (E) low protein with NaBr, (F) high protein with NaBr, (G) low protein with KCl salt, and (H) high protein with KCl salt.

system. These issues make it more time consuming to generate sufficient quantities of accurate ground-truth data by accurately segmenting real data, for training purposes. To alleviate this limitation, we developed a method to generate and use synthetic data for model training step.

In order to generate synthetic data, we have used the open source library (PoresPy) to generate 3D synthetic binary ground truth images from overlapping spheres or blobs with different porosity values and shapes, and then they are split into 2D slices (Keilegavlen *et al.*, 2019). First we scale the brightness of the generated 2D image by shifting the binary image values [0,255] to the peaks of the original grayscale image. Then a Gaussian noise is added to the 2D binary image as shown in equation (1),

$$I_{\text{noise}}(i, j) = I_{\text{binary}}(i, j) + N(i, j). \qquad (1)$$

Finally, I_{noise} is convoluted with a Gaussian blur kernel followed by a sharpening function to create good contrast
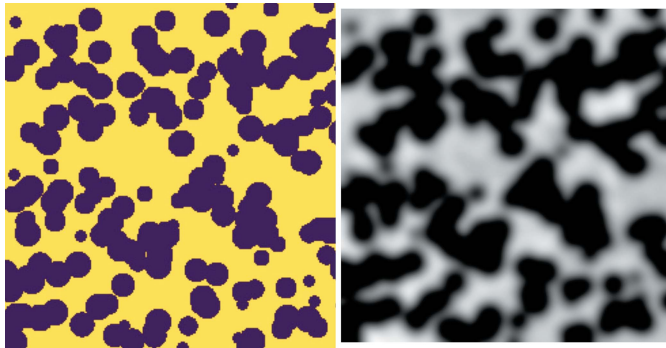
Figure 6
Synthetic images generated using the PoresPy library. Left: binary image. Right: grayscale synthetic image.

images similar to the original images. Figure 6 shows 2D slice of a synthetic image generated using this method.

The synthetic training data need to represent the characteristics and structure of the dataset. We demonstrate that they are not required to be closely matching the testing dataset. As shown in Figure 7, the synthetic images are generated from the combination of shapes and blobs consisting of images with porosity ranges from 0.4 to 0.6, and different blob sizes to generate images with varying sizes of the pore. These values are selected as they are in the range of real bread porosities. As the bread CT images are taken during
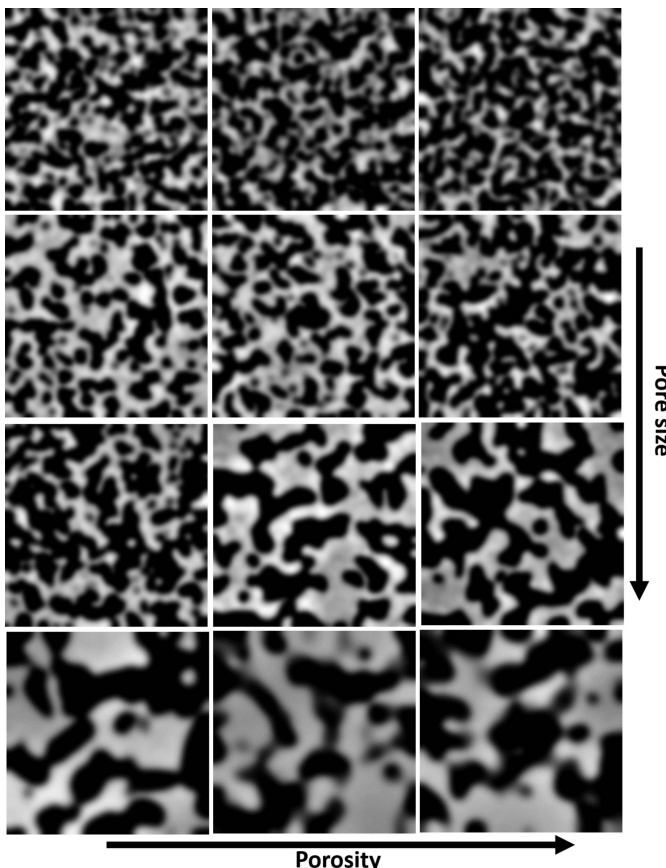

Figure 7
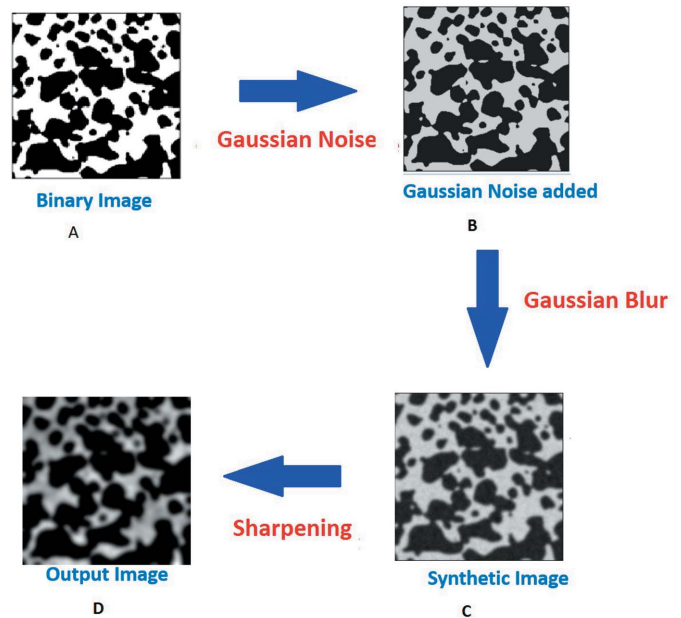Synthetic images with different shapes and appearances.

Figure 8
Steps to generate synthetic training images. (A) Input binary image, (B) imaged with added noise, (C) image with blurring effect, and (D) the output image.

rising and baking, and the pores tend to change over time, the synthetic images are generated from overlapping spheres at different radius sizes to simulate change of pore sizes of the real images, and to ensure the generated images are not repeated and have different appearances.

In this way, we can create many images with different values of noise, shapes and porosities for training the segmentation models and verifying the accuracy of the model without any need to create more ground truth images manually. The outline of the proposed method is shown in Figure 8.

2.4. Training the U-Net deep neural network architecture

Deep learning is a subset of machine learning, which learns the pattern from the data. It learns the hidden features of the data and does not require the training dataset to close match the predicted dataset (LeCun *et al.*, 2015). Our work employs the U-Net model as it is commonly used for medical image segmentation. The U-Net architecture is illustrated schematically in Figure 9. This architecture consists of three sections: the contraction, the bottleneck and the expansion section. The contraction section is initialized with the pre-trained models used for classification such as ResNet and VGG models (Yakubovskiy, 2019). That usually gives a better result than initializing the weights randomly. This part is a down-sampling path and it enables extracting coarse contextual information. The bottleneck layer is the intermediates between the contraction and the expansion layers. The last section is the expansion, which is called an up-sampling path. It enables extracting a precise localization of the features, with the contracting path features using skip connections. The output of the U-Net is a prediction for the class of each pixel of the input image (*e.g.* pore versus dough in the case of our data).
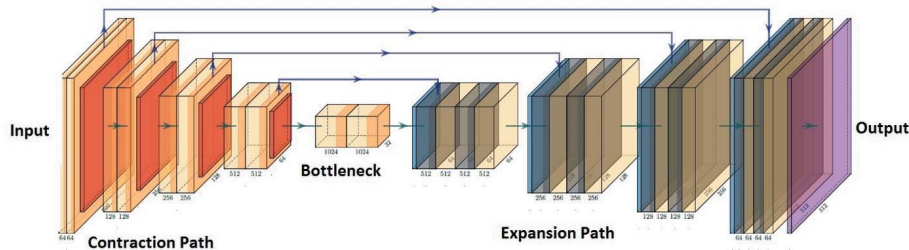
Figure 9
The U-Net architecture (Ronneberger *et al.*, 2015).

The model was trained on the CSIRO High Performance Computing system. The resources allocated were a K40 NVIDIA GPU, 40 GB memory. The inference time is measured to be around 29.5 ms, which enables it to be used for the segmentation of large datasets. The model was trained using 1400 synthetic 2D images generated from different 3D images, and different augmentation methods to enhance the ability of the model to generalize its learning. As shown in Figure 10, the augmentation methods add more variance to the training images, which removes the need for generating closely similar images to the testing images.

3. Results and discussion

3.1. Comparison of U-Net and Otsu segmentation

The test images selected from the bread CT scans were segmented into pore and dough using the trained U-Net model and the conventional Otsu threshold technique for comparison. Different types of denoising filters such as bilateral filter, non-local means and Gaussian filter were applied to the images before using the Otsu threshold to improve its accuracy. The best accuracy was achieved with the Gaussian filter, which was applied for testing images as a preprocessing step for both the U-Net model and Otsu threshold.
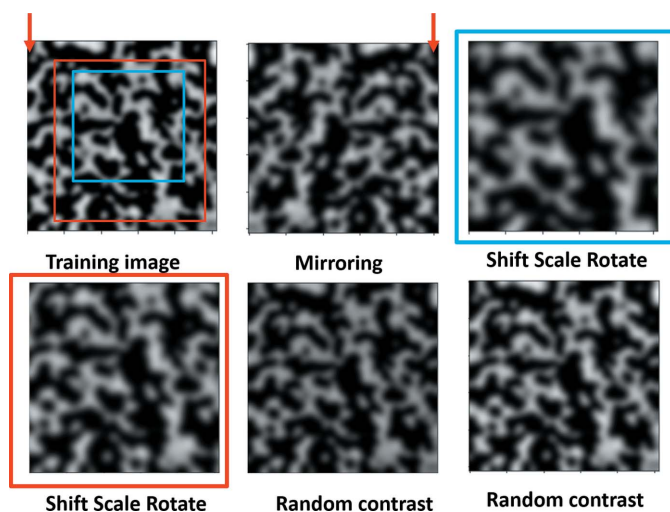
The accuracy of each segmentation is calculated in terms of pixel accuracy (PA) defined as follows,

$$PA = \frac{TP + TN}{TP + TN + FP + FN}, \quad (2)$$

where TP stands for true positive, TN stands for true negative, FP stands for false positive, and FN stands for false negative. PA can be calculated separately for pore, dough and the overall segmentation. Table 4 shows the pore, dough and overall PA of the test images using for both methods. Best results are shown in bold. The relative performance of each method is summarized in Figure 11.

As we can see in Table 4, the accuracy of the U-Net model for each phase (pore and bread) outperformed the Otsu thresholding (TH) for almost every measurement. Most importantly, it is able to pick up the small pores, as shown in Figure 12, which are very critical for the quantification analysis that is usually performed after extracting the binary image. Therefore, this model not only outperforms the automatic thresholding in terms of the pixel accuracy but it also picks up the difficult regions, which are missed by most of the automatic segmentation techniques. Therefore, using this model will give a very accurate binary image, which is the most challenging task for quantification analysis.

Figure 12 shows a comparison of overlaying regions of the proposed technique and Otsu thresholding segmentation. In panels D and G of Fig. 12, the yellow labels are regions for which the method successfully predicts the output (true positives), while the red labels are regions for which the method fails to predict the pores pixels (false positives). The green labels are regions for which the method fails to predict bread pixels (true negatives). It is clear that the proposed method has fewer green and red regions, which indicates that it outperforms the Otsu threshold method. The overall accuracy is shown in panels E and H, where the red regions indicate the mislabelling of the prediction compared with the

Figure 10
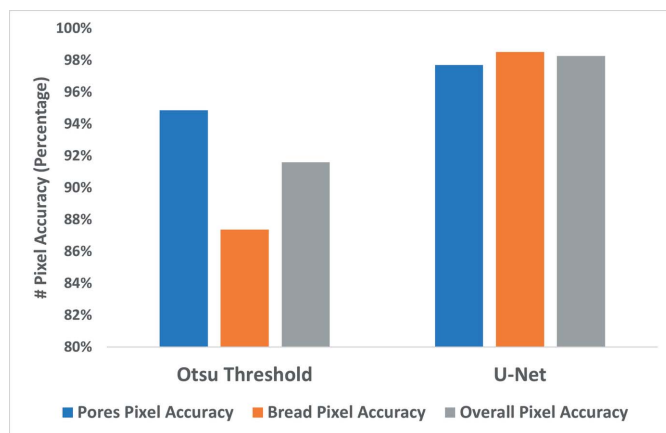Different augmentation methods applied on training images.

Figure 11
Comparison of accuracy for the eight samples of bread using the U-Net model trained on synthetics images generated from the Porespy library and automatic Otsu TH.

Table 4
Comparison of accuracy for eight samples of bread using the U-Net model trained on synthetic images generated from the Porespy library and Otsu TH.

| Sample name	Pore PA		Bread PA		Overall PA	
	Otsu TH	U-Net	Otsu TH	U-Net	Otsu TH	U-Net
No salt–low protein	90.80%	**96.45%**	92.91	**98.65%**	91.84%	**97.62%**
No salt–high protein	96.88%	**99.49%**	87.76%	**97.67%**	94.05%	**99.05%**
Salt NaCl–low protein	96.21%	**99.20%**	89.24%	**98.42%**	93.37%	**98.94%**
Salt NaCl–high protein	96.70%	**99.30%**	89.35%	**97.91%**	94.08%	**98.81%**
Salt NaBr–low protein	93.62%	**97.13%**	86.89%	**99.21%**	89.67%	**98.33%**
Salt NaBr–high protein	**95.55%**	95.41%	80.74%	**99.52%**	88.24%	**97.23%**
Salt KCl–low protein	93.15%	**95.08%**	81.91%	**98.38%**	86.91%	**97.04%**
Salt KCl–high protein	95.88%	**99.47%**	90.08%	**98.28%**	93.65%	**99.04%**
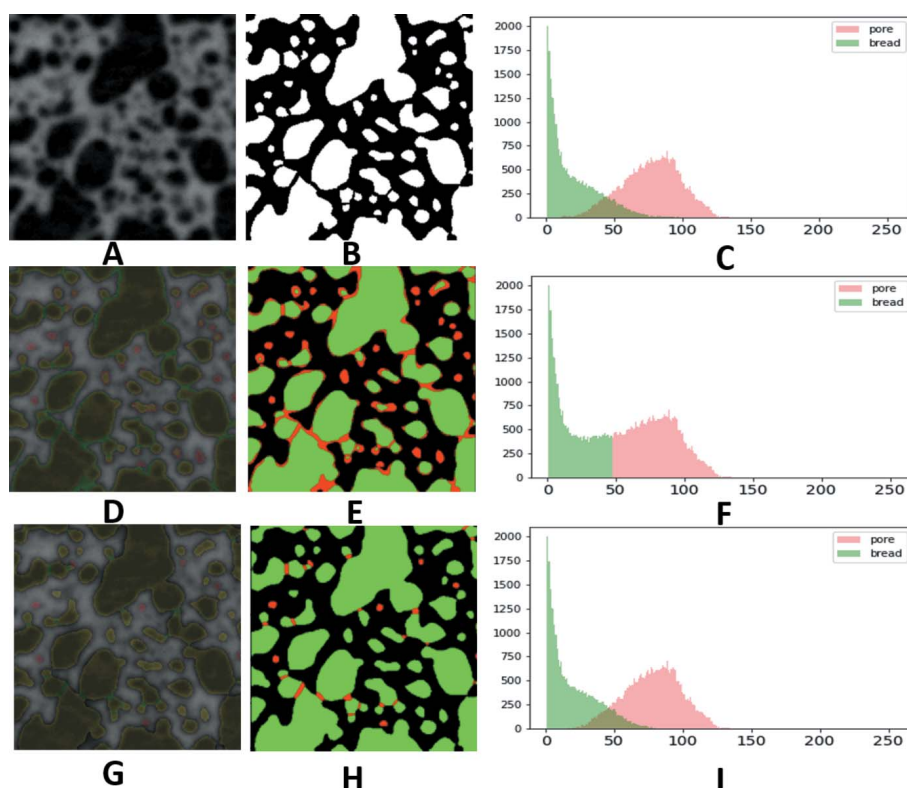
Figure 12
Comparison of using U-Net and Otsu thresholding on the original bread image, showing the overlapping of the pixel intensities histogram. (A) Input grayscale image, (B) ground truth image, (C) histogram of the ground truth of the pores and bread pixels, (D) Otsu threshold image overlaying on the grayscale input image, (E) intersection of segmented image using Otsu threshold and ground truth image, (F) histogram of the prediction of the pores and bread pixels using Otsu threshold, (G) U-net predicted image overlaying on the grayscale input image, (H) intersection of segmented image using U-net and ground truth image, and (I) histogram of the prediction of the pores and bread pixels using U-net.

ground truth. The histogram of both methods in panels C, F and I show that the proposed method is able to detect the overlapping of pixel intensities, while the Otsu threshold method fails to do the same, resulting in more true negative and false positive regions.

The accuracy of the initial segmentation has a flow-on effect in quantitative analyses based on the segmented data. To examine this effect, the output of both above-described methods is processed to extract typical quantitative measurements: average pore diameter and number of pores.

The error in these measurements is calculated using the following formula,

$$\text{Error} \% = \left| \frac{PV - GTV}{GTV} \right|, \quad (3)$$

where GTV is the ground truth value extracted from the manually segmented test data, PV is the predicted value. In simple terms, it is the fraction of correctly classified pixels. The results are reported in Tables 5 and 6 with the best results highlighted in bold.

The U-Net method gives as expected more accurate results for all the parameters. As the Otsu threshold-based segmentation fails to pick up the small pores, and it merges pores together, the number of pores will be underestimated. Table 5 shows that the error in the extracted number of pores from the output of the Otsu threshold segmentation is underestimated in the range between 5% and 16%. On the other hand, our proposed method successfully outputs very accurate results, with smaller errors than the Otsu-based analysis; however, in contrast, it sometimes overestimates the number of pores.

The other parameter we have measured is the average pore diameter, which is important for porosity analysis. As shown in Table 6, the accuracy of our proposed method outperforms automatic thresholding. Incorporating the proposed method will ensure extraction of more accurate parameters for the bread analysis.

4. Conclusions

In this paper, we have proposed a workflow to automatically segment X-ray computed micro-tomography images of bread dough scanned at the Australian Synchrotron using a U-Net model. An important contribution of this paper is to devise a method for the generation of a synthetic annotated/labelled dataset for training of a deep neural network for segmentation of microCT images. As such, we devised and employed an end-to-end deep learning method to automatically segment the porous material and in particular the bread data using synthetically generated images. We showed that having accurate segmentation results improves the accuracy of the statistical metrics extracted from binary images. This approach can be extended to other porous materials such as rocks. Usually, the most challenging part in

Table 5
Comparison of the number of pores between the Otsu threshold and deep learning based segmentation.

Sample name	Ground truth images	Otsu threshold		U-Net method	
	Number of pores	Number of pores	Error %	Number of pores	Error %
No salt–low protein	87	72	16.33%	78	**10.26%**
No salt–high protein	51	47	5.8%	53	**3.96%**
Salt NaCl–low protein	56	49	9.59%	53	**4.99%**
Salt NaCl–high protein	59	52	11.45%	56	**4.87%**
Salt NaBr–low protein	92	79	13.91%	88	**4.52%**
Salt NaBr–high protein	63	56	**10.87%**	72	15.72%
Salt KCl–low protein	76	67	12.11%	83	**9.03%**
Salt KCl–high protein	64	56	11.54%	62	**3.32%**

Table 6
Comparison of average pore diameter between the Otsu threshold and deep learning based segmentation.

Sample name	Ground truth images	Otsu threshold		U-Net method	
	Average pore diameter	Average pore diameter	Error %	Average pore diameter	Error %
No salt–low protein	11.43	12.51	9.61%	12.05	**5.22%**
No salt–high protein	14.90	16.16	9.02%	15.24	**1.64%**
Salt NaCl–low protein	14.79	15.82	7.62%	15.26	**3.25%**
Salt NaCl–high protein	14.18	16	12.77%	15.12	**6.5%**
Salt NaBr–low protein	10.53	12.15	15.28%	10.79	**2.45%**
Salt NaBr–high protein	13.16	16.10	23.46%	12.80	**2.42%**
Salt KCl–low protein	11.64	13.36	14.96%	11.12	**4.53%**
Salt KCl–high protein	13.85	15.61	12.78%	14.29	**3.24%**

evaluating or training any deep learning model is the creation of the ground truth training images. This is even more difficult in the case of porous materials, as it is sometimes difficult to distinguish between the phases. Therefore, the proposed approach eliminates the need to annotate ground truth datasets. The proposed workflow outputs more accurate parameters such as number of pores, average area and average pore diameter as these parameters are sensitive to the accuracy of the binary segmentation. This approach needs further research to evaluate its performance for 3D images. Extending this work to multi-phase materials would be a challenging but worthwhile avenue of further study. Another potential application of using this workflow is to extend it to 4D datasets where it is challenging to manually segment the whole data. One drawback with this approach is that the model needs to be trained on synthetic images that have the features and attributes of the original data. Therefore, this requires studying the characteristics of the original images in order to gain information about the shapes needed to create the corresponding ground truth images. For example, bread dough images usually have circular shapes interconnected with each other, while other porous material such as other foams and metallurgical coke have irregular shapes as these are mostly single materials.

References

Babin, P., Valle, G. D., Dendievel, R., Lassoued, N. & Salvo, L. (2005). *J. Mater. Sci.* **40**, 5867–5873.

Bernklau, I., Lucas, L., Jekle, M. & Becker, T. (2016). *Food. Res. Int.* **89**, 812–819.

Cafarelli, B., Spada, A., Laverse, J., Lampignano, V. & Del Nobile, M. A. (2014*a*). *J. Food Eng.* **124**, 64–71.

Cafarelli, B., Spada, A., Laverse, J., Lampignano, V. & Del Nobile, M. A. (2014*b*). *Food. Res. Int.* **66**, 180–185.

Chakrabarti-Bell, S., Wang, S. & Siddique, K. H. (2014). *Food. Res. Int.* **64**, 587–597.

DeCost, B., Lei, B., Francis, T. & Holm, E. (2019). *Microsc. Microanal.* **25**, 21–29.

Dutta, A., Gupta, A. & Zissermann, A. (2016). *VGG image annotator (VIA)*, http://www.robots.ox.ac.uk/vgg/software/via/.

Falcone, P. M., Baiano, A., Conte, A., Mancini, L., Tromba, G., Zanini, F. & Del Nobile, M. A. (2006). *Adv. Food Nutr. Res.* **51**, 205–263.

Gao, J., Wang, Y., Dong, Z. & Zhou, W. (2018). *Int. J. Food Sci. Technol.* **53**, 858–872.

Guntoro, P. I., Tiu, G., Ghorbani, Y., Lund, C. & Rosenkranz, J. (2019). *Miner. Eng.* **142**, 105882.

Gureyev, T., Nesterets, Ternovski, Thompson, D., Wilkins, S. W., Stevenson, A., Sakellariou, A. & Taylor, J. A. (2011). *Proc. SPIE*, **8141**, 81410B.

Jensen, S., Samanta, S., Chakrabarti-Bell, S., Regenauer-Lieb, K., Siddique, K. & Wang, S. (2014). *J. Microsc.* **256**, 100–110.

Keilegavlen, E., Berge, R., Fumagalli, A., Starnoni, M., Stefansson, I., Varela, J. & Berre, I. (2019). *arXiv*:1908.09869.

Koksel, F., Aritan, S., Strybulevych, A., Page, J. H. & Scanlon, M. G. (2016). *Food. Res. Int.* **80**, 12–18.

Koksel, F. & Scanlon, M. (2016). *Imaging Technologies and Data Processing for Food Engineers*, pp. 129–167. Cham: Springer.

Koksel, F., Strybulevych, A., Aritan, S., Page, J. H. & Scanlon, M. G. (2017a). *J. Cereal Sci.* **78**, 10–18.

Koksel, F., Strybulevych, A., Page, J. H. & Scanlon, M. G. (2017b). *J. Food Eng.* **204**, 1–7.

Lai, Z. & Chen, Q. (2019). *Acta Geotech.* **14**, 1–18.

LeCun, Y., Bengio, Y. & Hinton, G. (2015). *Nature*, **521**, 436–444.

Mayo, S. C., McCann, T., Day, L., Favaro, J., Tuhumury, H., Thompson, D. & Maksimenko, A. (2016). *AIP Conf. Proc.* **1696**, 020006.

Paganin, D., Mayo, S. C., Gureyev, T. E., Miller, P. R. & Wilkins, S. W. (2002). *J. Microsc.* **206**, 33–40.

Park, W. B., Chung, J., Jung, J., Sohn, K., Singh, S. P., Pyo, M., Shin, N. & Sohn, K.-S. (2017). *IUCrJ*, **4**, 486–494.

Petrich, L., Westhoff, D., Feinauer, J., Finegan, D. P., Daemi, S. R., Shearing, P. R. & Schmidt, V. (2017). *Comput. Mater. Sci.* **136**, 297–305.

Ronneberger, O., Fischer, P. & Brox, T. (2015). *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, edited by N. Navab, J. Hornegger, W. M. Wells & A. F. Frangi, pp. 234–241. Cham: Springer International Publishing.

Smal, P., Gouze, P. & Rodriguez, O. (2018). *J. Petrol. Sci. Eng.* **166**, 198–207.

Suresh, A. & Neethirajan, S. (2015). *J. Cereal Sci.* **63**, 81–87.

Trinh, L., Lowe, T., Campbell, G., Withers, P. & Martin, P. J. (2013). *Chem. Eng. Sci.* **101**, 470–477.

Turbin-Orger, A., Babin, P., Boller, E., Chaunier, L., Chiron, H., Della Valle, G., Dendievel, R., Réguerre, A. L. & Salvo, L. (2015). *Soft Matter*, **11**, 3373–3384.

Van Dyck, T., Verboven, P., Herremans, E., Defraeye, T., Van Campenhout, L., Wevers, M., Claes, J. & Nicolaï, B. (2014). *J. Food Eng.* **123**, 67–77.

Wang, S. (2015). *Ab initio bread design: the role of bubbles in doughs and pores in breads.* PhD thesis, The University of Western Australia, Australia.

Wang, S., Karrech, A., Regenauer-Lieb, K. & Chakrabati-Bell, S. (2013). *J. Food Eng.* **116**, 852–861.

Yakubovskiy, P. (2019). *Segmentation models*, https://github.com/qubvel/segmentation_models.